Gottfried Wilhelm Leibniz Universität Hannover
Fakultät für Mathematik und Physik
Institut für Theoretische Physik
AG Quanteninformation

Master's Thesis

Free Probabilistic Subordination for Asymptotic Spectral Distributions of Functions in Random-Matrix-Variables

Tim Frederic Heine

Supervisors

Prof. Dr. Tobias J. Osborne

Prof. Dr. Reinhard F. Werner

November 10, 2023

Abstract

Spectral information about Hermitian operators play the omnipresent role in dealing with quantum systems. The classical, deterministic diagonlization routines for Hamilton operators, not to say smooth interpolations of those, quickly become computationally infeasible as the system size increases. In this project, we artificially add randomness to the Quantum Adiabatic Algorithm (QAA) [FGGS00]. Using this approach, we avoid the numerical problem of computationally expensive diagonalization by transferring the problem to an additive convolution of probability measures. After studying free probability and its subordination formalism [BB07], we extend the Subordination Algorithm to a Free Adiabatic Spectral Gap Estimating algorithm (FASGE). This approach provides a classical stochastic method, which can estimate spectral gaps for randomly influenced adiabatic Hamilton operators for several hundreds of qubits.

Contents

1. Introduction

Quantum computers are the candidates for outperforming commercially available, classical computers of our time. The inherent nature of quantum mechanics provides computationally novel and classically inaccessible tools, which can be appropriately redesigned for operational tasks. As the era of quantum computing begins, many distinct fields of scientific research come together to enable a full-stack approach of building these devices. Just like in classical computer science, the interplay between hardware and software plays the crucial role for a successful technological revolution.

In quantum information, we deal with mathematical structures of quantum mechanics, that lead to solutions to quantum computational problems. Studying and understanding these structures carefully, paves the way for future software applications on quantum computers.

Besides the widely studied, gate-based framework of quantum computation, there exists an operationally equivalent approach, which is known as *adiabatic quantum computing (AQC)* [FGGS00],[AL18]. The underlying quantum algorithm of this model is known as the Quantum Adiabatic Algorithm (QAA). Since realizable quantum systems cannot be prepared without bias, the idealized concept of adiabatic quantum computing is captured within the terminology of *quantum annealing (QA)* [HKL+20]. The circuit model and the concept of quantum annealing were shown to simulate each other with an at most polynomial resource overhead [AL18].

Ever since Farhi et al. [FGGS00] introduced quantum computation based on the adiabatic evolution, much progress dedicated to solving combinatorial optimization problems has been made. In this context, quantum search by measurement [CDF+02], the Quantum Approximate Optimization Algorithm (QAOA) [FGG14], and its generalization to the Quantum Alternating Operator Ansatz (QAO) [HWO+19], are to be mentioned.

Due to the fact that the QAA usually converges slowly compared to the QAOA, modern algorithms hardly make use of the original QAA any more. The main criterion for the QAA to converge, relates to a careful spectral analysis of the involved Hamilton operator. More precisely, the spectral gap between the eigenvalue curves of the energy levels related to the ground state and the first excited state [FGGS00]. The convergence speed of the QAA crucially depends on this spectral gap.

In this project, we revisit the original QAA as proposed in [FGGS00]. Motivated by finding a method which allows us to classically estimate the spectral gap (before any quantum annealing device is switched on), we extensively study a non-commutative probability theory suitable for Hermitian operators: *free probability* [Voi97],[NS06],[MS17]. By applying a randomized, unitary transformation to one (!)[1] of the operators involved in the QAA, we open the mathematical toolbox provided by free probability theory. Subsequently, we conceptually and numerically study this stochastically influenced model of adiabatic quantum computing, which we call *free adiabatic quantum computing (FAQC)*. This approach allows us to develop an algorithm (FASGE), that estimates the minimal spectral gap in the FAQC model. We analyze our algorithm for a theoretical problem instance given by a typical initial Hamiltonian, and its randomized unitary

[1] If we applied a unitary transformation to the entire system (either deterministic or random), we would not change anything, due to the spectral invaraince under unitary transformations.

transformation. Moreover, we study spectral gaps for Grover's search problems [Gro96] within our FAQC model. Thereby, we empirically demonstrate the power of the FASGE to estimate spectral gaps for operators on Hilbert spaces with large, finite dimensions. We present spectral distributions for operators with dimensions corresponding to 100 qubits and possibly more.

1.1. Quantum Mechanics and Spectral Theory

The axiomatic formalism of quantum mechanics states that the dynamics of a quantum system is fully described by the Schrödinger equation. Theoretically speaking, this means to inevitably solve a partial differential equation, where physical states correspond to eigenfunctions, and energy levels (the "spectrum") to eigenvalues of the Hamilton operator. Due to (essential) self-adjointness, the spectrum of the Hamiltonian is compactly supported on the real line. If the Hilbert space is finite-dimensional, the Schrödinger equation simplifies to an eigenvalue problem of a self-ajoint matrix. We consider quantum systems of multi-qubit Hilbert spaces. Throughout, we consider N qubits. Let $\mathcal{H}$ be a Hilbert space of dimension $\dim(\mathcal{H}) = d$, where $d = 2^N$. If $N \geq 20$, diagonalizing operators of this dimension ($d \geq 10^6$) will become practically infeasible, since diagonalization routines include memory storage of arrays of size $10^6 \times 10^6 \cong 10^{12}$. For comparison, the classical singular value decomposition converges in $\mathcal{O}(d^3)$ operations.

1.1.1. Combining Spectra of Distinct Operators?

Consider a quantum system S, described by a Hamiltonian H_S, which is a linear combination of k Hamiltonians

$$H_S = \sum_{i=1}^{k} \alpha_i H_i, \qquad \alpha_i \in \mathbb{R}, \tag{1.1}$$

where the eigenvalue equation reads

$$H_S|\psi\rangle = \lambda|\psi\rangle, \qquad |\psi\rangle \in \mathcal{H}, \quad \lambda \in \mathbb{R}.$$

The ground state energy corresponds to the minimal eigenvalue of H_S. It is now tempting to assume that one could solve this eigenvalue problem by diagonalizing each operator H_i and take the union of these distinct, partial spectra. However, such an approach clearly does not work, unless each H_i would act on a disjoint set of S. Then, $[H_i, H_j] = 0$ for all $i, j \in [k]$, i.e. we would have a complete set of commuting operators (CSCO), thus one common diagonal basis. If the diagonal bases are different, there is no deterministic approach which takes the spectra of two operators A, B (denoted by $\sigma(\cdot)$), $\sigma(A), \sigma(B)$ and computes $\sigma(A+B)$. In the finite-dimensional case, one reason for this obstacle is given by the non-linearity of the determinant function.

Although the quest of uniting several spectra fails in this setting, we stick to the idea that one could develop a quantitative theory which overcomes this barrier of "spectral combinations" by suitably modifying the system.

1.2. Adding Randomness to Spectral Theory

A way out of this is enabled by consulting probability theory. We will artificially include a randomized, unitary basis transformation [Spe19b, Chapter 7] to some parts of the system, in order to lift the problem of uniting spectra to the task of determining a convolution of probability measures.

Although this ansatz of applying a randomized transformation seems relatively straightforward, the probability theory, one must study in this context, is classified differently: random variables are replaced by operators on Hilbert spaces. Notions from probability theory for measurable functions are to be adjusted, in order to respect the non-commutative structure of multiplying operators. Moreover, the induced operator topology generalizes the concept of stochastic convergence, which leads to a different central limit theorem, known as Wigner's semicircle law [Wig67], [Jia21]. Hence, all results about uniting spectra of self-adjoint operators, which we produce along this way, have a probabilistic flavour: we are forced to derive all implications from histograms that display the spectral statistics, rather than from deterministic eigenvalues. The price we pay is inherited in *probabilistic,* rather than *deterministic* spectral distributions.

1.2.1. One Canonical Application: The Quantum Adiabatic Algorithm

Nevertheless, we will be able to produce histograms displaying spectral statistics for operators of the form (1.1). One specific application in quantum computation, where spectral information about linear combinations of operators like (1.1) is of utmost interest, is the QAA [FGGS00]. Hence, we can apply these results directly to QAA Hamiltonians by taking $k = 2$ and choosing $\alpha_1 = 1 - s$, $\alpha_2 = s$, for $s \in [0, 1]$, we are able to study

$$H = (1 - s)H_1 + sH_2', \tag{1.2}$$

for H_2' being unitarily transformed at random. We refer to Chapter 2 and 5 for a detailed analysis of convex combinations of the form (1.2). In the remainder of this introductory survey, we summarize the specific type of probability theory that is needed to deal with operators instead of measurable functions: free probability [Voi97].

1.3. Free Probability and its Applications

Free probability theory links two modern fields of mathematical research: operator algebras and random matrices [MS17], [Voi95]. The term "free" (from "freeness"[2]) refers to a modified version of statistical independence transferred to the framework of linear operators. The main feature that distinguishes free independence from (classical) statistical independence is, that freeness respects non-commutativity [Spe19b, Chapter 1]. Therefore, it naturally enables studying non-commutative random variables that will be represented by operators on Hilbert spaces. In this project, we focus on applications of free probability to random matrix theory which has both an analytic and a combinatorial side [NS06]. It turns out that asymptotic freeness is governed by many random matrix models [MS17]. For our purposes, it is sufficient to transfer the fact that Wigner random matrices (self-adjoint random matrices), deterministic matrices and random unitary matrices are asymptotically freely independent, to numerical computations (Chapter 4, 5).

1.4. Random Matrices and Quantum Information

Applying results from random matrix theory (RMT) to quantum information is not new [CN16]. Naturally, random quantum states and random quantum channels are studied using tools from

[2]This is the correct terminology. It must not to be mixed with *liberty* or *freedom.*

RMT. Several more applications cover entanglement thresholds, the output set of quantum channels, and the minimum output entropy of random channels [CN16].

There even exists a random matrix model for the QAA [MALW05]. In [MALW05], Mitchell et al. study the variation of a "Brody parameter", in order to distinguish spectral regularity (Poissonian) from a chaotic (Wigner type) distribution of normalized nearest neighbor spacings.

However, a model of fusing freeness with the QAA, as we propose in this thesis, where a randomized basis transformation is applied to either the initial or the target Hamiltonian, is - to our knowledge - not yet existent. The key concept in our approach of estimating QAA-spectral gaps relies on adapting the subordination algorithm, proposed by Belinschi et al. (2007) [BB07] to compute the free additive convolution of the empirical spectral measures. We elaborate this carefully in Chapter 4 and 5.

1.5. A Guide to Read

We assume that the reader is familiar with both foundations of quantum mechanics[3] and basic notions from quantum information[4]. After introducing the required mathematical framework, we present relevant notions from random matrices intersected with free probability theory in Chapter 3. Although mathematically advanced, these two Chapters do not present any novel insights, but nevertheless build the compactified journey to understand the combination between asymptotic freeness and adiabatic quantum computation, which is necessary for Chapter 4 and 5. Within these two chapters, we carefully study the foundational algorithm that allows the numerical computation of adding spectra of random matrices: subordination iteration. We demonstrate its working principle and provide numerical parameters that are suitable for our computations. Based on the subordination algorithm, we develop an algorithm that computes spectral gaps of linearly combined, self-adjoint deterministic and random matrices in Chapter 5: the FASGE. More specifically, it allows to determine spectral gaps of two linearly combined deterministic matrices A and B, where the basis of B is unitarily rotated at random compared to the basis of A. This model of a randomized addition of distinct spectra is transferred to the adiabatic quantum computing setup and declared as "free adiabatic quantum computation". We discuss its applicability and possible thresholds of this kind of model in Chapter 5 and 6.

[3] see e.g. [Sch07], [Sch08]
[4] e.g. [Wil17a]

2. Basic Mathematical Background

We begin this chapter with reviewing the functional analytic formalism, in which free probability is classified. Subsequently, we recall the algebraic formalism of classical probability theory and present the transition to the non-commutative generalization. Finally, we combine these two fields by the functional calculus on so called C^* probability spaces.

2.1. Spectral Theory for Bounded Normal Operators

Free probability mainly interconnects operator algebras and random matrices. Since probability measures in the algebra of random matrices both have an operator theoretic and a stochastic viewpoint, we shall recall deterministic, functional analytic foundations that will be generalized to a stochastic functional calculus later on. Two foundational concepts of functional analysis are the theorem of Stone-Weierstrass and the Riesz representation theorem. The theorem by Stone-Weierstrass enables approximating continuous functions via polynomials. The simplest version, known as Weierstrass approximation theorem, states that polynomials are dense in $(C[a,b], \|\cdot\|_\infty)$ for $a, b \in \mathbb{R}, a < b$ [Wer18, Theorem I.2.11]. Throughout, X denotes a Banach space, $\mathcal{H}$ a finite, $d-$dimensional Hilbert space over $\mathbb{C}$, and $\mathcal{B}(\cdot)$, denotes the set of bounded linear operators on X, respectively $\mathcal{H}$. Probability measures are positive and normalized. If we define those on the compactly supported, real-valued spectrum of Hermitian operators, we need a measure that respects both the compactness and positivity structure. This leads to [Def. C.1][Mai17]

Definition 2.1. *A positive measure $\mu : \mathcal{B}(X) \to [0, \infty]$ is called*

- ***inner regular,** if*
$$\mu(B) = \sup_{K \overset{compact}{\subset} B} \mu(K), \qquad B \in \mathcal{B}(X).$$
- ***locally finite** (or Borel measure), if each $x \in X$ has an open neighborhood U, for which $\mu(U) < \infty$ holds true.*
- ***Radon measure,** if it is both inner regular and locally finite.*

It is mentioned that the Hausdorff property enables the separation of points, both for the Weierstrass theorem and for well-defining open neighborhoods for Borel measures. Due to the fact that finding joint probability distributions is related to determining a unique probability measure, we review Riesz' representation theorem in the vector space of measurable functions, following [Mai17, Corollary C.5]

Theorem 2.2. *Let X be a Hausdorff topological vector space. If $I : C(X) \to \mathbb{C}$ is a positive linear functional, then there exists a unique and finite Radon measure $\mu : \mathcal{B}(X) \to [0, \infty)$, such that*
$$I[f] = \int_X f(x)\, d\mu(x), \qquad f \in C(X). \tag{2.1}$$

We proceed by collecting some elementary results about spectral theory, following [Hal13], [Mor17], and [Wer18].

Definition 2.3. *Let $A \in \mathcal{B}(\mathcal{H})$. The **spectrum** $\sigma(A)$ is defined as:*

$$\sigma(A) := \{z \in \mathbb{C} | \, (z\mathbb{1}_{\mathcal{H}} - A) \;\; \textit{is not invertible in} \;\; \mathcal{B}(\mathcal{H})\}.$$

Before stating the spectral theorem for bounded self-adjoint and normal operators, we recall the notion of a spectral measure, or - more commonly used in the quantum mechanics literature - a projection-valued measure.

Definition 2.4 ([Hal13], Def. 7.10)**.** *Let Ω be a set and $\mathcal{F}$ a $\sigma-$algebra on X. A map $\mu : \mathcal{F} \to \mathcal{B}(\mathcal{H})$ is called **projection-valued measure,** if the following conditions hold true:*

1. *for each $E \in \mathcal{F}$, $\mu(E)$ is an orthogonal projection,*
2. *$\mu(\emptyset) = 0$, and $\mu(\Omega) = \mathbb{1}_{\mathcal{H}}$,*
3. *if $E_1, E_2, E_3, ...$ are disjoint in $\mathcal{F}$, then for all $v \in \mathcal{H}$:*

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) v = \sum_{j=1}^{\infty} \mu(E_j) v,$$

   (convergence in norm topology on $\mathcal{H}$)

4. *for all $E_1, E_2 \in \mathcal{F}$, we have $\mu(E_1 \cap E_2) = \mu(E_1)\mu(E_2)$.*

Consequently, $\mu(E_1)\mu(E_2)$ is an orthogonal projection, if E_1 and E_2 are disjoint. For $\psi \in \mathcal{H}$, an ordinary, positive real-valued measure can be recovered from a spectral measure by taking $\mu_\psi = \langle \psi, \mu(E)\psi \rangle$, $\forall \, E \in \Omega$ [Hal13]. We recall the definition of a spectral subspace (see [Hal13], Def. 7.14)

Definition 2.5. *Let $A \in \mathcal{B}(\mathcal{H})$ and μ^A be the spectral measure associated to A. Note that μ^A can be extended to an ordinary measure by defining $\mu^A(\mathbb{R} \setminus \sigma(A)) = 0$. Then, for any Borel set $E \subset \mathbb{R}$, the spectral subspace $V_E \subset \mathcal{H}$ is defined by $V_E = ran\,(\mu^A(E))$.*

If we consider $d \times d$ matrices ($d := 2^N$) over $\mathbb{C}$ as elements in $\mathcal{B}(\mathcal{H})$ over the underlying, finite-dimensional Hilbert space $\mathcal{H} = \mathbb{C}^N$, we deal with the the simplest case of bounded operators. For our purpose, it is hence sufficient to work with spectral theorems on $\mathcal{B}(\mathcal{H})$. The self-adjoint version is presented as follows (compare with [Hal13], Thm 7.12).

Theorem 2.6 (Spectral Theorem for bounded self-adjoint operators)**.** *Let $A \in \mathcal{B}(\mathcal{H})$ be self-adjoint. Then, there exists a unique spectral measure μ^A on $\sigma(A)$ such that*

$$A = \int\limits_{\lambda \in \sigma(A)} \lambda \, d\mu^A(\lambda). \tag{2.2}$$

By Definition 2.4, μ^A takes values in projections on $\mathcal{H}$ and since $\sigma(A)$ of A is bounded, the function $f(\lambda) := \lambda$ is bounded on $\sigma(A)$.

Selfadjoint and unitary operators are both subclasses of normal operators. The spectrum of a selfadjoint operator is compactly supported on $\mathbb{R}$ and of a unitary operator on the unit disk $\mathbb{S}^1 \subset \mathbb{C}$. For the sake of completeness, we state the spectral decomposition theorem for bounded normal operators as follows (see e.g. [Mor17], Ch. 8.4.1).

Theorem 2.7. *Let $A \in \mathcal{B}(\mathcal{H})$ be normal, i.e. $AA^* = A^*A$ and let*

$$\lambda : \mathbb{R}^2 \to \mathbb{C}, \qquad (x, y) \mapsto \lambda(x, y) \equiv \lambda := x + iy$$

Moreover, let $\mathfrak{B}(\mathbb{R}^2)$ be the Borel σ-algebra over $\mathbb{R}^2$.
Then, the following holds true.

(a) There exists unique bounded spectral measure $\mu^A : \mathfrak{B}(\mathbb{R}^2) \to \mathcal{B}(\mathcal{H})$ such that

$$A = \int_{\mathrm{supp}(\mu^A)} \lambda \, d\mu^A(x, y). \tag{2.3}$$

(b) μ^A is concentrated on its support and supp $(\mu^A) = \sigma(A)$.

(c) If f is a bounded, measurable function on $\sigma(A)$, then the operator $\int_{\sigma(A)} f(x, y) \, d\mu^A(x, y)$ commutes with every operator in $\mathcal{B}(\mathcal{H})$ that commutes with A and A^*.

The spectral decomposition theorem allows to establish a functional calculus, i.e. computing functions of operators [Hal13]. If $f : \sigma(A) \to \mathbb{C}$ is a bounded measurable function for some $A \in \mathcal{B}(\mathcal{H})$, we can evaluate $f(A)$ as

$$f(A) = \int_{\sigma(A)} f(\lambda) \, d\mu^A(\lambda). \tag{2.4}$$

If we impose more structure on f, rather than $f \in \mathcal{M}_b(\sigma(A))$, the functional calculus can be extended accordingly. For instance, if f is a holomorphic function, many results from complex analysis are transferred to functions of operator-valued variables [Sch23]. Let $R(\lambda, A) := (\lambda \mathbb{1}_{\mathcal{H}} - A)^{-1}$ be the resolvent operator, where $()^{-1}$ denotes the Banach space inversion. If γ is a piecemeal C^1 curve, closed around $\sigma(A)$ and f is holomorphic on an open domain $D \subset \mathbb{C}$, we can define the Cauchy integral formula in A via

$$f(A) = \frac{1}{2\pi i} \int_{\gamma} f(\lambda) R(\lambda, A) d\lambda \in \mathcal{B}(\mathcal{H}). \tag{2.5}$$

Note that $\mathcal{H}$ being a Banach space is sufficient, for the above to be valid. Combining a holomorphic functional calculus and the empirical spectral measure of a random Hermitian matrix enables us to rigorously define the Cauchy transform in Chapter 3. Some aspects about functional calculi in C^* probability spaces are presented at the end of this chapter.

2.2. The Quantum Adiabatic Evolution

Transferring the rich field of free probability to a stochastic spectral gap analysis for adiabatic quantum computing, is the key for this project. In this part, we briefly describe the quantum adiabatic evolution applied to quantum computation, as it was proposed by Farhi et al. [FGGS00]. The adiabatic evolution is a smooth interpolation of Hamilton operators, based on the adiabatic theorem of quantum mechanics. We refer to [AL18] for a general, comprehensive review about adiabatic quantum computing.

The main object of interest in adiabatic quantum computing, is a convex combination of two self-adjoint operators. Throughout, we denote these operators simply by A and B and define the adiabatic Hamilton operator as $H(\cdot) \in C^2([0,1], \mathcal{B}(\mathcal{H}))$

$$H(s) := (1-s)A + sB, \qquad s \in [0,1]. \tag{2.6}$$

The C^2 requirement ensures the well-definition of minimizing $s \mapsto H(s)$, i.e. $\frac{d}{ds}H(s) = B - A$, $\frac{d^2}{ds^2}H(s) = 0$. Here, $s \in [0,1]$ corresponds to a rescaled time parameter, $s := \frac{t}{T}$, $0 \leq t \leq T$, where $\hat{H}(t)$ is the Hamilton operator [FGGS00]

$$\frac{d}{dt}|\psi(t)\rangle = -i\hat{H}(t)|\psi(t)\rangle, \tag{2.7}$$

such that

$$\hat{H}(t) = H\left(\frac{t}{T}\right) = H(s). \tag{2.8}$$

Solving the time-dependent Schrödinger equation (2.7) for the slowly varying Hamiltonian $H(s)$ in terms of the unitary evolution operator $U_T(s)$, it holds that [Sim17]

$$\frac{d}{ds}U_T(s) = -iTH(s)U_T(s), \qquad s \in [0,1], \qquad U_T(0) = \mathbb{1}. \tag{2.9}$$

We recall the quantum adiabatic theorem in T.Kato's formulation, following [Sim17, Theorem 17.2].

Theorem 2.8 (Adiabatic Theorem)**.** *Let $H(s)$ be a C^2-family of bounded self-adjoint operators on a complex, separable Hilbert space $\mathcal{H}$. Suppose, there is a C^2-function, $\lambda(s)$ so that for all s, $\lambda(s)$ is an isolated point in the spectrum of $H(s)$ and so that*

$$g_{\min} \equiv \inf_{s\in[0,1]} \operatorname{dist}\left(\lambda(s), \sigma(H(s)) \setminus \{\lambda(s)\}\right) > 0. \tag{2.10}$$

Let $P(s)$ be the spectral projection onto the eigenspace, associated to $\lambda(s)$ of $H(s)$. Then,

$$\lim_{T\to\infty} (\mathbb{1} - P(s))U_T(s)P(0) = 0 \tag{2.11}$$

uniformly in $s \in [0,1]$.

The above relation is equivalent to

$$\lim_{T\to\infty} U_T(s)P(0) = \lim_{T\to\infty} P(s)U_T(s)P(0).$$

Let $\psi_0 \in \operatorname{ran}(P(0))$. In the large-$T$ limit, this means that $U_T(s)\psi_0 \approx P(s)U_T(s)\psi_0$ [Sim17]. In words, the eigenspace during the adiabatic evolution stays sufficiently close to the ground state for large values of T.

In the QAA, the initial operator A, often called "initial Hamiltonian" typically encodes an easy to prepare ground state, that leaves N qubits in uniform superposition. By tuning the renormalized time parameter s sufficiently slowly from 0 to 1, the adiabatic Hamiltonian evolves to the operator B - known as "problem Hamiltonian" or "target Hamiltonian", whose ground state encodes the solution to the given problem. The minimal spectral gap $g_{\min}$ [FGGS00, Chapter 2.1],

$$g_{\min} := \min_{s\in[0,1]} (\lambda_1(s) - \lambda_0(s)) \tag{2.12}$$

is the crucial, performance-indicating quantity of interest. Here, $\lambda_0(s)$ denotes the eigenvalue curve corresponding to the ground state energy of the operator, where $\lambda_1(s)$ refers to the eigenvalue curve, that corresponds to the energy level of the first excited state. The run time of the adiabatic algorithm is proportional to $g_{\min}^{-2}$ [FGGS00]. In order to solve combinatorial optimization problems within this quantum computational framework, one needs a quantum annealing device [RSDC23]. Quantum annealing is usually referred to the real-world application of the idealized model of adiabatic quantum computing. We stick to a mathematical analysis of (2.6) and refer for discussions regarding experimental and technical realizations of quantum annealers to [RSDC23], [HKL+20].

2.3. Classical Probability Theory

Free probability theory is based on a generalized notion of statistical independence, which is called free independence. The key difference comes with non-commutativity of random variables. In order to compare the notions that are introduced in the free probability chapter, we briefly recall some results from classical probability theory. These notions are elementary and can be found in any textbook on probability theory. We mainly follow part I in [Geo15].

Definition 2.9. *(Probability Space, Random Variable) Let Ω be a set ($\Omega \neq \emptyset$), $\mathcal{F}$ be a σ-algebra of events over Ω and let $\mathbb{P} : \mathcal{F} \to [0,1]$ be a normalized probability measure. The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called probability space. The pair $(\Omega, \mathcal{F})$ is called measurable space of events. A measurable function $f : (\Omega_1, \mathcal{F}_1) \to (\Omega_2, \mathcal{F}_2)$ is called **random variable**, if for any $A \in \mathcal{F}_2 \implies f^{-1}(A) \in \mathcal{F}_1$.*

Here, the requirement of $\mathbb{P}$ being a normalized measure especially means

$$\mathbb{P}(\Omega) = 1, \qquad \mathbb{P}\left(\bigcup_{i\geq 1} A_i\right) = \sum_{i\geq 1} \mathbb{P}(A_i), \qquad \text{for any } A_i \cap A_j = \emptyset, \quad i \neq j. \tag{2.13}$$

We consider $\mathbb{K} = \{\mathbb{R}, \mathbb{C}\}$ valued random variables, i.e. $\Omega = \mathbb{K}$ and $\mathcal{F} = \mathfrak{B}(\mathbb{K})$. The random variables are hence the bounded, $\mathbb{K}-$ valued measurable functions, which are denoted by $\mathcal{M}(\mathbb{K})$. Depending on the probability measure, one can compute for each random variable between measurable spaces the expectation value. We understand the expectation value as a map. More precisely, a linear functional from the algebra of bounded measurable functions to the ground field, i.e.

$$\mathbb{E}[\cdot] \in \mathcal{L}(\mathcal{M}(\mathbb{K}), \mathbb{K}), \qquad \mathbb{E}[X] := \int_{\omega\in\Omega} X(\omega) d\mathbb{P}(\omega). \tag{2.14}$$

Note that for discrete probability spaces, we canonically choose $\mathbb{P}$ to be the (normalized) counting measure. Therefore, equation (2.14) remains true, even in the case of discrete probability spaces. We emphasize that $\mathcal{M}(\mathbb{K})$ is a unital algebra and $\mathbb{E} : \mathcal{M}(\mathbb{K}) \to \mathbb{K}$ its corresponding unital linear functional. The axioms of a unital algebra are quickly verified by taking into account that measurability is closed under point-wise multiplication. Due to normalization of $\mathbb{P}$, unitality of the functional is obtained via

$$\mathbb{E}[1_{\mathcal{M}(\mathbb{K})}] = \int_{\omega\in\Omega} 1 \, d\mathbb{P}(\omega) = 1. \tag{2.15}$$

The free probabilistic definition of non-commutative probability spaces is formulated in the language of unital (C^*-) algebras. By taking the pair $(\mathcal{M}(\mathbb{K}), \mathbb{E})$, the above shows that we can construct classical probability theory in this algebraic framework, as well.

In the following, we restrict to probability measures with continuous densities. Standard examples are the Gaussian, Poisson, and the uniform distribution (on Borel measurable sets). If the distribution has a density $\rho \in C^1(\mathbb{R})$, we will use the convention of the Riemann-Stieltjes integral, i.e.

$$d\mathbb{P} \equiv d\rho(t) := \rho'(t)dt \tag{2.16}$$

One of the motivating questions for this thesis is, whether and how distinct spectra of random matrices add up. The key concept in the addition of random matrices is computing a joint probability distribution in some suitable probability space. Abstractly speaking, this relates to computing the (additive) convolution of the underlying probability measures. In classical probability, this corresponds to determining the joint distribution of statistically independent random variables.

Definition 2.10 (Statistical Independence, [Geo15])**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Two events $E_1, E_2 \in \mathcal{F}$ are said to be statistically independent with respect to $\mathbb{P}$, if $\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1)\mathbb{P}(E_2)$.*

A direct corollary from the definition of statistical independence is that the expectation functional for independent random variables $X, Y \in \mathcal{M}_b(\mathbb{K})$ factorizes with respect to $\mathbb{E}$, i.e. for $m, n \in \mathbb{N}$,

$$\mathbb{E}[X^m Y^n] = \mathbb{E}[X^m]\mathbb{E}[Y^n]. \tag{2.17}$$

Instead of the measure-based notion of statistical independence from Definition 2.10, we will review the construction of a state-based notion of independence, as it is written in (2.17). The joint distribution of a collection of random variables $X_1, ..., X_n \in L^\infty(\Omega, \mathbb{P})$ is the push forward of the map [Mai17, Chapter I.1.1]:

$$X : \Omega \to \mathbb{R}^n, \qquad \omega \mapsto (X_1(\omega), ..., X_n(\omega)),$$

where the probability measure μ_X is defined for any Borel set $B \subset \mathbb{R}^n$ as:

$$\mu_X(B) := \mathbb{P}\left(\{\omega \in \Omega | (X_1(\omega), ..., X_n(\omega)) \in B\}\right). \tag{2.18}$$

From now on, we focus on the algebraic formalism of probability theory. Joint distributions of random variables are given by the collection of all their moments with respect to the expectation functional. This implicitly means, that we can characterize the probability measure in 2.18, without explicitly referring to $\mathbb{P}$. By consulting the Stone-Weierstrass approximation theorem and Riesz' representation theorem for Radon measures, we write the information about the joint distribution of a polynomial P in the X_i, $i \in [n]$ [Mai17, Example I.1.11]:

$$\begin{aligned}\mathbb{E}[P(X_1, ..., X_n)] &= \int_\Omega P(X_1(\omega), .., X_n(\omega))\, d\mathbb{P}(\omega) \\ &= \int_{\mathbb{R}^n} P(x_1, ..., x_n)\, d\mu_X(x_1, ..., x_n).\end{aligned}$$

Consequently, the information about this joint distribution is encoded into the linear functional $\mathbb{E}$:

$$\mathbb{E} \in \mathcal{L}(\mathbb{C}[x_1, ..., x_n], \mathbb{C}), \qquad P(X_1, ..., X_n) \mapsto \mathbb{E}[P(X_1, ..., X_n)] \tag{2.19}$$

This correspondence justifies describing joint distributions in (non-commuting) random variables by the collection of their moments with respect to the state. Hence, we can formulate probability theory in the framework of unital C^* algebras. Statistical independence is encoded into the state on these algebras without explicitly regarding the underlying probability measure. Both the transition from a classical to a non-commutative probability theory and functional analytic tools in C^* probability spaces, are described in the remainder of this chapter.

2.4. Non-Commutative ∗-Probability Spaces

In free probability theory, random variables are operators in unital algebras, which usually do not commute. This non-commutativity structure is inherited in the notion of free independence. In order to define freeness properly, we present the minimal algebraic structure needed to formulate free probability [Spe19b], [NS06].

Definition 2.11 (Non-Commutative *-Probability Space)**.** *Let $\mathcal{A}$ be a unital *-algebra over $\mathbb{C}$, i.e. an algebra $\mathcal{A}$ equipped with an antilinear *-operation $a \mapsto a^* \in \mathcal{A}$, which is involutive $((a^*)^* = a)$ and satisfies $(ab)^* = b^*a^*$. Furthermore, let $\varphi : \mathcal{A} \to \mathbb{C}$ be a unital linear functional, i.e. $\varphi(1_{\mathcal{A}}) = 1$. Then, the pair $(\mathcal{A}, \varphi)$ is called* ***non-commutative *-probability space.*** *The elements $a \in \mathcal{A}$ are called* ***non-commutative random variables.*** *Moreover, $(\mathcal{A}, \varphi)$ is called* ***tracial,*** *if $\varphi(ab) = \varphi(ba)$ for all $a, b \in \mathcal{A}$. The functional φ is called a* ***state,*** *if φ is positive, i.e. $\varphi(a^*a) \geq 0$ and a state is called* ***faithful,*** *if additionally $\varphi(a^*a) = 0 \Rightarrow a = 0 \ \forall a \in \mathcal{A}$.*

We abbreviate a non-commutative probability spaces with NCPS. Note that the *-property could be omitted both in the abstract definition of $(\mathcal{A}, \varphi)$ and in the traciality property. Since the C^* algebra of bounded Hermitian operators will be the main object in this thesis, we just define the *-probability space and leave the concept of non-commutative probability spaces - without $*-$operation - for more abstract purposes.

From now on, we suppress the non-commutativity property, i.e. we implicitly assume that two elements $a, b \in \mathcal{A}$ do not commute and thereby simply denote $a, b \in \mathcal{A}$ as **random variables.**

The subsequent proposition is a simple yet extremely relevant result in random matrix theory. Properties of the underlying *-probability space $(\mathcal{A}, \varphi)$ are preserved, when one considers the matrix algebra over that particular space. This result follows Exercise 1.23 of [NS06].

Proposition 2.12. *Let $(\mathcal{A}, \varphi)$ be a *-probability space and let $M_d(\mathcal{A})$ be the matrix algebra over $\mathcal{A}$, i.e.*

$$M_d(\mathcal{A}) := \{(a_{ij})_{i,j=1}^d, a_{ij} \in \mathcal{A} \text{ for all } 1 \leq i, j \leq d\}$$

*The *-operation is canonically defined as the component wise *-operation and transposition that is induced by the underlying algebra $\mathcal{A}, (a_{ij})^* := a_{ji}^*$. Then, one considers the unital linear functional $\varphi_d : M_d(\mathbb{C}) \otimes \mathcal{A} \to \mathbb{C}$, given by:*

$$\varphi_d(A) = \frac{1}{d} \sum_{i=1}^{d} \varphi(a_{ii}) \tag{2.20}$$

for $A \in M_d(\mathcal{A})$ and $A = (a_{ij})_{i,j=1}^d$ for all $a_{ij} \in \mathcal{A}$, where $M_d(\mathcal{A}) \cong M_d(\mathbb{C}) \otimes \mathcal{A}$. Under this isomorphism, $\varphi_d = \operatorname{tr} \otimes \varphi$. Then,

1. *$(M_d(\mathcal{A}), \varphi_d)$ is a *-probability space.*
2. *If $(\mathcal{A}, \varphi)$ is tracial, then so is $(M_d(\mathcal{A}), \varphi_d)$.*

3. If $\varphi : \mathcal{A} \to \mathbb{C}$ is faithful, then so is $\varphi_d : M_d(\mathcal{A}) \to \mathbb{C}$.

Proof. The proof of all three properties follows immediately by the definition of a *-probability space. The linearity of φ_d is preserved due to the linearity of both φ (by definition) and the trace functional tr. □

2.5. Functional Calculus on C^* Probability Spaces

In light of Section 2.1, this last part is dedicated to generalizing functional calculus from $\mathcal{B}(\mathcal{H})$ to stochastic functional calculus on abstract C^* algebras. We review some results of probability theory with C^* algebras, closely following Chapter 3 in [NS06]. The probabilistic flavor relates to φ being unital and a *state;* the rest is basic functional calculus with C^* algebras. Clearly, $\mathcal{B}(\mathcal{H})$ is a C^* algebra [Hei20, Corollary 2.7.3].

Definition 2.13. *[Mai17, Def. I.1.7] A C^* probability space is a pair $(\mathcal{A}, \varphi)$ consisting of a unital C^* algebra $\mathcal{A}$ and a state φ on $\mathcal{A}$.*

Topological notions naturally go over to the norm topology. If the underlying algebra $\mathcal{A}$ is weakened to a von Neumann algebra, one can consider so called W^* probability spaces, where topological structures are naturally induced by the weak$-*$ topology. Since random matrices are closed within a specific norm topology, which is composed by the essential supremum of the operator norm, our analysis shall be restricted to C^* probability spaces. Within this abstract formalism, the algebra of complex square matrices equipped with the normalized trace functional $(M_d(\mathbb{C}), \mathrm{tr}_d)$ is mentioned as an example for a C^* probability space. Moreover, the trace functional tr_d is a faithful, tracial state [Mai17].

Since φ is a *state* on $\mathcal{A}$, φ is positive, i.e. $\varphi(a^*a) \geq 0$, for all $a \in \mathcal{A}$ [Mai17]. Positive elements $a_+ \in \mathcal{A}^+$ are defined as

$$\mathcal{A}^+ := \{a_+ \in \mathcal{A} | a_+ = a_+^*, \text{ and } \sigma(a_+) \subset [0, \infty)\}.$$

Thus, we have

Proposition 2.14. *This set of positive elements forms a convex cone, i.e. it is closed under linear combinations with positive coefficients. Consequently, for $s \in [0,1]$ and $\tilde{a} \in \mathcal{A}$, defined as*

$$\tilde{a}(s) := (1-s)a + sb, \qquad a, b \in \mathcal{A}^+, \tag{2.21}$$

we have $\tilde{a} \in \mathcal{A}^+$.

Proof. Let $a, b \in \mathcal{A}^+, \alpha, \beta \in \mathbb{R}^+$ and define $\tilde{a} := \alpha a + \beta b$. $\tilde{a}$ is self-adjoint due to

$$\tilde{a}^* = (\alpha a + \beta b)^* \overset{\alpha,\beta\in\mathbb{R}}{=} \alpha a^* + \beta b^* \overset{a,b \text{ s.a.}}{=} \alpha a + \beta b = \tilde{a}.$$

Thus, $\sigma(\tilde{a}) \subset \mathbb{R}$. We apply a standard estimate for the spectral radius: $r(a) < \|a\|_{\mathrm{op}}$ (see e.g. [Bau22, Lemma 23]). Then,

$$r(\tilde{a}) < \|\tilde{a}\|_{\mathrm{op}} = \|\alpha a + \beta b\|_{\mathrm{op}} \overset{\Delta-\text{ineq.}}{\leq} |\alpha|\|a\|_{\mathrm{op}} + |\beta|\|b\|_{\mathrm{op}} \overset{\alpha,\beta\geq 0}{=} \alpha\|a\|_{\mathrm{op}} + \beta\|b\|_{\mathrm{op}} \overset{a,b\in\mathcal{A}^+}{<} \infty,$$

i.e. $r(\tilde{a}) < \infty$. If $\tilde{a}$ is positive, then there exists $\tilde{b}$ such that $\tilde{a} = \tilde{b}\tilde{b}^*$. We make the ansatz $\tilde{b} := x + iy$. Then, $\tilde{b}\tilde{b}^* = x^2 + y^2$, from which the choice $x := \sqrt{\alpha a}$ and $y := \sqrt{\beta b}$ shows that there exists $\tilde{b}$ such that $\tilde{a} = \tilde{b}\tilde{b}^*$ holds true. Note that $\sqrt{\cdot} : \sigma(aa^*) \to \mathbb{R}^+$ is well defined by positivity of a and b, respectively. The convex combination in (2.21) is a direct application of the above relation with $\alpha := 1 - s$ and $\beta := s$. □

In case of $\mathcal{A} = \mathcal{B}(\mathcal{H})$, the previous proposition ensures that convex combinations of positive, bounded linear operators remain positive. For our purposes, the quantity $\tilde{A}(\cdot) \in C^2([0,1], \mathcal{B}(\mathcal{H}))$

$$\tilde{A}(s) := (1-s)A + s\,\Phi_U(B), \qquad A, B \in \mathcal{B}(\mathcal{H}) \qquad s \in [0,1], \tag{2.22}$$

is of particular interest, where $\Phi_U(\cdot) := U(\cdot)U^*$ indicates a unitary transformation. Since such a transformation does not change the spectrum of B, i.e.

$$B' := \Phi_U(B) \quad \Longrightarrow \quad \sigma(B') = \sigma(B), \qquad \text{for } B, B' \in \mathcal{B}(\mathcal{H}), \tag{2.23}$$

Proposition 2.14 remains true, if

$$\tilde{A}(s) = (1-s)A + sB'. \tag{2.24}$$

We will return to equation (2.24) in Chapter 5. Note that the spectral invariance of B under Φ_U holds true, no matter if U is a *deterministic,* or a *random* unitary transformation.

3. Random Matrices and Free Probability Theory

In this chapter, we introduce the central object for this project: random matrices. The theory of random matrices started in the beginning of the 20'th century contemplating fields like number theory, condensed matter physics, wireless communications, and most importantly, quantum mechanics [EW13] [TV+04]. Its relevance to quantum physics was established in the 1950s, when Wigner modelled large, heavy nuclei, using self-adjoint random matrices [EW13]. Besides Wigner in physics, contributions by both Wishart (1928), and Marcenko, Pastur (1967) in statistics, paved the way for this nowadays rich and active field of research [Ize08] [TV+04].

We aim to study functions, that take random matrices as indeterminates. We begin with sums, convex and linear combinations. Specifically, we study spectra of linear combinations of random matrices in the asymptotic limit. That is, starting with a finite dimensional $d \times d$ matrix, we consider spectral distributions for $d \to \infty$.

It turns out that *Free Probability* [Voi97] yields the mathematical toolbox, needed to answer questions of this kind. At first, we study the relevant notions of random Hermitian matrices linked to free probability. Second, we focus on the calculus of *free variables* in non-commutative C^* probability spaces. Finally, we study the free additive convolution and asymptotic freeness between deterministic and unitary random matrices.

3.1. Basic Properties

We start by defining what the algebra of random matrices actually is. Roughly speaking, a random matrix is a Hermitian symmetric, complex valued square matrix, say of dimension $d \in \mathbb{N}$, where each entry is a random number. However, this rather naive picture is incomplete. First, we must require that all (scalar valued) random variables in each of the d^2 matrix entries are drawn from the same probability distribution. Secondly, the assumption that the random variables are independent and identically distributed (i.i.d) - up to Hermitian symmetry - reduces the degrees of freedom and provides statistical independence between matrices. Finally, we provide the adequate norm induced by the C^* algebra of random matrices.

We introduce random matrices rigorously, following [NS06, Def. 22.6]

Definition 3.1. *Let $M_d(\mathbb{C})$ be the algebra of $d \times d$ complex matrices and let $L^{\infty-}(\Omega, \mathbb{P})$ be the algebra over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, defined as*

$$L^{\infty-}(\Omega, \mathbb{P}) := \bigcap_{1 \leq p < \infty} L^p(\Omega, \mathbb{P}).$$

Then, the ****-probability space of** $d \times d$ **random matrices** is given by*

$$(M_d(L^{\infty-}(\Omega, \mathbb{P})), \mathrm{tr} \otimes \mathbb{E}),$$

where $M_d(L^{\infty-}(\Omega, \mathbb{P})) \simeq M_d(\mathbb{C}) \otimes L^{\infty-}(\Omega, \mathbb{P})$. The linear functional $\varphi_d \equiv \mathrm{tr} \otimes \mathbb{E}$ is induced by the normalized trace functional on $M_d(\mathbb{C})$ together with the (expectation) functional $\mathbb{E}$ on

$(\Omega, \mathcal{F}, \mathbb{P})$, *i.e.*

$$\varphi_d(A) \equiv (\mathrm{tr} \otimes \mathbb{E})(A) = \int_{\omega \in \Omega} \mathrm{tr}(A(\omega))\, d\mathbb{P}(\omega). \tag{3.1}$$

for any ***random matrix*** $A \in M_d(L^{\infty-}(\Omega, \mathbb{P}))$. *The above, abstract formulation can be concretely written as:*

$$\varphi_d(A) \equiv (\mathrm{tr} \otimes \mathbb{E})(A) = \mathbb{E}[\mathrm{tr}(A)] = \frac{1}{d} \sum_{i=1}^{d} \mathbb{E}[a_{ii}], \tag{3.2}$$

where $A = (a_{ij})_{i,j=1}^{d}$ *and* $a_{ij} \in L^{\infty-}(\Omega, \mathbb{P})$.

Hence, random matrices are constructed by taking the algebraic tensor product between a set of matrices and a set of probability distributions. This means for self-adjoint random matrices: deterministic, square matrices over $\mathbb{C}$ with $L^{\infty-}(\mathbb{C}, \mathbb{P})$, because all moments of the distribution are assumed to exist. A simple analysis provides the number of independent elements of both real symmetric, and complex self-adjoint random matrices. In the real-valued case, one may think of the lower left triangle of the matrix (including its diagonal entries), which is then transposed to the upper right triangle. The number of elements in the lower triangle corresponds to the famous Gauss sum $N_{\text{symmetric}} = \frac{d(d+1)}{2}$. A (complex) Hermitian matrix has d independent (real) diagonal elements and consequently $N_{\text{complex, im}} = N_{\text{orthogonal}} - d = \frac{d(d-1)}{2}$ independent (complex) matrix elements. Each element consists of a real and imaginary part, and adding the d real diagonal elements makes in total

$$N_{\text{self-adjoint}} = 2\frac{d(d-1)}{2} + d = d^2$$

independent real-valued random variables.

3.2. Gaussian Ensembles

The commonly considered ensembles in random matrix theory are **Gaussian ensembles.** These ensembles contain real or complex symmetric matrices, whose entries are (standard) Gaussian variables [Spe19c][1], [Meh04]. One distinguishes between three Gaussian ensembles: Gaussian orthogonal ensemble (GOE), Gaussian unitary ensemble (GUE) and the Gaussian symplectic ensemble (GSE). These abbreviations refer to the specific symmetry transformation under which their joint probability distribution remains invariant. This means that an N-dimensional GUE matrix, call it A, is invariant under a unitary transformation V, i.e.:

$$A \mapsto A' := V^* A V, \tag{3.3}$$

where $V \in SU(d)$. The definitions for the GOE and the GSE are completely analogue, but one has to replace the underlying transformation matrix V by a (real) orthogonal matrix $O \in SO(d)$ and a (quarternionic) symplectic matrix $S \in Sp(d)$, where $Sp(d)$ denotes the N-dimensional symplectic group [Meh04]. In the following, we will restrict our analysis to the GUE case, which is the relevant ensemble in quantum mechanics because it contains the Hermitian symmetric matrices. A Gaussian random matrix of the GUE is hence defined as follows [Spe19c]:

[1] supplemented by online lecture videos [Spe19d]

Definition 3.2. *In the sense of Definition 3.1, i.e.* $\Omega = \mathbb{C}$, $\mathbb{P}(x,y) = \frac{1}{\pi}\exp\{-(x^2+y^2)\}$, *the* ***Gaussian random matrix*** $A_d \in M_d(L^{\infty-}(\Omega,\mathbb{P}))$ *is of the form* $A_d = (\widehat{a}_{kl})_{k,l=1}^d = \frac{1}{\sqrt{d}}(a_{kl})_{k,l=1}^d$, *where*

- $a_{kl} \sim \mathcal{CN}(0,1)$
- $a_{kl} = a_{lk}^*$
- $\{a_{kl} : l \leq k\}$ *are independent*

for all $k,l \in \{1,...,d\}$. *Writing* $a_{kl} = \frac{1}{\sqrt{2}}(x_{kl} + iy_{kl})$ *for* $k < l$, A_d *reads:*

$$A_d = \begin{cases} (\widehat{a}_{kl})_{k,l=1}^d = \frac{1}{\sqrt{2d}}(x_{kl} + iy_{kl}), & k < l \\ (\widehat{a}_{kl})_{k,l=1}^d = \frac{1}{\sqrt{2d}}(x_{kl} - iy_{kl}), & l < k \\ (\widehat{a}_{kk})_{k=1}^d = \frac{1}{\sqrt{d}}x_{kk}, & k = l. \end{cases} \tag{3.4}$$

This means that $x_{kl} \sim \mathcal{N}(0,\frac{1}{2})$ *and* $y_{kl} \sim \mathcal{N}(0,\frac{1}{2})$. *The set of these random matrices is denoted by* ***GUE(d).***

Note that the above normalization is chosen so that $\mathbb{E}[|\widehat{a}_{kl}|^2] = \frac{1}{2d}$ if $k \neq l$ and $\mathbb{E}[|\widehat{a}_{kk}|^2] = \frac{1}{d}$ if $k = l$. Moreover, the analytic expression for moments of standard Gaussian random variables is

$$\mathbb{E}[X^m] = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} t^m \exp\left(-\frac{t^2}{2}\right) dt = \begin{cases} 0, & m \text{ odd} \\ (m-1)!!, & m \text{ even.} \end{cases} \tag{3.5}$$

If one generalizes the Gaussian distribution in Definition 3.2 to arbitrary probability distributions supported on $\mathbb{C}$ but keeps Hermitian symmetry, the *Wigner random matrix* is defined as follows:

Definition 3.3. *A* ***Wigner matrix*** *is a Hermitian symmetric random matrix consisting of complex-valued random variables, which are drawn according to a fixed distribution on the off-diagonal, and its real-valued restriction on the diagonal.*

Definition 3.3 is formally the same as Definition 3.2, but one replaces $a_{kl} \sim \mathcal{CN}(0,1)$ by an arbitrary distribution. Following [MS17, Ch. 1.3-1.6], we construct the joint probability density function of a GUE matrix, depending on the entries of the matrix.

Lemma 3.4. *Let* $A_N \in GUE(N)$. *By rewriting the components of* A_N *to a* N^2*-dimensional Gaussian random vector*

$$(X_1,...,X_{N^2}) = (x_{11},...,x_{NN},x_{12},...,x_{1N},x_{2N},...,x_{N-1,N},y_{12},...,y_{1N},y_{2N},...,y_{N-1,N}),$$

the joint probability density $f_{A_N}(X_1,...,X_{N^2})$ *of a GUE matrix* A_N, *defined in equation (3.4) is*

$$f_{A_N}(X_1,...,X_{N^2}) = c_N \exp\{-N \operatorname{Tr}\left(A_N^2\right)\} dX, \tag{3.6}$$

where $dX = \prod_{k=1}^N dx_{kk} \prod_{\substack{k=1 \\ k<l}}^N dx_{kl}dy_{kl}$ *is the Lebesgue measure on* $\mathbb{R}^{N^2}$ *and*

$$c_N = \frac{1}{2^{\frac{N}{2}}}\left(\frac{N}{\pi}\right)^{\frac{N^2}{2}}$$

Proof. First step. We assume that the multivariate random variable $X = (X_1, ..., X_{N^2})$ is real valued. The probability density of this Gaussian vector is [MS17, Ch. 1.3]

$$f(X_1, ..., X_{N^2}) = c_N \exp\left\{ \frac{-1}{2} \langle C^{-1}X, X\rangle \right\} dX,$$

with $c_N = \frac{1}{\sqrt{\det(C)(2\pi)^{N^2}}}$, $dX = \prod_{k=1}^{N^2} dX_k = \prod_{k=1}^{N} dx_{kk} \prod_{\substack{k=1 \\ k<l}}^{N} dx_{kl} dy_{kl}$, C being the covariance matrix and $\langle \cdot, \cdot \rangle$ denoting the Euclidean inner product on $\mathbb{R}^{N^2}$. The covariance matrix is diagonal iff the random variables $X_1, ..., X_{N^2}$ are independent. We note that the independence assumption is fulfilled by recalling Definition 3.2.

Second step. Compute the covariance matrix C and the quantity $\langle C^{-1}X, X\rangle$. Since the X_i, $i \in [N^2]$ are mutually statistically independent, the covariance matrix C is diagonal with its diagonal entries being $\mathrm{Var}[X_i] = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2$. Because of $\mathbb{E}[X_i] = 0$, $\mathrm{Var}[X_i] = \mathbb{E}[X_i^2]$ and by construction:

$$\mathbb{E}[X_i^2] = \begin{cases} \frac{1}{N}, & \text{for the } X_i \text{ corresp. to the } N \text{ diagonal entries } x_{ii} \\ \frac{1}{2N}, & \text{for the } X_i \text{ corresp. to the } N^2 - N \text{ upper triangular entries } x_{kl}, k < l. \end{cases}$$

Thus, the N^2-dim. matrix C contains N times $\frac{1}{N}$ and $N^2 - N$ times $\frac{1}{2N}$ on the diagonal. Inverting this (diagonal) matrix is elementary and thus we find

$$C^{-1} = \mathrm{diag}(N, N, ..., N, 2N, 2N, ..., 2N), \qquad C = \mathrm{diag}\left(\frac{1}{N}, \frac{1}{N}, ..., \frac{1}{N}, \frac{1}{2N}, \frac{1}{2N}, ..., \frac{1}{2N}\right)$$

Then,

$$\begin{aligned} \langle C^{-1}X, X\rangle &= N \left(\sum_{i=1}^{N} {x_{ii}}^2 + 2 \sum_{\substack{i,j=1 \\ i<j}}^{N} ({x_{ij}}^2 + {y_{ij}}^2) \right) = N \left(\sum_{i=1}^{N} {x_{ii}}^2 + \sum_{\substack{i,j=1 \\ i\neq j}}^{N} ({x_{ij}}^2 + {y_{ij}}^2) \right) \\ &= N \left(\sum_{i=1}^{N} {x_{ii}}^2 + \sum_{\substack{i,j=1 \\ i\neq j}}^{N} |a_{ij}|^2 \right) = N \,\mathrm{Tr}\left({A_N}^2\right). \end{aligned}$$

The last identity can be found in Corollary B.1.

Third step. Finally, we evaluate the constant c_N. With $\det(C) = N^{\frac{N^2}{2}} 2^{\frac{N^2-N}{2}}$ and some algebraic manipulation, we have reached the assertion. □

Lemma 3.4 demonstrates, why joint probability densities corresponding to Gaussian ensembles are invariant under the respective symmetry transformation: the traciality of Tr is the reason for this invariance. For instance, for $A \in \mathrm{GUE}(N)$, the joint probability density is of the form

$$f_A(X_1, ..., X_{N^2}) = c \exp\{-N \mathrm{Tr}\,(A^2)\} dX.$$

Due to the cyclic permutation invariance of the trace, we have for $A' := UAU^*$

$$\begin{aligned} f_{A'}(X_1, ..., X_{N^2}) &= c \exp\{-N \mathrm{Tr}\,(A'^2)\} dX = c \exp\{-N \mathrm{Tr}\,((UAU^*)^2)\} dX \\ &= c \exp\{-N \mathrm{Tr}\,(A^2)\} dX = f_A(X_1, ..., X_{N^2}), \end{aligned}$$

where the Lebesgue measure dX is clearly invariant under unitary transformations. Similarly, one obtains the joint probability density function for the orthogonal and the symplectic ensemble, and shows invariance under their respective symmetry transformation in analogy to the GUE case.

3.3. Wigner's Semicircle Law

The semicircle law is the most prominent result in random matrix theory [Meh04]. It generalizes central limit theory for measurable functions, and is hence to be understood as *the* central limit theorem for self-adjoint random matrices. However, if Hermitian symmetry is dropped, the situation gets tremendously more complicated. The empirical spectral measures loose compact support on $\mathbb{R}$ and any combinatorial feature that we present in this chapter, gets lost. Until 2008, there was no rigorous proof of the widely conjectured, more general, *circular* law for random matrices. It was proven by Tao and Vu in [TV08]. In the following, we collect some results from combinatorial random matrix theory which yield an averaged version of Wigner's semicircle law for GUE random matrices. Astonishingly, a direct proof of the semicircle law makes use of purely combinatorial tools without using any notions of convergence. We closely follow chapter two of Roland Speichers lecture notes on random matrices [Spe19c].

Let X, Y be real valued, independent, (standard) Gaussian random variables (short notation: $X, Y \sim \mathcal{N}(0,1)$) and let $Z = \frac{X+iY}{\sqrt{2}}$ be the corresponding complex Gaussian random variable. Later on, we calculate the m'th moment of a GUE random matrix, i.e. $\mathbb{E}[\mathrm{tr}(A_N^m)]$. Classically independent random variables commute with respect to $\mathbb{E}$ and hence factorize according to (2.17). Due to linearity of the expectation functional and statistical independence, a direct computation verifies the following properties for real X, Y and for complex Z as defined above:

Proposition 3.5. *Let $X, Y \sim \mathcal{N}(0,1)$ and $Z = \frac{X+iY}{\sqrt{2}}$. Then,*

$$\mathbb{E}[Z] = 0, \quad \mathbb{E}[Z^2] = 0, \quad \mathbb{E}[\overline{Z}] = 0, \quad \mathbb{E}[\overline{Z}^2] = 0, \quad \mathbb{E}[|Z|^2] = 1$$

The combinatorial nature of moments of GUE random matrices arises from the above properties together with the statistical independence of X and Y. Let us consider an N-dimensional real random vector, $(X_1, ..., X_N)$ where each $X_k \in \{X, Y\}$ for $k \in \{1, ..., N\}$ Due to X and Y being standard Gaussian random variables, the above proposition makes the covariance matrix diagonal. In the combinatorial picture, this provides a "counting functional" which counts second moments,

$$\mathbb{E}[X_i X_j] = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{3.7}$$

Therefore, this functional counts "bonds" between equal random variables [Spe19c, Chapter 2]. This means that $\mathbb{E}[X_i X_j] = 1$ implies that both X_i and X_j belong to X. The applicability to GUE matrices stems from the following observation: from Proposition 3.5, we see that the expectation functional on complex random variables Z does not count bonds between equal $Z's$ but rather connects Z with its complex conjugate $\overline{Z}$:

$$\mathbb{E}[Z_m Z_n] = \begin{cases} 1, & Z_n = \overline{Z}_m \\ 0, & else. \end{cases} \tag{3.8}$$

Adapting this result to the definition of a GUE matrix, where we have necessarily $a_{ij} = \overline{a}_{ji}$, we observe for the second moments:

Corollary 3.6. *Let $A_d = \frac{1}{\sqrt{d}}(a_{ij})_{i,j=1}^d$ be a complex valued (self adjoint) GUE matrix. By using Hermitian symmetry and equation (3.8), it holds for the second moments:*

$$\mathbb{E}[a_{ij} a_{kl}] = \delta_{il}\delta_{jk}. \tag{3.9}$$

Proof. Due to Hermitian symmetry, one finds the constraint

$$\mathbb{E}[a_{ij}a_{kl}] = \mathbb{E}[a_{ij}\overline{a}_{lk}] = \mathbb{E}[\overline{a}_{ji}\overline{a}_{lk}] = \mathbb{E}[\overline{a}_{ji}a_{kl}].$$

Using a bijection $(m,n) \mapsto (a_{i(m)j(m)}, a_{k(n)l(n)})$ and eq. (3.8), one writes:

$$\mathbb{E}[a_{i(m)j(m)}a_{k(n)l(n)}] = \begin{cases} 1, & a_{k(n)l(n)} = \overline{a}_{i(m)j(m)} = a_{j(m)i(m)} \\ 0, & \text{else,} \end{cases} \tag{3.10}$$

which gives $i = l$ and $j = k$. □

Note that we can and will replace X, Y with any finite family of standard Gaussian random variables, e.g. $\{Y_1, ..., Y_p\}$. We further sketch the relation between random matrix theory and combinatorics, closely following [Spe19b, Ch. 2.1-2.3]. For $m \in \mathbb{N}$, consider the ordered set $[m]$, i.e. $[m] = \{1, ..., m\}$.

Definition 3.7. *[Spe19b, Def. 2.4] A pairing (or pair-partition) π of $[n]$ is a decomposition of $[n]$ into disjoint subsets of size 2, i.e. $\pi = \{V_1, ..., V_k\} \subset \mathcal{P}([n])$, $k = \frac{n}{2}$, such that $\forall\ i, j \in [k]$, and $i \neq j$*

- $V_i \subset [n]$
- $\#V_i = 2$
- $V_i \cap V_j = \emptyset$
- $\bigcup_{i=1}^{k} V_i = [n]$,

where $\mathcal{P}([n])$ denotes the power set over the set $[n]$. The set of all pairings of $[n]$ is denoted by

$$\mathcal{P}_2(n) = \{\pi \in \mathcal{P}([n]) | \pi \text{ is a pairing of } [n]\}.$$

Since any pairing decomposes $[n]$ into pairs, we can associate a unique map to $\pi : [n] \to [n]$ in the following sense: suppose that $l \in [k]$, and $(l_1, l_2) = V_l \in \pi$. Hence, $\pi(l_1) = l_2$ and $\pi(l_2) = l_1$, and $\pi \circ \pi|_{V_l} \equiv \mathrm{id}$. Thus, π (as a pair partition) is associated to a permutation (more precisely, a transposition) $\pi \in S_n$, that decomposes $[n]$ into $k := \frac{n}{2}$ disjoint *orbits* V_l, $l \in [k]$. Some group theoretic basics are listed in Appendix C. *Non-crossing* pair partitions are defined in the following [Spe19b, Defintion 2.17].

Definition 3.8. *$\pi \in \mathcal{P}_2(n)$ is said to be **non-crossing,** if there are no pairs (i,k) and (j,l) in π with $i < j < k < l$. We define the lattice of non-crossing pairings as*

$$\mathcal{NC}_2(n) = \{\pi \in \mathcal{P}_2(n) | \pi \text{ is non-crossing}\}.$$

The subsequent theorem follows [Spe19c, Theorem 2.8]

Theorem 3.9. *(Wick-Isserlis Formula) Let $\pi \in \mathcal{P}_2(n)$. Let Y_i, $i \in [p]$ be independent standard Gaussian random variables and let $x_1, ..., x_n \in \{Y_1, ..., Y_p\}$. Then, the **Wick formula** holds true:*

$$\mathbb{E}[x_1 \cdot ... \cdot x_n] = \sum_{\pi \in \mathcal{P}_2(n)} \prod_{(i,j) \in \pi} \mathbb{E}[x_i x_j] \tag{3.11}$$

Since A is self-adjoint, $D_A = VAV^*$ for some $V \in U(N)$, thus

$$D_A^m = (VAV^*)^m = VA^mV^* \quad \Rightarrow \quad \mathrm{tr}\{D_A^m\} = \mathrm{tr}\{A^m\}, \qquad m \in \mathbb{N},$$

which means that the spectral measure is naturally encoded into the moments, of both deterministic and random matrices:

$$\mathrm{tr}\{A^m\} = \frac{1}{N}\sum_{i=1}^{N} \lambda_i^m. \tag{3.12}$$

Thus, convergence of the spectral distribution is directly linked to the convergence of moments. This correspondence between *moments* and *histograms* is crucial for (the averaged version of) Wigner's semi-circle law. The next theorem provides an expression for the m'th moment of a random matrix. The theorem and its proof follow [NS06, Chapter 22] and [Spe19c, Chapter 2]. The proof is reproduced in Appendix B.

Theorem 3.10. *Let $A_N \in GUE(N)$. Let $j \in [m]$, $\pi \in \mathcal{P}_2(m)$ and define $\gamma\pi \in S_m$ as $\gamma(\pi)[j] := (\pi(j)+1) \mod m$. Then, for all $m \in \mathbb{N}$, it holds:*

$$\mathbb{E}[\mathrm{tr}\{A_N^m\}] = \sum_{\pi \in \mathcal{P}_2(m)} N^{\#(\gamma\pi) - \left(\frac{m}{2}+1\right)}, \tag{3.13}$$

where $\#(\cdot)$ denotes the number of cycles of a permutation.

Theorem 3.10 states that the m'th moment of a GUE matrix contains, at least *asymptotically*, corresponds to $C_{\frac{m}{2}}$, i.e. the $\frac{m}{2}$-th Catalan number. This quantity counts the number of non-crossing pair partitions. The next proposition ensures that the exponent in (3.13) is bounded by 0, so that the overall sum does not diverge.

Proposition 3.11. *[Spe19b, Prop. 2.20] Let $m \in \mathbb{N}$ be even and let $\pi \in \mathcal{P}_2(m)$, which is identified with a permutation $\pi \in S_m$. Again, let $\gamma \in S_m$ with $\gamma(j) := (j+1) \mod m$, for $j \in [m]$. Then, it holds:*

1. *$\#(\gamma\pi) - (\frac{m}{2}+1) \leq 0$ for all $\pi \in \mathcal{P}_2(m)$.*
2. *$\#(\gamma\pi) - (\frac{m}{2}+1) = 0$ if and only if $\pi \in \mathcal{NC}_2(m)$.*

If $\#(\gamma\pi) - (\frac{m}{2}+1) = 0$ holds true for $\pi \in S_m$, Proposition 3.11 yields that $\pi \in \mathcal{NC}_2(m)$. Then, $N^{\#(\gamma\pi)-\left(\frac{m}{2}+1\right)} = N^0 = 1$, i.e. each non-crossing pairing makes a contribution $\propto 1$, where the other pairings contribute $\propto \frac{1}{N} \to 0$, for N large. Thus, $\mathbb{E}[\mathrm{tr}\{A_N^m\}]$ *counts* $\mathcal{NC}_2(m)$-pairings in the limit $N \to \infty$.

Moreover, Theorem 3.10 is known as *Genus expansion* for the GUE [Spe19b, Remark 2.2], since

$$\mathbb{E}[\mathrm{tr}\{A_N^m\}] = \sum_{\pi \in \mathcal{P}_2(m)} N^{-2g(\pi)} \tag{3.14}$$

refers to the genus of an underlying Euler surface, where $\chi(g) = 2 - 2g$, and $\chi(g)$ being the algebraic Euler characteristic. This means that a crossing pairing has genus $g = 1$. Non-crossing pairings have $g = 0$.

Finally we have collected all preliminaries to present the averaged version of Wigner's semi-circle law [Spe19b, Theorem 2.21]. The semicircular law is the central limit theorem for random Hermitian matrices.

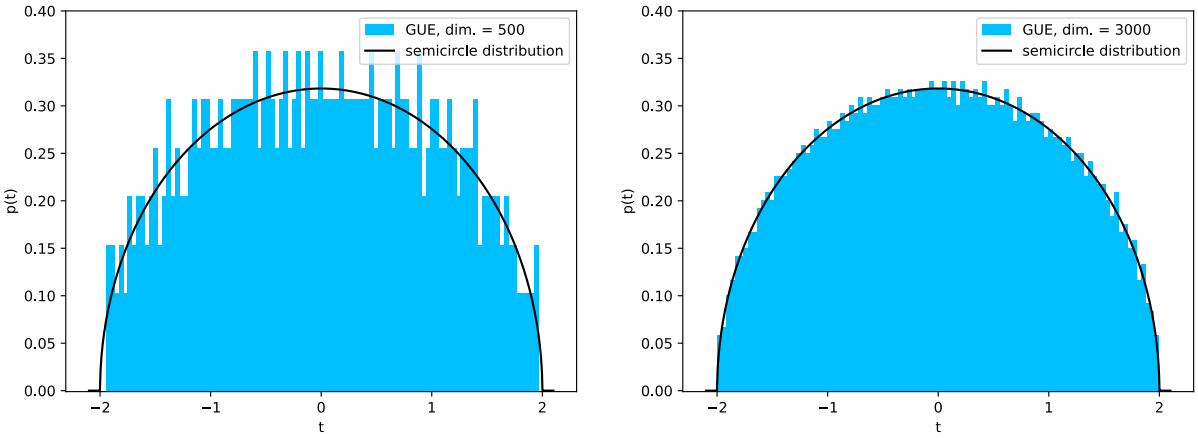
Theorem 3.12 (Wigner's semicircle law for the GUE, averaged version)**.** *Let $A_N \in GUE(N)$. Then, for all $m \in \mathbb{N}$:*

$$\lim_{N\to\infty} \mathbb{E}[\mathrm{tr}\{A_N^m\}] = \frac{1}{2\pi}\int_{-2}^{2} t^m\sqrt{4-t^2}dt = \begin{cases} 0, & \text{if } m \text{ is odd,} \\ C_{\frac{m}{2}}, & \text{if } m \text{ is even.} \end{cases} \tag{3.15}$$

The above version of the semi-circle law corresponds to convergence in average, since the sequence of moments $(\mathbb{E}[\mathrm{tr}\{A_N^m\}])_{N\in\mathbb{N}}$ uniformly converges (for m even) to the (large, but finite) Catalan number i.e.

$$\mathbb{E}[\mathrm{tr}\{A_N^m\}] \overset{N\to\infty}{\longrightarrow} C_{\frac{m}{2}} = \frac{1}{\frac{m}{2}+1}\binom{m}{\frac{m}{2}} \qquad \forall\, m \in \mathbb{N}. \tag{3.16}$$

The averaged semicircle law for *Wigner matrices* is the same as the expression in (3.15). However, the semicircle law also holds true in terms of *almost sure* convergence. Proving the almost sure case is substantially more difficult than the averaged, combinatorial formalism. We refer to Tao [Tao23], Tao and Vu [TV08] for details. So far, we have collected central results about random matrices. In order to treat those matrices as generalized random variables in some generalized probability space, we introduce non-commutative probability theory: *free probability.*



(a) GUE(d) for $d = 500$

(b) GUE(d) for $d = 3000$

Figure 3.1.: Wigner's semi-circle law shows that the spectral density of self-adjoint random matrices converges almost surely to the semi-circle distribution (3.15). We present a comparison between histograms displaying spectra of GUE(d) random matrices A_d, of dimension $d = 500$ (a), resp. $d = 3000$ (b) with the semicircle distribution. We see that $\|A_d\| \leq 2$. Moreover, the semi-cirle distribution becomes clearly visible, even for $d \propto 10^3$. The strong coincidence between spectral simulations and their limiting distributions builds the foundation of our numerical analysis in Chapter 4 and 5.

3.4. Free Independence

We leave random matrices for now to proceed with free probability. At the end of this chapter, we return to random matrices and see how these fields are glued together.

We present the main definition which distinguishes free probability from classical probability. Statistical independence is replaced by *free independence* [NS06, Chapter 5]. Free independence implies that the joint distribution of a family of (free) random variables is completely determined by the knowledge of the individual distributions [NS06, Ch. 5.5]. Hence, the free convolution machinery (Chapter 3.5) enables to determine joint distributions, assuming that the corresponding operator-valued random variables are freely independent.

Definition 3.13. *[Spe19b, Def .1.1] Let $(\mathcal{A}.\varphi)$ be a non-commutative $*-$probability space and let I be a finite index set. Unital subalgebras $(A_i)_{i\in I}$ of $\mathcal{A}$ are called* ***free*** *in $(\mathcal{A},\varphi)$, if for $k\in\mathbb{N}$, $\varphi(a_1\cdot\ldots\cdot a_k)=0$, whenever*

- *$i(j)\in I$ for all $j\in[k]$,*
- *$a_j\in\mathcal{A}_{i(j)}$, $j\in[k]$,*
- *$i(1)\neq i(2)\neq\ldots\neq i(k)$,*
- *$\varphi(a_j)=0$ for all $j\in[k]$.*

The first two aspects of Definition 3.13 form the abstract setting for non-commuting unital subalgebras. Thus, random variables cannot be arbitrarily permuted under φ, so that e.g. $\varphi(xyz)\neq\varphi(yxz)$ for $x\in\mathcal{A}_1, y\in\mathcal{A}_2, z\in\mathcal{A}_3$ is respected. The third point means that neighboring elements are from different subalgebras. However, this does not mean that e.g. $i(1)$ and $i(3)$ must be different. The last point says that the random variables a_j are *centered:* this is the free analogue of $\mathbb{E}[X]=0$. We require these subalgebras to be "unitally embedded" into $\mathcal{A}$, i.e. all subalgebras share the unit $1_{\mathcal{A}}$ of the larger algebra $\mathcal{A}$ [Mai17]. A family of random variables $(X_i)_{i\in I}$ is called free in the sense of Definition 3.13, if the subalgebras $\mathcal{A}_i$, that are generated by the X_i, are free. Definition 3.13 provides that joint moments in free variables factorize. Hence, we have reached the free analogue of (2.17), i.e.

$$\varphi(a^m b^n)=\varphi(a^m)\,\varphi(b^n),\qquad m,n\in\mathbb{N},\tag{3.17}$$

whenever $a,b\in\mathcal{A}$ are *free.* In a non-commutative *-probability space $(\mathcal{A},\varphi)$, semicircular variables $s\in\mathcal{A}$ satisfy for some $\sigma\in\mathbb{R}^+$

$$\varphi(s^k)=\begin{cases}0, & \text{if } k \text{ is odd},\\ \sigma^k C_{\frac{k}{2}}, & \text{if } k \text{ is even}.\end{cases}\tag{3.18}$$

If $\sigma=1$, s is called a *standard semicircular variable.* It is mentioned that in this case - in analogy to the standard normal random variable - $\varphi(s)=0$ and $\mathbb{V}_\varphi(s):=\varphi(s^2)-\varphi(s)^2=1$, where $\mathbb{V}_\varphi(\cdot)$ denotes the variance with respect to φ.

Furthermore, we define convergence in distribution for free variables, which is similar to the notion of convergence in distribution for measurable functions. We have [NS06, Definition 8.1]

Definition 3.14. *Let $(\mathcal{A}_n,\varphi_n)$, $(\mathcal{A},\varphi)$ be non-commutative probability spaces and let $a_n\in\mathcal{A}_n$ for all $n\in\mathbb{N}$. A sequence of random variables $(a_n)_{n\in\mathbb{N}}\subset\mathcal{A}$ is said to* ***converge in distribution*** *to $a\in\mathcal{A}$, if*

$$\lim_{n\to\infty}\varphi_n(a_n^m)=\varphi(a^m),\qquad\forall\, m\in\mathbb{N}.$$

The convergence is denoted by $a_n \xrightarrow{distr.} a$.

In the following, we recall the classical central limit theorem in the algebraic language and then present Voiculescu's version of a free central limit theorem. The latter shows, how Wigner's semicircle law for random matrices falls into the framework of free probability. We follow [Spe19b, Thm. 2.7 and 2.8][2].

Theorem 3.15 (Classical Central Limit Theorem)**.** *Let* $(\mathcal{A}, \varphi)$ *be a classical (i.e. commutative), algebraic probability space and let* $(a_i)_{i\in\mathbb{N}}$ *be a family of (classically) independent random variables. Since all* a_i *commute, joint moments factorize, i.e. for all* $i : [k] \to \mathbb{N}$*, with* $i(1) < i(2) < i(3) < ... < i(k)$*, and* $r_1, ..., r_k \in \mathbb{N}$*,*

$$\varphi\left(a_{i(1)}^{r_1} a_{i(2)}^{r_2} ... a_{i(k)}^{r_k}\right) = \varphi\left(a_{i(1)}^{r_1}\right) \varphi\left(a_{i(2)}^{r_2}\right) ... \varphi\left(a_{i(k)}^{r_k}\right).$$

Moreover, we assume that all variables are ***independent and identically distributed (i.i.d)****, with mean zero and variance one, i.e.*

$$\varphi(a_i^r) = \varphi(a_j^r), \qquad \varphi(a_i) = 0, \qquad \varphi(a_i^2) = 1,$$

for all $i, j, r \in \mathbb{N}$ *and define for every* $n \in \mathbb{N}$, $S_n := \frac{1}{\sqrt{n}}(a_1 + a_2 + ... + a_n)$*. Then,* $(S_n)_{n\in\mathbb{N}}$ *converges in distribution to a* ***normal*** *random variable, i.e. for all* $k \in \mathbb{N}$ *:*

$$\lim_{n\to\infty} \varphi(S_n^k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} t^k e^{-\frac{t^2}{2}}\, dt = \begin{cases} 0, & \text{if } k \text{ is odd,} \\ (k-1)!!, & \text{if } k \text{ is even.} \end{cases} \tag{3.19}$$

The analogue in free probability theory was established by Voiculescu as follows.

Theorem 3.16 (Free Central Limit Theorem, Voiculescu 1985)**.** *Let* $(\mathcal{A}, \varphi)$ *be a non-commutative probability space and let* $(a_i)_{i\in\mathbb{N}}$ *be a family of freely independent random variables in* $(\mathcal{A}, \varphi)$*. Again, we assume that all variables are i.i.d with* $\varphi(a_i) = 0$ *and* $\varphi(a_i^2) = 1$*, and define for every* $n \in \mathbb{N}$, $S_n := \frac{1}{\sqrt{n}}(a_1 + a_2 + ... + a_n)$*. Then,* $(S_n)_{n\in\mathbb{N}}$ *converges in distribution to a* ***semicircular*** *variable, i.e. for all* $k \in \mathbb{N}$ *:*

$$\lim_{n\to\infty} \varphi(S_n^k) = \frac{1}{2\pi} \int_{-2}^{2} t^k \sqrt{4 - t^2}\, dt = \begin{cases} 0, & \text{if } k \text{ is odd,} \\ C_k, & \text{if } k \text{ is even.} \end{cases} \tag{3.20}$$

Comparing (3.20) to Theorem 3.12 indicates the connection between random matrices and free probability. Taking $\mathcal{A} := M_d(L^{\infty-}(\mathbb{C}, \mathcal{CN}(0,1)))$ with state $\varphi(\cdot) = \mathbb{E}[\mathrm{tr}(\cdot)]$ demonstrates how random matrices follow Voiculescu's free central limit theorem, if one considers random matrices to be freely independent random variables in a tracial, non-commutative C^* probability space. However, a purely combinatorial formalism of freeness does not provide analytical tools to compute joint probability densities. Hence, we introduce the analytical description in the following.

[2] supplemented by online lecture videos [Spe19a]

3.5. The Cauchy-Stieltjes Transform

A classical random variable can be described via its characteristic function and its moment-generating function. For a random variable with probability density, the characteristic function corresponds to its Fourier transform. In order to encode distributions of free random variables analytically (a random variable is fully described by the knowledge of all moments), there exists a similar notion in free probability: the Cauchy transform. Up to the sign, the terms Cauchy transform and Stieltjes transform are used synonymously [Mai17]. Throughout, we denote the upper complex plane by $\mathbb{C}^+ := \{z \in \mathbb{C} | \mathrm{Im}\{z\} > 0\}$ and $\mathbb{C}^-$ similarly.

Definition 3.17 (Cauchy transform)**.** *Let μ be a Borel probability measure on $\mathbb{R}$ and $z \in \mathbb{C}^+$. The **Cauchy transform** of μ, $G_\mu : \mathbb{C}^+ \to \mathbb{C}^-$ is defined by*

$$G_\mu(z) := \int_{\mathbb{R}} \frac{1}{z-t} d\mu(t). \tag{3.21}$$

The Stieltjes transform is defined by $S_\mu(z) := -G_\mu(z)$. Note that this minus sign changes the image from $\mathbb{C}^-$ to $\mathbb{C}^+$.

Since $z \in \mathbb{C}^+$ and μ being supported on $\mathbb{R}$, $G_\mu(\cdot)$ is bounded. The Cauchy transform provides an analytic tool that lifts the distributional information from the level of probability measures to the level of holomorphic functions[3]. The subsequent theorem (compare with [Spe19c, Theorem 4.9]) summarizes some properties about the Cauchy transform that are used for the free convolution, later on.

Theorem 3.18. *Let G_μ be the Cauchy transform for some compactly supported probability measure μ and $z := x + iy \in \mathbb{C}^+, x, y \in \mathbb{R}$. Then,*

- *G_μ is analytic on $\mathbb{C}^+$*
- *it holds:*
$$\lim_{y\to\infty} iy\, G_\mu(iy) = 1, \qquad \sup_{y>0, x\in\mathbb{R}} y|G_\mu(x+iy)| = 1. \tag{3.22}$$
- *Probability measures are uniquely encoded into G, i.e. if $G_\mu(z) = G_\nu(z)$ for all $z \in \mathbb{C}^+$, then $\mu = \nu$. Moreover, the probability measure μ can be uniquely recovered from G_μ via the **Stieltjes inversion formula:***
$$-\lim_{\varepsilon\to 0^+} \frac{1}{\pi} \int_a^b Im\left(G_\mu(x+i\varepsilon)\right) = \mu((a,b)) + \frac{1}{2}\mu(\{a,b\}) \tag{3.23}$$

The *Stieltjes inversion formula* for the probability density of the measure is captured via

$$\mu(x) = \frac{-1}{\pi} \lim_{\varepsilon\to 0^+} \mathrm{Im}\left(G_\mu(x+i\varepsilon)\right). \tag{3.24}$$

In addition to Theorem 3.18, the converse of (3.22) holds true, as well: if an analytic function $G : \mathbb{C}^+ \to \mathbb{C}^-$ satisfies $\limsup_{y\to\infty} y|G(iy)| = 1$, then there exists a unique probability measure

[3] Small remark on the mathematical beauty of this object: the Cauchy transform is characterized within "free harmonic analysis". This field connects operator algebras, probability theory, functional calculus, unitary representation theory and complex analysis. This intersection demonstrates and motivates the depth of random matrix theory and justifies its well-deserved role in modern mathematical research.

μ on $\mathbb{R}$ such that $G \equiv G_\mu$. The Stieltjes inversion formula inherently includes a theorem for a sequence of Borel measures that converge weakly in the dual space of bounded continuous functions. More precisely, if the density (3.24) is used to parameterize a family of measures, i.e.

$$d\mu_\varepsilon(t) = \frac{-1}{\pi}\mathrm{Im}\,(G_\mu(x+i\varepsilon)),$$

then, $d\mu_\varepsilon$ defines for any $\varepsilon > 0$ an absolutely continuous probability measure μ_ε on $\mathbb{R}$ [Mai17, Theorem I.1.29]. Then,

$$\lim_{\varepsilon\to 0^+}\int_{\mathbb{R}} f(t)d\mu_\varepsilon(t) = \int_{\mathbb{R}} f(t)d\mu(t), \qquad f \in C_b(\mathbb{R},\mathbb{C}). \tag{3.25}$$

Similarly to the Cauchy integral formalism in complex analysis, the *moments* of μ are defined via the Laurent series of the Cauchy transform. Due to the compact support of μ, let $r := \sup_{t\in\mathbb{R}}\{|\mu(t)| > 0\}$, so $\mu([-r,r]) = 1$, G_μ admits a Laurent series expansion around ∞, as follows [Spe19b], [Mai17]:

$$G_\mu(z) = \sum_{n=0}^{\infty}\frac{m_n(\mu)}{z^{n+1}}, \qquad z \in \mathbb{C}^+, \quad |z| > r. \tag{3.26}$$

Consequently, the $n-th$ moment of μ is derived from G_μ to

$$m_n(\mu) := \int_{\mathbb{R}} t^n d\mu(t). \tag{3.27}$$

Hence, a probability measure μ is said to be *determined by its moments*, if all moments $m_n(\mu)$, $n \in \mathbb{N}_0$ exist. Thus, the existence of all moments for compactly supported probability measures ensures the well-definition of the notion of *convergence in distribution.*

It is left to combine the above notions with free variables in a C^* probability space $(\mathcal{A},\varphi)$. Let $a \in \mathcal{A}$ be self-adjoint, and let μ_a be the associated, (real-valued) probability distribution. The Cauchy transform is constructed as follows. The state φ encodes the distribution of a, due to $\varphi(\cdot) := \int_{\mathbb{R}}(\cdot)\,d\mu_a$ (construction via Riesz representation, Chapter 2). Applying the holomorphic functional calculus for C^* probability spaces, any $f \in C_b(\mathbb{R},\mathbb{C})$ applied to the operator $a \in \mathcal{A}$ takes the form

$$f(a) = \int_{\mathrm{supp}(\mu_a)} f(t)\,d\mu_a(t) \in \mathbb{C}.$$

By taking $\tilde{f} : \mathbb{C}^+ \times \mathcal{A} \to \mathbb{C}$, $f \in C_b(\mathbb{C}^+\times\mathbb{R},\mathbb{C})$, the functional calculus takes the form

$$(z,a) \mapsto \tilde{f}(z,a) := \int_{\mathrm{supp}(\mu_a)} f(z,t)\,d\mu_a(t) \in \mathbb{C}. \tag{3.28}$$

With

$$f \equiv f(z,x) := \frac{1}{z-x}, \qquad x \in \mathbb{R}, \qquad z \in \mathbb{C}^+, \tag{3.29}$$

we get

$$\tilde{f}(z,a) = \int_{\mathrm{supp}(\mu_a)} f(z,t)\,d\mu_a(t) = \int_{\mathrm{supp}(\mu_a)} \frac{1}{z-t}\,d\mu_a(t), \qquad a = a^* \in \mathcal{A}. \tag{3.30}$$

Then, $G_{\mu_a}(z)$ is actually the mapping $G_{\mu_a}(z,a) \equiv \tilde{f}(z,a)$, as in (3.30), where the Cartesian product with the operator algebra is suppressed, but indexed in the underlying spectral distribution,

$$G_{\mu_a}(z) := \int_{\mathbb{R}} \frac{1}{z-t} d\mu_a(t) = \varphi((z\mathbb{1}_{\mathcal{A}} - a)^{-1}), \qquad z \in \mathbb{C}^+. \tag{3.31}$$

In this way, we conclude that G_{μ_a} is a holomorphic mapping that has an analytic extension to the resolvent set $\mathbb{C} \setminus \operatorname{supp}(\mu_a) = \mathbb{C} \setminus \sigma(a) \equiv \rho(a)$ [Mai17, Remark I.1.33].

Example 3.19. *Let $A \in M_d(\mathbb{C})$ and take the (classical) C^* probability space $(M_d(\mathbb{C}), \frac{1}{N}\operatorname{tr})$. Then, by taking $\mu_A = \sum_{i=1}^d \delta_{\lambda_i}$, the Cauchy transform is for $z \in \mathbb{C}^+$,*

$$G_{\mu_A}(z) = \int_{\mathbb{R}} \frac{1}{z-t}\, d\mu_A(t) = \int_{\mathbb{R}} \frac{1}{z-t} \left(\frac{1}{d} \sum_{i=1}^d d\delta_{\lambda_i} \right) = \frac{1}{d} \sum_{i=1}^d \frac{1}{z-\lambda_i}.$$

3.6. Free Additive Convolution

We have already provided the notion of free independence together with the encoding of probability distributions into holomorphic functions. The last step towards determining joint distributions in free variables consists of establishing the *free additive convolution*. This object replaces the convolution integral for classical random variables. We follow Speicher [Spe19b, Chapter 4]. An introduction to the theory of free additive convolution semigroups can be found in [NS06, Chapter 14.5]. Support properties are discussed in [BES20].

Definition 3.20. *Let $(\mathcal{A}, \varphi)$ be a *-NCPS and $x, y \in \mathcal{A}$ self-adjoint random variables with probability measures μ_x, and μ_y. The probability measure μ_{x+y} is called the free additive convolution of μ_x and μ_y and denoted by:*

$$\mu_{x+y} = \mu_x \boxplus \mu_y \tag{3.32}$$

Before collecting some properties about this type of convolution, we recall a theorem that provides its existence. This goes back to the construction of a product between algebras that respects freeness [NS06, Chapter 6]. We present the final result, but omit details [Spe19b, Theorem 4.5].

Theorem 3.21. *For two compactly supported probability measures μ, ν on $\mathbb{R}$, there exists a *-probability space $(\mathcal{A}, \varphi)$ and self-adjoint $x, y \in \mathcal{A}$ such that:*

1. *the distributions μ_x and μ_y satisfy: $\mu_x = \mu$ and $\mu_y = \nu$,*
2. *x, y are free in $(\mathcal{A}, \varphi)$.*

The sequence of joint moments $(\varphi(z^n))_{n\in\mathbb{N}}$ for $z := x + y \in \mathcal{A}$ is exponentially bounded and thus determines uniquely a compactly supported probability measure $\mu_{x+y} \equiv \mu \boxplus \nu$ on $\mathbb{R}$. $(\varphi(x^n))_{n\in\mathbb{N}}$, $(\varphi(y^n))_{n\in\mathbb{N}}$ and $(\varphi(z^n))_{n\in\mathbb{N}}$ only depend on freeness between x and y, and on μ and ν, but not on concrete realizations of x and y.

Theorem 3.21 comprises that mixed cumulants in free variables vanish (e.g. [Spe09, Thm. 22.4.2]). The state φ can be defined as a n-linear functional $\varphi_n : \mathcal{A}^n \to \mathbb{C}$, via $\varphi_n(a_1, ..., a_n) :=$

$\varphi(a_1...a_n)$. Hence, the *cumulant functionals* $\kappa_n : \mathcal{A}^n \to \mathbb{C}, n \in \mathbb{N}$ are defined by moment-cumulant relations for free variables, i.e. $\kappa = \varphi * \mu$. If $x, y \in \mathcal{A}$ are free, then the cumulant functional κ_n is separated as follows:

$$\kappa_n^{x+y} = \kappa_n^x + \kappa_n^y. \tag{3.33}$$

We refer to [Spe09], [Spe19b, Chapter 3], [NS06, Part 2] for details. Hence, the linearization of free cumulants separates the generating power series, as well. This is a key feature of *Voiculescu's R-transform,* which boils down to the vanishing mixed cumulants of free variables [Spe19b]:

Definition 3.22. *Let μ be a compactly supported, real-valued probability measure. We define the power series expansion*

$$R_\mu(z) := \sum_{n=1}^{\infty} \kappa_n z^{n-1}, \qquad |z| < c\frac{1}{r}, \tag{3.34}$$

where $c > 0$, $(\kappa_n)_{n\in\mathbb{N}}$ is the collection of cumulants corresponding to the moments $(m_n)_{n\in\mathbb{N}}$ (3.27) and $r \in \mathbb{R}^+$ is sufficiently large, such that $\mu[-r, r] = 1$.

Thus, $R_{\mu\boxplus\nu}(z) = R_\mu(z) + R_\nu(z)$, for z on a sufficiently small, suitable domain. We refer to [Spe19b, Chapter 4] for details regarding domain properties. Since $\boxplus$ is commutative, associative, translation invariant, and has a neutral element, the free additive convolution forms a (convolution) semigroup, i.e.

$$t \mapsto (\mu^{\boxplus t}), \qquad t \geq 1. \tag{3.35}$$

Here, the neutral element is the Delta distribution δ_0, i.e. $\delta_0 \boxplus \mu = \mu$, for all measures μ. Conceptually, the semigroup structure is obtained for the classical convolution as well. Given two arbitrary (real-valued) probability measures, how can we actually compute their free additive convolution? This is answered in Chapter 4 both analytically and numerically.

3.7. Freeness between Unitary Random and Deterministic Matrices

Computing the free convolution for two spectral measures requires - as the name suggests - freeness between the corresponding operators. Given two deterministic operators $A, B \in \mathcal{B}(\mathcal{H})$, how can we suitably modify the situation, to apply the free convolution machinery to the spectral measures of A and B? One answer to this question is given by applying a random unitary transformation $\Phi_U(\cdot) := U(\cdot)U^*$ to either A or B. W.l.o.g., we apply this transformation to $B' := \Phi_U(B)$. In the final part of this chapter about random matrices and free probability, we demonstrate how (asymptotic) freeness arises between A and B'. Together with the free convolution theory, we will have completed the compendium on free probability to treat linear combinations of free variables analytically. We follow [MS17, Chapter 4.3] and [NS06, Chapter 23].

Definition 3.23. *Let $(\mathcal{A}, \varphi)$ be a *-NCPS. $u \in \mathcal{A}$ is called **Haar unitary,** if*

- $uu^* = u^*u = \mathbb{1}_\mathcal{A}$.
- $\varphi(u^k) = \delta_{0,k}$ *for all* $k \in \mathbb{Z}$.

In practice, unitary random matrices are generated by applying the Gram-Schmidt orthogonalization procedure (column-wise) to a (non-self-adjoint) Gaussian random matrix. Due to representation theory of unitary groups, $\mathcal{U}(d)$ is a compact group and the Haar measure, dU is the unique, translation invariant measure on $\mathcal{U}(d)$ [CMN22]. It is basically the analogue of the Lebesgue measure, where the translation invariant shifts of Borel sets generalize to the rotational symmetry of the manifold $\mathcal{U}(d)$. Geometrically, $UU^* = \mathbb{1}$ is equivalent to the isometry condition for U, i.e. $\langle Uv, Uw\rangle = \langle v, w\rangle$, for all $v, w \in \mathcal{H}$.

However, Haar unitary random matrices are substantially more complicated compared to Wigner matrices. The entry-wise random variables cannot be uniquely determined by their mean and variance anymore [CMN22]. Due to $UU^* = \mathbb{1}$, the entries are nevertheless (somehow) correlated. Precisely this correlation is inherited in the **Weingarten calculus,** which now replaces the combinatorial Wick-Isserlis description for moments of Gaussian and Wigner random matrices. The main problem of Weingarten calculus boils down to computing Weingarten integrals of given representations of compact groups [CMN22]. The Weingarten integral is of the form [CMN22]

$$I_{ij} = \int_G \prod_{x=1}^{d} U_{i(x)j(x)}(g)\, dg, \tag{3.36}$$

where i, j are multi-indices, G the respective compact group of interest and dg being the Haar measure. The characteristic property of the Haar measure is for fixed $g_0 \in G$:

$$\int_G F(g_0 g)\, dg = \int_G F(g g_0)\, dg = \int_G F(g)\, dg \tag{3.37}$$

By taking G to be the unitary group, the Weingarten convolution formula reads [Nec18]

$$\mathbb{E}[u_{i_1,j_1} \cdots u_{i_n,j_n} \overline{u_{i'_1,j'_1}} \cdots \overline{u_{i'_n,j'_n}}] = \sum_{\sigma,\tau \in S_n} \prod_{r=1}^{n} \delta_{i_r,i'_{\sigma(r)}} \delta_{j_r,j'_{\tau(r)}} \mathrm{Wg}_N(\tau\sigma^{-1}), \tag{3.38}$$

where $\mathrm{Wg}_N(\tau\sigma^{-1})$ is the combinatorial Weingarten function, which can be seen as a generalization of the counting function from the genus expansion (Theorem 3.10). Using the Weingarten formula (3.38) to expand joint moments in deterministic matrices and Haar unitaries is the key idea to prove the subsequent Theorem 3.25. For a graph-theoretical viewpoint, one might consult Nechita et al. [Nec18]. The subsequent results by Voiculescu [Voi91] build the foundation for Chapter 4 and 5. In the subsequent theorems (see [MS17, Ch. 4.3, Theorem 8, 9]), u_i are Haar unitaries and d_j denote deterministic limits for the sequence of deterministic random matrices.

Theorem 3.24 (Voiculescu, 1991)**.** *Let $p, q \in \mathbb{N}$ and let $\{U_N^{(i)}\}_{i=1,\dots,p}$ be a family of p independent Haar unitary random matrices and let $\{D_N^{(j)}\}_{j=1,\dots,q}$ be family of q deterministic matrices. We assume that*

$$\begin{aligned} U_N^{(1)}, \dots, U_N^{(p)} &\xrightarrow{distr.} u_1, \dots, u_p \\ D_N^{(1)}, \dots, D_N^{(q)} &\xrightarrow{distr.} d_1, \dots, d_q \qquad N \to \infty. \end{aligned}$$

Then, for $N \to \infty$,

$$\{U_N^{(i)}, U_N^{(i)*}\}_{i=1}^{p}, \{D_N^{(j)}\}_{j=1}^{q} \xrightarrow{distr.} \{u^{(i)}, u^{(i)*}\}_{i=1}^{p}, \{d^{(j)}\}_{j=1}^{q},$$

where $\{u_1, u_1^*\}, \dots, \{u_p, u_p^*\}, \{d_1, \dots, d_q\}$ are free.

A proof is located at [NS06, Chapter 23]. The above theorem holds also true for almost sure convergence. The next theorem follows directly from Theorem 3.24.

Theorem 3.25. *Let $A_N, B_N \in M_N(\mathbb{C})$ and consider the sequences $(A_N)_{N\in\mathbb{N}}$, $(B_N)_{N\in\mathbb{N}}$ which are assumed to converge in distribution, i.e.*

$$A_N \xrightarrow{distr.} a, \qquad B_N \xrightarrow{distr.} b, \qquad N \to \infty,$$

*for $a, b \in \mathcal{A}$ in some *-probability space $(\mathcal{A}, \varphi)$. Moreover, let $(U_N)_{N\in\mathbb{N}}$ be a sequence of Haar unitary random matrices. Then*

$$A_N, U_N B_N U_N^* \xrightarrow{distr.} a, b$$

where a, b are free.

Strengthening the above theorem to almost sure convergence holds true, as well. In particular, A_N and $U_N B_N U_N^*$ are **almost surely asymptotically free.**

4. The Subordination Algorithm

In this chapter, we analyze the main algorithm that builds the foundation of combining distinct spectra of random matrices: the subordination algorithm. This algorithm provides the numerical method to compute the free additive convolution of two real-valued, compactly supported probability measures. Its scalar-valued formulation was provided by Belinschi and Bercovici in 2007 [BB07]. A decade later, the operator-valued extension of the subordination algorithm was published by Belinschi, Mai, and Speicher [BMS17]. The key principle of the algorithm is a fixed point method of two concatenated Cauchy transforms. The two Cauchy transforms belong to the corresponding probability measures, of which the joint probability distribution shall be computed.

4.1. Subordinating Cauchy Transforms: Solving a Fixed Point Equation

At first, we restrict ourselves to the scalar-valued suborindation method [BB07]. We construct the fixed point equation of the subordinated Cauchy transforms, closely following the survey, which is presented in Roland Speicher's lecture notes [Spe19b, Chapter 5] and refer to Belinschi et al. [BB07] for details regarding the convergence proofs. Following Belinschi et al. [BB07], the free additive convolution has the following subordination property (first shown in [Voi94] and [Bia98]): There exists an analytic function $\omega : \mathbb{C}^+ \to \mathbb{C}^+$ such that

$$G_{\mu \boxplus \nu} = G_\mu \circ \omega, \tag{4.1}$$

such that

$$\lim_{y \to \infty} \frac{\omega(iy)}{iy} = 1. \tag{4.2}$$

Subordinating the function ω to G_μ inherits its name to the subsequent theory and the numerical algorithm. Here, one can (at least) guess, how the information about the additive convolution about two (!) measures is encoded into one (!) measure.

4.1.1. Towards a Subordination Formalism

Let μ and ν be two compactly supported probability measures on $\mathbb{R}$ and let G denote the Cauchy transform, as usual. By observations from *free harmonic analysis* [MS17, Chapter 3], one has the following implicit relation between the formal Laurent series representation of Voiculescu's $R-$ transform and the Cauchy transform:

$$G\left(R(z) + \frac{1}{z}\right) = z, \tag{4.3}$$

and

$$R(G(z)) + \frac{1}{G(z)} = z. \tag{4.4}$$

Since the $R-$ transform linearizes the free additive convolution of the measures,

$$R_{\mu\boxplus\nu}(z) = R_\mu(z) + R_\nu(z), \tag{4.5}$$

we get for the Cauchy transform (at least implicitly)

$$G_{\mu\boxplus\nu}\left(R_{\mu\boxplus\nu}(z) + \frac{1}{z}\right) = z. \tag{4.6}$$

Since the relation between the $G-$ and the $R-$ transform is not distributive, the free additive convolution is a non-distributive operation as well. Moreover, the brute force attempt, to solve equation (4.6) numerically, fails in most cases [Spe19b].

The subsequent ansatz passes on its name to the resulting algorithm, which numerically solves the problem of finding freely independent, joint probability distributions: the ***subordination*** algorithm. Due to its significance to the present thesis, we will develop its foundations - closely following Chapter five in [Spe19b] - in detail.

Instead of using the $R-$ transform in equation (4.6), we define two new power series expansions in the following way:

$$\omega_1(z) := z - R_\nu(G_{\mu\boxplus\nu}(z)), \qquad \omega_2(z) := z - R_\mu(G_{\mu\boxplus\nu}(z)). \tag{4.7}$$

W.l.o.g., we consider ω_1. Inserting this ansatz into the Cauchy transform of μ gives

$$G_\mu(\omega_1(z)) = G_\mu(z - R_\nu(G_{\mu\boxplus\nu}(z))). \tag{4.8}$$

Due to equation (4.5), $R_{\mu\boxplus\nu}(G_{\mu\boxplus\nu}(z)) = R_\mu(G_{\mu\boxplus\nu}(z)) + R_\nu(G_{\mu\boxplus\nu}(z))$ so that equation (4.8) becomes

$$G_\mu(\omega_1(z)) = G_\mu\left(z - \{R_{\mu\boxplus\nu}(G_{\mu\boxplus\nu}(z)) - R_\mu(G_{\mu\boxplus\nu}(z))\}\right). \tag{4.9}$$

We use the functional relation (4.4), to obtain $R_{\mu\boxplus\nu}(G_{\mu\boxplus\nu}(z)) = z - \frac{1}{G_{\mu\boxplus\nu}(z)}$ and thus:

$$\begin{aligned} G_\mu(\omega_1(z)) &= G_\mu\left(z - \left\{z - \frac{1}{G_{\mu\boxplus\nu}(z)} - R_\mu(G_{\mu\boxplus\nu}(z))\right\}\right) \\ &= G_\mu\left(R_\mu(G_{\mu\boxplus\nu}(z)) + \frac{1}{G_{\mu\boxplus\nu}(z)}\right). \end{aligned}$$

Finally, we define $\tilde{z} := G_{\mu\boxplus\nu}(z)$ and take equation (4.3) into account, i.e. $G_\mu\left(R_\mu(\tilde{z}) + \frac{1}{\tilde{z}}\right) = \tilde{z}$. Thus, we get

$$G_\mu(\omega_1(z)) = G_\mu\left(R_\mu(G_{\mu\boxplus\nu}(z)) + \frac{1}{G_{\mu\boxplus\nu}(z)}\right) = G_{\mu\boxplus\nu}(z). \tag{4.10}$$

In this way, the function $\omega_1(z)$ is *subordinated* to $G_\mu(z)$, so that knowledge about $\omega_1(z)$ provides knowledge about $G_{\mu\boxplus\nu}(z)$ via Stieltjes inversion. In complete analogy, we construct the subordination formalism for $\omega_2(z)$. We summarize:

$$G_\mu(\omega_1(z)) = G_{\mu\boxplus\nu}(z), \qquad G_\nu(\omega_2(z)) = G_{\mu\boxplus\nu}(z). \tag{4.11}$$

However, we have not yet discussed how to determine $\omega_1(z)$, and $\omega_2(z)$ respectively. This is demonstrated in the following.

$$\begin{aligned}\omega_1(z)+\omega_2(z) &= z+z-R_\nu(G_{\mu\boxplus\nu}(z))-R_\mu(G_{\mu\boxplus\nu}(z))\\ &= z+z-(R_\nu(G_{\mu\boxplus\nu}(z))+R_\mu(G_{\mu\boxplus\nu}(z)))\\ &\overset{(4.5)}{=} z+z-R_{\mu\boxplus\nu}(G_{\mu\boxplus\nu}(z)) \overset{(4.4)}{=} z+\frac{1}{G_{\mu\boxplus\nu}(z)}.\end{aligned}$$

Consequently,

$$\omega_1(z)+\omega_2(z)=z+\frac{1}{G_{\mu\boxplus\nu}(z)} \overset{(4.11)}{=} z+\frac{1}{G_\mu(\omega_1(z))}.$$

Equivalently,

$$\omega_1(z)+\omega_2(z)=z+\frac{1}{G_\nu(\omega_2(z))}.$$

Therefore, we arrive at

$$\omega_1(z)+\omega_2(z)-\frac{1}{G_\mu(\omega_1(z))}=z, \qquad \omega_1(z)+\omega_2(z)-\frac{1}{G_\nu(\omega_2(z))}=z.$$

In order to combine the above equations, we define

$$H_i(z):=\frac{1}{G_i(z)}-z, \qquad i\in\{\mu,\nu\}. \tag{4.12}$$

Thus,

$$\begin{aligned}\omega_2(z) &= z+\frac{1}{G_\mu(\omega_1(z))}-\omega_1(z)=z+H_\mu(\omega_1(z)),\\ \omega_1(z) &= z+\frac{1}{G_\nu(\omega_2(z))}-\omega_2(z)=z+H_\nu(\omega_2(z)),\end{aligned}$$

Inserting the first of the above equations into the second, we finally get the relation

$$\omega_1(z)=z+H_\nu(z+H_\mu(\omega_1(z))), \tag{4.13}$$

or, (in terms of the Cauchy transform)

$$\omega_1(z)=\omega_1(z)+\frac{1}{G_\nu\left(z-\omega_1(z)+\frac{1}{G_\mu(\omega_1(z))}\right)}-\frac{1}{G_\mu(\omega_1(z))} \tag{4.14}$$

As it was shown by Belinschi and Bercovici [BB07], this ω_1-dependent mapping forms a contraction mapping on $\mathbb{C}^+$. As a consequence, $\omega_1(z)$ can be determined by a fixed point algorithm, which determines for each $z\in\mathbb{C}^+$ a unique fixed point $\omega_1(z)\in\mathbb{C}^+$. The subsequent theorem follows [Spe19b, Thm. 5.11].

Theorem 4.1 (Subordination Convergence Theorem)**.** *Let μ,ν be two real-valued probability measures. Then, the mapping $g:\mathbb{C}^+\times\mathbb{C}^+\to\mathbb{C}^+$,*

$$g(z,w):=w+\frac{1}{G_\nu\left(z-w+\frac{1}{G_\mu(w)}\right)}-\frac{1}{G_\mu(w)} \tag{4.15}$$

is well defined. For fixed $z \in \mathbb{C}^+$, it naturally induces a contraction mapping in the $w-$component, i.e.

$$g_z : \mathbb{C}^+ \to \mathbb{C}^+, \qquad w \mapsto g_z(w) := g(z, w),$$

which converges for each $z \in \mathbb{C}^+$ to a unique fixed point $w(z) \in \mathbb{C}^+$, i.e.

$$w(z) = \lim_{n\to\infty} g_z^{\circ n}(w_0) \qquad \text{for any arbitrary } w_0 \in \mathbb{C}^+. \tag{4.16}$$

Moreover, the resulting fixed point function $z \mapsto w(z)$ is analytic.

These results allow to calculate the free additive convolution $\mu \boxplus \nu$ of any two real-valued probability measures μ and ν. Since the problem of computing joint probability functions of (asymptotically) free random variables boils down to determining the convolution of the empirical spectral measures, we have reached a cornerstone for answering the question, how spectra of random matrices are added up. In general, we cannot use the above machinery analytically. However, a numerical fixed point iteration procedure is straightforward to implement [BB07, Spe19b]. In the next section, we present a concrete fixed point algorithm, from which we can compute arbitrary linear combinations of random matrices that are asymptotically freely independent.

4.2. The Numerical Subordination Algorithm

Here, we present a numerical subordination algorithm resulting from Theorem 4.1. The core of the algorithm is the fixed point method in Theorem 4.1. It is emphasized that there might be several other (numerical) ways to determine the fixed point in (4.16). Like any numerical algorithm, the tuning of the initial parameters has significant relevance for the results. Since a concrete realization of the implicit numerical algorithm in Theorem 4.1 is rarely described in the literature, we explain some details in the following. Moreover, we implement Algorithm 1, which does the job.

4.2.1. Working Principle

Before presenting numerical details, we summarize the main steps of the Subordination Iteration Algorithm. According to Theorem 4.1, the algorithm operates as follows:

1. Compute the Cauchy transforms G_μ and G_ν of the measures μ and ν.
2. Define the fixed point equation (4.15).
3. For each $z \in \mathbb{C}^+$ alongside the iteration domain, iterate (4.15) in the $w-$component.
4. Step 3 yields the map $w(z)$, which is inserted into G_μ, so that $G_{\mu\boxplus\nu}(z) = G_\mu(w(z))$.
5. Recover the free additive convolution $\mu \boxplus \nu$ from $G_\mu(w(z))$ via Stieltjes inversion formula (3.24).

4.2.2. Choice of Parameters

Although one has some flexibility in setting the numerical parameters, we present a suitable choice of values. The probability measures μ and ν are obligatory to input. In case of adding Hermitian matrices, these correspond to the (normalized) list of their eigenvalues. However, there are other numerical parameters, which rather correspond to fine tuning. Unless the problem specific application requires modifications in these parameters, we choose them as presented in Algorithm 1. In the following, we present a list of the involved numerical parameters. If the

α	conergence accuracy
ε	shift into the upper complex half plane
M	maximal number of iterations (in the $w-$component)
N	number of equidistant steps fot the interval partitioning
w_0	initial value for the fixed point iteration

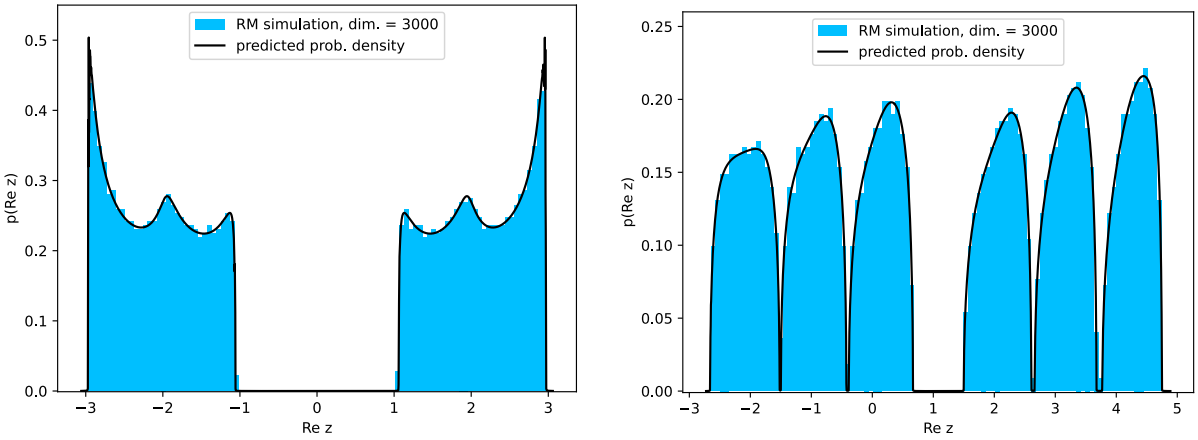
Algorithm 1 `Sub` $(\mu, \nu, \mathcal{D})$

Input: μ, ν : probability measures on $\mathbb{R}$ **and** $\mathcal{D} := \{t_0, ..., t_{D-1}\} \subset \mathbb{R}$: iteration domain
Input: **(optional)** $\alpha, \varepsilon, M, w_0$
  $\alpha \leftarrow 10^{-8}$ ▷ default values
  $\varepsilon \leftarrow 10^{-6}$
  $M \leftarrow 400$ ▷ maximal number of iterations
  $w_0 \leftarrow 0.5$
  $N \leftarrow \text{length}(\mathcal{D})$
  $\tilde{\mathcal{D}} := \{t + i\varepsilon | t \in \mathcal{D}\}$ ▷ stay ε - close to $\mathcal{D}$ for later Stieltjes inversion
  Define $G_\mu(z)$, $G_\nu(z)$ for $z \in \tilde{\mathcal{D}}$. ▷ compute Cauchy transforms to define $g(z, w)$
  Define $g(z, w)$ according to eq. (4.15).
  for $j \in \{0, ..., N-1\}$ **do** ▷ iterate through domain $\tilde{\mathcal{D}}$
    $z_j \leftarrow t_j + i\varepsilon$ ▷ put z_j in $\tilde{\mathcal{D}}$
    $w_j \leftarrow t_j + iw_0$ ▷ initial guess value on $\mathbb{C}^+$
    for $k \in \{0, ..., M-1\}$ **do**
      if $|g(z_j, w_k) - g(z_j, w_{k-1})| \leq \alpha$ **then** ▷ check, if convergence accuracy is reached
        store $w_{k_{\text{thresh}}}$ ▷ fixed point $w_{k_{\text{thresh}}} \equiv w(z_j)$
      end if
      compute $\mu \boxplus \nu\,(t_j) \leftarrow \frac{-1}{\pi} \text{Im}\, G_\mu\,(w_{k_{\text{thresh}}})$ ▷ density via Stieltjes inversion
    end for
  end for
Output: $(\mathcal{D}, \mu \boxplus \nu(\mathcal{D}))$ ▷ outputs the graph of the joint density on the domain $\mathcal{D}$.

number of distinct eigenvalues (i.e. not counting the algebraic multiplicity) l_i, $i \in \{\mu, \nu\}$ is not too large, the Subordination Algorithm mainly depends on the numerical parameters N and M. Assuming that $l_i \propto 10^1$, the Subordination Algorithm converges in $\mathcal{O}(NM)$ operations.

4.2.3. Examples

Here, we apply the results from Chapter 3.7 that essential for Chapter 5. We consider examples for two real-valued deterministic, diagonal (thus self-adjoint) matrices A, B, where $B' = \Phi_U(B)$ is unitarily rotated at random against A. Theorem 3.24 and Theorem 3.25 ensure that A and



(a) RM simulation (blue) vs. $\mu_{a_1} \boxplus \mu_{b_1}(t)$ (black) (b) RM simulation (blue) vs. $\mu_{a_2} \boxplus \mu_{b_2}(t)$ (black)

Figure 4.1.: Example 4.2 and 4.3 of the free additive convolution for spectral distributions of self-adjoint operators A_i and B_i, $i = 1, 2$. We generate a random unitary matrix of dimension $d = 3000$ by using the SciPy routine `scipy.stats.unitary_group` [Doc23a]. We generate the matrices from Example 4.2 and 4.3 and generate histograms which show the spectral density of the function $f(A_i, B_i, U)$, for f as in (4.17). We compare the random-matrix simulation (blue histograms) with the results, predicted by running `Sub` (Algorithm 1) with the densities from Example 4.2 Example 4.3.

B are asymptotically free. Thus, we can compute the probability density over $\sigma(A + B')$ using the subordination algorithm to calculate the free convolution of the spectral measures. The procedure is as follows.

1. $A = \mathrm{diag}\,(a_1, ..., a_d), \qquad B = \mathrm{diag}\,(b_1, ..., b_d),$
2. $\mu_A = \frac{1}{d}(a_1, ..., a_d), \qquad \mu_B = \frac{1}{d}(b_1, ..., b_d)$
3. Define $\mathcal{D} = \{t_0, ..., t_{N-1}\}$, where $t_0 := \inf\{a_1, b_1\}, \quad t_{N-1} := \sup\{a_d, b_d\}$
4. Run `Sub` $(\mu_A, \mu_B, \mathcal{D})$ (Algorithm 1).
5. $\mu_A \boxplus \mu_B(\mathcal{D}) =$ `Sub` $(\mu_A, \mu_B, \mathcal{D})$.

It is mentioned that the eigenvalues of matrices A and B are counted with multiplicity, i.e. $a_1 \leq a_2 \leq ... \leq a_d$ and similarly for B. The iteration domain $\mathcal{D}$ is hence an interval partitioning between t_0 and t_N. The finer this partitioning is chosen, the more accurate becomes the outputted distribution. We present results for $A, B \in M_d(\mathbb{R})$ of dimension $d = 3000$. Figure 4.1 shows plots of the subsequent examples. We compare those results that are predicted by Algorithm 1 (black curve) with random matrix simulations of size 3000×3000 (blue histogram). This convention is kept throughout the thesis.

Example 4.2. *This first example is visualized in Figure 4.1a.*

$$A_1 = \mathrm{diag}(-2^{\times 1500}, 2^{\times 1500}), \qquad B_1 = \mathrm{diag}(-1^{\times 1000}, 0^{\times 1000}, 1^{\times 1000}),$$

where the notation $-2^{\times 1500}$ *means that the eigenvalue* -2 *has an algebraic multiplicity of* 1500. *By construction, the operator sequences* $(A_1)_{d\in\mathbb{N}} := \operatorname{diag}(-2^{\frac{d}{2}}, 2^{\frac{d}{2}})$ *and* $(B_1)_{d\in\mathbb{N}} := \operatorname{diag}(-1^{\frac{d}{3}}, 0^{\frac{d}{3}}, 1^{\frac{d}{3}})$ *converge to the following probability densities*

$$A_1 \xrightarrow{distr.} a_1(t) := \frac{1}{2}\left(\delta(t+2) + \delta(t-2)\right),$$
$$B_1 \xrightarrow{distr.} b_1(t) := \frac{1}{3}\left(\delta(t+1) + \delta(t) + \delta(t-1)\right).$$

We then generate a random unitary matrix $U \in \mathcal{U}(d)$, *for* $d = 3000$ *and define the function*

$$f(A, B, U) := A + UBU^* \tag{4.17}$$

Furthermore, we simulate the spectral distribution of $f(A_1, B_1, U)$ *and compare it to* $\mu_{a_1} \boxplus \mu_{b_1}(t)$, *where* $t \in \mathcal{D}$,

$$\mathcal{D} := \mathbb{R} \cap [\inf\{\sigma(A) \cup \sigma(B)\}, \sup\{\sigma(A) \cup \sigma(B)\}].$$

$\mu_{a_1} \boxplus \mu_{b_1}(t)$ *is computed by running Sub* $(\mu_{a_1}, \mu_{b_1}, \mathcal{D})$.

Example 4.3. *This second example is visualized in Figure 4.1b.*

$$A_2 = \operatorname{diag}(-2^{\times 500}, -1^{\times 500}, 0^{\times 500}, 2^{\times 500}, 3^{\times 500}, 4^{\times 500}),$$
$$B_2 = \operatorname{diag}(-1^{\times 750}, 0^{\times 750}, 0.5^{\times 750}, 1^{\times 750}),$$

where we take the same f *as in (4.17), generate the histogram for* $f(A_2, B_2, U)$ *and compare it to* $\mu_{a_2} \boxplus \mu_{b_2}(t)$ *with* $t \in \mathcal{D}$ *as above.* $\mu_{a_2} \boxplus \mu_{b_2}(t)$ *is computed by running Sub* $(\mu_{a_2}, \mu_{b_2}, \mathcal{D})$.

Figure 4.1 verifies numerically, what Voiculescu's results (Theorem 3.24 and Theorem 3.25) predict. Randomly rotating B_i compared to A_i, $i \in \{1, 2\}$ generates asymptotic freeness. Then, we can apply subordination theory to determine the joint spectral distribution. We observe the similar accuracy between theory and simulation as for the semi-circle law (Figure 3.1).

For simplicity, we choose A, B to be diagonal. However, this is not necessary. If A and B are non-diagonal, one must diagonalize A and B at first, in order to receive the spectral measures. In this case, Algorithm 1 takes with the diagonal matrices corresponding to A and B. Due to self-adjointness, the spectral theorem guarantees that there are exactly d real eigenvalues of A and B so that no problem with the normalization of μ_A and μ_B arises here.

We extensively use this subordination theory in Chapter 5 to simulate spectra of quantum adiabatic Hamiltonians. The scalar-valued subordination framework allows to compute linear combinations of matrices. We summarize this in the following Lemma.

Lemma 4.4. *Let* $d \in \mathbb{N}$ *and* $A, B \in M_d(\mathbb{C})$ *self-adjoint. Let* D_A *and* D_B *be the corresponding diagonal matrices. Moreover, let* $D'_B := \Phi_U(D_B)$ *be the matrix, which is unitarily transformed at random. Then, for any* $\alpha, \beta \in \mathbb{R}$, $\tilde{D}_A = \alpha D_A$ *and* $\tilde{D}_B = \beta D'_B$ *are almost surely asymptotically free.*

Proof. This directly follows from Theorem 3.25 by applying the theorem to $\tilde{D}_A$ and $\tilde{D}_B$. □

The above Lemma emphasizes that linear (and consequently convex) combinations fall into the framework of scalar-valued subordination. Therefore, scalar-valued free convolution suffices to treat QAA Hamiltonians (see Chapter 5). Although free probability relates to random-matrix applications, there is not as much randomness involved as one would suggest due to the name. The only randomness that infects the above examples comes from a basis randomization between

distinct eigenspaces. This relative randomness between A and B generates freeness and activates the powerful free convolution machinery. The numerical costs which are needed to determine those convolutions, reduces significantly compared to diagonalization routines for large matrices.

However, subordination theory has been developed much further in the context of free probability [BMS17] [Spe20a]. We present some modern developments of this field in the remainder of this Chapter.

4.3. More General Functions in Random-Matrix-Variables

So far, the subordination algorithm covers linear combinations of free variables. We take these free variables to be random operators on a (finite-dimensional) Hilbert space. In light of the previous section, we emphasize that constant matrices play the analogous role of constant functions in classical probability theory. Consider the following problem: instead of computing the spectral convolution of functions of the form,

$$f(A,B) := \alpha A + \beta B, \qquad \alpha, \beta \in \mathbb{R}, \quad A, B \in \mathcal{B}(\mathcal{H}), \tag{4.18}$$

can we possibly compute spectral convolutions of functions of more general *random-matrix variables,* like polynomials? The answer is yes! The first result was provided by Haagerup et al. [HT05], who showed that for a self-adjoint polynomial P in k GUE(N) random matrices,

$$\sigma\left(P\left(X_1^N, ..., X_k^N\right)\right) \subset \operatorname{supp} \mu_p + (-\varepsilon, \varepsilon), \qquad \varepsilon > 0, \tag{4.19}$$

where μ_P corresponds to the measure of the polynomial P. In words: the spectral measure of the polynomial in random-matrix variables is concentrated on its support [Par22]. This has been refined for other types of random matrices ever since [Par22]. However, the theoretical framework of dealing with more complicated functions becomes substantially more involved. Nevertheless, it can be done within the theory of operator-valued free probability [Jek18], [Wil17b], [BMS17], [MS17, Chapter 9-10]. A motivating example is the theory of block random matrices [MS17, Chapter 9.1]. Consider a sequence of random matrices within a self-adjoint block structure, i.e.

$$X_N = \frac{1}{\sqrt{3}} \begin{pmatrix} A_N & B_N & C_N \\ B_N & A_N & B_N \\ C_N & B_N & A_N \end{pmatrix}, \qquad A_N, B_N, C_N \in \mathrm{GUE}\,(N). \tag{4.20}$$

Note that such a block structure does not give a GUE(3N) matrix. This would only be true for $N = 1$, but the symmetry (due to self-adjointness) distinguishes a GUE($3 \cdot (2 \times 2)$) from a GUE(6) matrix. Numerical experiments show that X_N converges also to deterministic limiting distribution that corresponds to a deformed semicircle (see e.g. Figure 9.1 in [MS17]), but not the usual semicircular shape (Figure 3.1) any more. This gives rise to the assumption that a hidden, deeper theory predicts such a modified central limiting object. We summarize some results from *operator-valued free probability theory* in the following. Roughly speaking, the block matrix (4.20) is a random matrix consisting of random matrices (or operators), which yields the name "operator-valued" free probability.

The central object in scalar-valued free probability theory is a non-commutative probability space (see Definition 2.11). Non-commutative probability spaces are generalized to the operator-valued setting, as follows [Spe20a, Definition 2.4][1].

[1] supplemented by online lecture videos [Spe20b]

Definition 4.5. *An **operator-valued non-commutative probability space** is a triple $(\mathcal{A}, \mathcal{B}, E)$, where $\mathcal{A}$ is a unital Banach algebra, $\mathcal{B} \subset \mathcal{A}$ is unital subalgebra $(1_{\mathcal{A}} \in \mathcal{B})$ and $E \in \mathcal{L}(\mathcal{A}, \mathcal{B})$ is a conditional expectation[2], i.e.*

- $E(1_{\mathcal{A}}) = 1_{\mathcal{A}}$
- *E has the bimodule property:*

$$E[b_1 A b_2] = b_1 E[A] b_2, \qquad b_1, b_2 \in \mathcal{B}, \quad A \in \mathcal{A} \tag{4.21}$$

and especially $E[b] = b$ for all $b \in \mathcal{B}$.

Any modified assumption about the topology of $\mathcal{A}$ strengthens the space topologically, e.g. if $\mathcal{A}$ is a C^* algebra and E is positive, then $(\mathcal{A}, \mathcal{B}, E)$ is an operator-valued (non-commutative) C^* probability space. E being positive means, that there exists some $b \in \mathcal{B}$ such that $E[A^*A] = b^*b$. Due to Chapter 3.5, the scalar-valued Cauchy transform is $G_{\mu_A}(z) = \varphi\left((z\mathbb{1} - A)^{-1}\right)$. Starting with a scalar-valued C^*-NCPS, $(\tilde{\mathcal{A}}, \varphi)$, the operator-valued space is constructed by taking $\mathcal{A} := M_n(\mathbb{C}) \otimes \tilde{\mathcal{A}} \cong M_n(\tilde{\mathcal{A}})$, $E := \mathrm{id}_{M_n(\mathbb{C})} \otimes \varphi$ and $\mathcal{B} := M_n(\mathbb{C}) \subset \mathcal{A}$. In the previous example (4.20), $n = 3$ and $\tilde{\mathcal{A}}$ is the algebra of GUE(N) matrices with $\varphi := \mathrm{tr} \otimes \mathbb{E}$. Then,

$$E[X_N] = \frac{1}{\sqrt{3}} \begin{pmatrix} \varphi(A_N) & \varphi(B_N) & \varphi(C_N) \\ \varphi(B_N) & \varphi(A_N) & \varphi(B_N) \\ \varphi(C_N) & \varphi(B_N) & \varphi(A_N) \end{pmatrix} \in M_3(\mathbb{C}).$$

We note that $M_n(\mathbb{C}) \subset M_n(\tilde{\mathcal{A}})$ is indeed a subalgebra. Assuming that the Banach space inversion of $(b - X)$ exists, the operator valued analogue is [Spe20a, Def. 2.5]:

Definition 4.6. *Let $(\mathcal{A}, \mathcal{B}, E)$ be an operator-valued C^* probability space and $X = X^* \in \mathcal{A}$. Then, the **operator-valued Cauchy transform** $G_X : \mathcal{B} \to \mathcal{B}$ is defined by*

$$G_X(b) := E\left[(b - X)^{-1}\right]. \tag{4.22}$$

Moreover, $G_X(\cdot)$ is a Fréchet analytic function [BMS17]. Definition 4.6 is related to non-commutative functions. This specific class of functions is also known as the class of *fully matricial functions*. The key concept of fully matricial functions is a non-commutative generalization of the differential quotient for holomorphic functions, where $\mathbb{C}$ is replaced by the Banach algebra $\mathcal{B}$ [Voi10]. For topological details regarding these generalized Cauchy transforms, we refer to Liu [Liu21]. Returning to the block matrix from (4.20), we have to specify the notion of free independence (Definition 3.13) in the operator-valued setting [Spe20a, Chapter 5].

Definition 4.7. *Let $(\mathcal{A}, \mathcal{B}, E)$ be an operator-valued NCPS and let I be a finite index set. Subalgebras $\mathcal{B} \subset \mathcal{A}_i \subset \mathcal{A}$ $(i \in I)$ are called **free,** if for $k \in \mathbb{N}$, $E[a_1 \cdots a_k] = 0$, whenever*

- *$a_j \in \mathcal{A}_{i(j)}$, $i(j) \in I$ for all $j \in [k]$,*
- *$i(1) \neq i(2) \neq \ldots \neq i(k)$,*
- *$E[a_j] = 0$, for all $j \in [k]$.*

[2] to avoid misunderstandings, we denote the space of linear operators here by $\mathcal{L}(\cdot)$, not $\mathcal{B}(\cdot)$ as before.

Similarly to the scalar-valued setting, subalgebras $\mathcal{A}_i$ are the generated $\mathcal{B}-$subalgebras $\mathcal{A}_i = \text{alg}\,(\mathcal{B}, x_i)$. This generalized type of free independence is also called *freeness with amalgamation.* By construction, operator-valued moments amalgamate with respect to E; e.g. the third moment would be $m_3(X) = E[Xb_1Xb_2X]$ for $b_1, b_2 \in \mathcal{B}$. Moreover, the combinatorial theory of freeness survives in the operator-valued framework as well. The lattice of non-crossing pairings arises - similarly to the scalar-valued case - in the factorization of mixed moments of operator-valued free variables. Hence, the collection of all $\mathcal{B}-$valued moments forms the *operator-valued distribution* μ_X of $X \in \mathcal{A}$.

Let $P(X, Y)$ be a polynomial of two self-adjoint operators $X, Y \in \mathcal{A}$. Since the subordination formalism can deal with the convolution of two probability measures, the question remains, how to separate $P(X, Y)$ in two (!) variables $\tilde{X}, \tilde{Y}$ such that $\mu_{P(X,Y)} \equiv \mu_{\tilde{X}} \boxplus \mu_{\tilde{Y}}$. Only then, we can apply the operator-valued analogue of the subordination algorithm. This leads to the theory of **Linearization by enlargement** [Spe15], [HMS18].

4.3.1. Linearization and Operator-Valued Free Convolution

We consider self-adjoint polynomials in random-matrix-variables [Spe15], [Spe20a, Chapter 8]. In order to apply subordination theory, we will linearize those polynomials to a higher dimensional operator algebra. A linearization is basically a block matrix such that a suitable multiplication of these blocks leads to the original polynomial.

Definition 4.8 (Linearization)**.** *Given a polynomial $P \in \mathbb{C}\langle x_1, ..., x_k\rangle$, the block matrix*

$$\hat{P} = \begin{pmatrix} 0 & u \\ v & q \end{pmatrix} \in M_n(\mathbb{C}\langle x_1, ..., x_k\rangle), \qquad n, k \in \mathbb{N}, \tag{4.23}$$

where

- *$q \in M_{n-1}(\mathbb{C}\langle x_1, ..., x_k\rangle)$ is invertible as a matrix over polynomials,*
- *$u \in M_{1,n}(\mathbb{C}\langle x_1, ..., x_k\rangle)$ is a row and*
- *$v \in M_{n,1}(\mathbb{C}\langle x_1, ..., x_k\rangle)$ is column,*

*is a **linearization** of P, if*

1. *$\hat{P}$ is an affine matrix in $x_1, ..., x_k$, i.e., there are $b_0, b_1, ..., b_k \in M_n(\mathbb{C})$ such that*

$$\hat{P} = b_0 \otimes \mathbb{1} + \sum_{i=1}^{k} b_i \otimes x_i, \tag{4.24}$$

2. *$P = -uq^{-1}v$.*

The fact that $uq^{-1}v$ actually exists, is due to the Schur complement formula [BMS17, Proposition 3.2]. The existence of a linearization for self-adjoint polynomials is provided by the next theorem [Spe20a, Theorem 8.5].

Theorem 4.9. *There exists a linearization $\hat{P}$ for any polynomial $P \in \mathbb{C}\langle x_1, ..., x_k\rangle$. If P is self-adjoint, then $\hat{P}$ is self-adjoint, as well.*

Finally, we present [BMS17, Theorem 2.2], which generalizes Theorem 4.1 to a subordination theorem for the operator-valued free additive convolution. Moreover, we discuss the resulting subordination algorithm to compute the additive convolution of polynomials in random-matrix-variables.

Theorem 4.10 (Belinschi et al., 2017)**.** *Let $(\mathcal{A}, \mathcal{B}, E)$ be an operator-valued C^* probability space and let $X_1, X_2 \in \mathcal{A}$ be self-adjoint and free with respect to E. Then, there exists a unique pair of Fréchet analytic maps,*

$$\omega_1, \omega_2 : \mathcal{B}^+ \to \mathcal{B}^+,$$

such that

1. *Im $\omega_i(z) \geq$ Im z, for all $z \in \mathcal{B}^+$, $(i = 1, 2)$,*
2. *for all $z \in \mathcal{B}^+$,*

$$G_1^{-1}(\omega_1(z)) + z = G_2^{-1}(\omega_2(z)) + z = \omega_1(z) + \omega_2(z), \tag{4.25}$$

3. *for all $z \in \mathcal{B}^+$,*

$$G_1(\omega_1(z)) = G_2(\omega_2(z)) = G(z) \tag{4.26}$$

Moreover, for $z \in \mathcal{B}^+$ fixed, $\omega_1(z)$ is the unique fixed point of $f(z, \cdot) : \mathcal{B}^+ \to \mathcal{B}^+$, given by

$$f(z, w) := w + G_2^{-1}\left(z - w + G_1^{-1}(w)\right) - G_1^{-1}(w), \tag{4.27}$$

so that $\omega_1(z) = \lim_{n\to\infty} f^{\circ n}(z, w_0)$ for any arbitrary $w_0 \in \mathcal{B}^+$.

The similarity between Theorem 4.1 and Theorem 4.10 is clearly visible. As it was shown [BMS17], the following algorithm allows to compute the asymptotic spectral distribution of a self-adjoint polynomial in random-matrix variables. Theorem 4.10, together with the linearization from Definition 4.8 build an algorithm for operator-valued free additive convolution.

Theorem 4.11. *[BMS17, Thm. 4.1] Let P be a self-adjoint polynomial in k non-commuting indeterminates x_i. The following algorithm leads to the distribution of P.*

1. *Due to Theorem 4.9, P admits a linearization, i.e. there exist $b_i \in M_n(\mathbb{C})$ such that*

$$\hat{P} = b_0 \otimes \mathbb{1} + \sum_{i=1}^{k} b_i \otimes x_i, \quad \in M_n(\mathbb{C}) \otimes \mathcal{A}. \tag{4.28}$$

2. *The operators $b_i \otimes x_i$ are free in the op-val. C^*-NCPS $(M_n(\mathcal{A}), M_n(\mathbb{C}), E)$ with $E := \mathrm{id}_{M_n(\mathbb{C})} \otimes \varphi$, where $(\mathcal{A}, \varphi)$ is a scalar-valued C^*-NCPS.*
3. *For each $i \in [k]$, the operator-valued Cauchy transform is completely determined by the scalar-valued Cauchy transform via the following Bochner integral representation:*

$$G_{b_i \otimes x_i}(b) = \lim_{\varepsilon \to 0^+} \frac{-1}{\pi} \int_{\mathbb{R}} (b - tb_i)^{-1} Im\left(G_{x_i}(t + i\varepsilon)\right) dt \tag{4.29}$$

   for $b \in \mathcal{B}^+ \equiv M_n^+(\mathbb{C})$. Note that $G_{\hat{P}}(b) = G_{\hat{P} - b_0 \otimes \mathbb{1}}(b - b_0)$.
4. *The scalar-valued Cauchy transform is given by $G_{\mu_P}(z) = \lim_{\varepsilon \to 0^+} \{G_{\hat{P}}(\Lambda_\varepsilon(z))\}_{1,1}$*

5. *Recover the joint density from $G_{\mu_P}(z)$ via Stieltjes inversion (3.24).*

A few remarks on Theorem 4.11: $\Lambda_\varepsilon(z)$ corresponds to an ε-shift into the upper half plane of $\mathcal{B}$, i.e. (e.g. for $n = 3$)

$$\Lambda_\varepsilon(z) = \begin{pmatrix} z & 0 & 0 \\ 0 & i\varepsilon & 0 \\ 0 & 0 & i\varepsilon \end{pmatrix}$$

Both in the scalar- and the operator-valued setting, we iterate alongside the real line, but the Cauchy transform is only defined on the upper half plane. Furthermore, the evaluation of a Bochner integral of the form (4.29) is numerically challenging, if we aim to avoid slowly converging Riemann sums. The expression in (4.29) can be evaluated using J.W. Helton's reduction method[3] (see [Mai17, Algorithm I.2.11]). Moreover,

$$G_{\hat{P}+b_0}(b) \overset{\text{def.}}{=} E\left((b - (b_0 + \hat{P}))^{-1}\right) = E\left(((b - b_0) - \hat{P})^{-1}\right) = G_{\hat{P}}(b - b_0).$$

For a successful implementation of the algorithm in Theorem 4.11, we refer to [HM23].

[3] Due to personal correspondence with Tobias Mai and Roland Speicher

5. Quantum Adiabatic Spectra via Free Probabilistic Subordination

Free Probability yields mathematical tools that allow to compute the additive convolution of probability measures. Provided that Hermitian operators are asymptotically freely independent, the subordination formalism can be transferred to determine the joint distribution of the corresponding spectral measures. In this Chapter, we apply the (scalar-valued) Subordination Iteration Algorithm 1 to convex combinations of self-adjoint operators. We choose those operators, which are suitable for QAA [FGGS00]. By doing so, we modify the setting of adiabatic quantum computation, which we will call *free adiabatic quantum computation*. The difference is a randomized, unitary basis transformation, that is applied to the problem Hamiltonian. As it is discussed in Chapter 3, this generic operation generates asymptotic freeness between the initial and the final Hamiltonian. Furthermore, we extend our approach by developing a spectral gap estimating algorithm (FASGE), which we apply to the free adiabatic Grover problem.

5.1. Fusing the Quantum Adiabatic Evolution with Freeness

We have developed the main results about adding the asymptotic spectral distributions of random matrices in the previous two chapters. It remains to discuss, how this approach is transferable to the quantum adiabatic algorithm [FGGS00]. Again, we neglect physical dimensions and simplify the adiabatic Hamilton operator $H(\cdot) \in C^2([0,1], \mathcal{B}(\mathcal{H}))$ for $A, B \in \mathcal{B}(\mathcal{H})$ to

$$H(s) = (1-s)A + sB, \qquad s \in [0,1]. \tag{5.1}$$

such that $H(0) = A$ and $H(1) = B$. We will restrict our simulations to the standard (linear) schedule, i.e. $\alpha(s) := 1 - s$ and $\beta(s) := s$. There are approaches [AL18], in which

$$H(s) := \alpha(s)A + \beta(s)B, \qquad \alpha(s) + \beta(s) = 1, \tag{5.2}$$

and $\alpha(\cdot), \beta(\cdot)$ being more complicated functions to tune the interpolation velocity suited to the respective problem under investigation. In order to insert freeness into the QAA, we apply a random unitary basis transformation to B, i.e.

$$B \mapsto B' := \Phi_U(B), \qquad \Phi_U(\cdot) := U(\cdot)U^*, \qquad U \in \mathcal{U}(d). \tag{5.3}$$

$\mathcal{U}(d)$ denotes the set of unitary random matrices with Haar distribution on the unitary group $\mathfrak{U}(d)$. We take QAA operators that are self-adjoint and positive semi-definite; i.e. the corresponding C^* algebra $\mathcal{B}(\mathcal{H})$ forms a convex cone. Assume that $a = \lim_{d\to\infty} A_d$ and $b = \lim_{d\to\infty} B_d$, for $(A_d)_{d\in\mathbb{N}}, (B_d)_{d\in\mathbb{N}}$ sequences of deterministic $d \times d$ matrices.

Due to Chapter 3, applying $\Phi_U(\cdot)$ to B, makes the operator sequences $(A_d)_{d\in\mathbb{N}}, (B_d)_{d\in\mathbb{N}}$ for $d \to \infty$ freely independent. Let $(\mathcal{A}, \varphi)$ be the C^*-NCPS with $a, b \in \mathcal{A}$. Thus, we get a limit operator of the form $h(\cdot) \in C^2([0,1], \mathcal{A})$

$$h(s) := (1-s)a + sb, \qquad s \in [0,1]. \tag{5.4}$$

Since taking the limit (w.r.t the C^* norm topology) is a closed operation within the cone of positive operators, the limit operator (5.4) exists indeed. However, our theory only holds true in the asymptotic limit. If we aim to study the finite-dimensional analogue of (5.4), i.e.

$$H(s) = (1-s)A + sB', \tag{5.5}$$

we have to take the deviation from the asymptotic limit into account. Nevertheless, our empirical examples (see Figure 4.1) demonstrate that even for $d \propto 10^3$, the simulation gets tremendously close to the limiting distribution. Thus, we have to specify "how far away" are finite-dimensional realizations compared to the asymptotic situation? In order to answer this question, we apply results from [Kar12] and refer to [GM20, Par22, Par21] for modified bounds on the operator norm that also respect the operator-valued extension of freeness. Guionnet et al. [GM20] derive a large deviation principle for the law of the largest eigenvalue of $A + UBU^*$. Parraud [Par22] analyzes deviations of the operator norm of non-commutative polynomials in deterministic and i.i.d Haar unitary matrices from their asymptotic limiting distribution. However, such quantities have been studied earlier on in the context of second-order freeness [MŚS07]. As we will see, our model of free adiabatic quantum computation can be controlled by using Kargin's bounds [Kar12] on the quantity $A + UBU^*$. We provide errors (up to constant factors) which control, how much the finite-dimensional random matrix model deviates from the free probabilistic limit.

5.1.1. Bounds on Deviations from Limiting Distributions

We consider deterministic Hermitian matrices A_N, B_N of large dimension $N \in \mathbb{N}$, and random unitary matrices U_N and study the quantity

$$H_N = A_N + U_N B_N U_N^*. \tag{5.6}$$

Our model of free adiabatic quantum computation will be

$$H_{N,s} = D_{A_N,s} + U_N D_{B_N,s} U_N^*, \tag{5.7}$$

where $D_{A_N,s} := \alpha(s) D_{A_N}$ and $D_{B_N,s} := \beta(s) D_{B_N}$. Hence, it is a canonical, parameterized application of (5.6). In the next theorem, $\mathcal{F}_{H_N}(t) := \mu_{H_N}(-\infty, t], t \in \mathbb{R}$ denotes the cumulative distribution function (CDF) for eigenvalues of H_N over the respective interval. Under the assumption that the measure $\mu_{A_N} \boxplus \mu_{B_N}$ is absolutely continuous everywhere on $\mathbb{R}$, and its density is bounded by a constant T_N, the following theorem holds true [Kar12, Theorem 1].

Theorem 5.1. *Let $\mathcal{F}_{H_N}$ be the CDF for $H_N = A_N + U_N B_N U_N^*$ and $\mathcal{F}_{\boxplus,N}$ be the CDF for $\mu_{A_N} \boxplus \mu_{B_N}$. Then, for all $N \geq \exp\left((\frac{c_1}{\delta})^{\frac{4}{\varepsilon}}\right)$,*

$$\mathbb{P}\left\{\sup_{t\in\mathbb{R}} |\mathcal{F}_{H_N}(t) - \mathcal{F}_{\boxplus,N}(t)| > \delta\right\} \leq \exp\left(-c_2 \delta^2 N^2 \ln^{-\varepsilon}(N)\right), \tag{5.8}$$

where $c_1, c_2 > 0$ and depend only on $K_N := \max\{\|A_N\|, \|B_N\|\}$, T_N, and $\varepsilon \in (0, 2]$.

Lemma 5.2. *By taking $H_{N,s} = D_{A_N,s} + U_N D_{B_N,s} U_N^*$ to be the parameterized, self-adjoint convex combination from (5.7) and applying Theorem 5.1, we have the same bound on the cumulative distribution function, i.e. for all $N \geq \exp\left((\frac{c_1}{\delta})^{\frac{4}{\varepsilon}}\right)$*

$$\mathbb{P}\left\{\sup_{t\in\mathbb{R}} |\mathcal{F}_{H_N,s}(t) - \mathcal{F}_{\boxplus,N,s}(t)| > \delta\right\} \leq \exp\left(-c_2 \delta^2 N^2 \ln^{-\varepsilon}(N)\right), \tag{5.9}$$

where $c_1, c_2 > 0$ depend only on $K_{N,s} := \max\{\|D_{A_N,s}\|, \|D_{B_N,s}\|\}$ (and T_N, ε as in Thm. 5.1).

Proof. We directly apply Theorem 5.1 to the diagonal matrices $A_N := D_{A_N}$ and $B_N := D_{B_N}$ to find the exact same bounds. Now, we choose these diagonal matrices to be parameterized by s, i.e. $A_N := D_{A_N,s}$ and $B_N := D_{B_N,s}$, defined by

$$D_{A_N,s} := \alpha(s) D_{A_N}, \qquad D_{B_N,s} := \beta(s) D_{B_N}, \qquad \alpha(\cdot), \beta(\cdot) \in C^0([0,1],[0,1]).$$

Since $|\alpha(s)|, |\beta(s)| \leq 1$ for all $s \in [0,1]$, we especially have that $\|D_{A_N,s}\| = \|\alpha(s) D_{A_N}\| = |\alpha(s)| \|D_{A_N}\| \leq \|D_{A_N}\|$ and similarly for D_{B_N}. Hence, $K_{N,s} \leq K_N$ and T_N, ε can be chosen to be the same constants as in Theorem 5.1 [Kar12]. □

Since Theorem 5.1 and the immediate implication to our problem (Lemma 5.2) deal with the cumulative distribution function, we proceed with the subsequent Theorem 5.3 (adapted from [Kar12, Theorem 2]) and its Lemma to find a bound on the probability density function.

Theorem 5.3. *Let $\mathcal{N}_\eta(E) := \mathcal{N}_{(E-\eta,E+\eta]}$ denote the number of eigenvalues of H_N in an interval of width 2η centered at E. and let $\varrho_{\boxplus,N}(E)$ denote the density of $\mu_{A_N} \boxplus \mu_{B_N}$ at E. Suppose that $\eta = \eta(N)$ and $\frac{1}{\sqrt{\log(N)}} \ll \eta \ll 1$. Let the same assumption as in Theorem 5.1 hold with $T_N = T$. Assume that $K_N = \max\{\|A_N\|, \|B_N\|\} \leq K$ for all N. Then, for all sufficiently large N,*

$$\mathbb{P}\left\{ \sup_{E\in\mathbb{R}} \left| \frac{\mathcal{N}_\eta(E)}{2N\eta} - \varrho_{\boxplus,N}(E) \right| \geq \delta \right\} \leq \exp\left(-c\delta^2 \frac{(\eta N)^2}{(\log N)^2} \right), \tag{5.10}$$

where $c > 0$ depends only on K and T.

As a direct consequence, we have

Lemma 5.4. *By taking $H_{N,s} = D_{A_N,s} + U_N D_{B_N,s} U_N^*$ to be the parameterized, self-adjoint convex combination from (5.7), we have the following. Let $\mathcal{N}_\eta(E_s) := \mathcal{N}_{(E_s-\eta,E_s+\eta]}$ denote the number of eigenvalues of $H_{N,s}$ in an interval of width 2η centered at E_s, and let $\varrho_{\boxplus,N}(E_s)$ denote the density of $\mu_{D_{A_N},s} \boxplus \mu_{D_{B_N},s}$ at E_s for all $s \in [0,1]$. Let the same assumptions hold as in Theorem 5.3 and assume that $K_{N,s} = \max\{\|D_{A_N,s}\|, \|D_{B_N,s}\|\} \leq K_s$ for all N and for all s. Then, for all sufficiently large N and for each $s \in [0,1]$,*

$$\mathbb{P}\left\{ \sup_{E_s\in\mathbb{R}} \left| \frac{\mathcal{N}_\eta(E_s)}{2N\eta} - \varrho_{\boxplus,N}(E_s) \right| \geq \delta_s \right\} \leq \exp\left(-c\delta_s^2 \frac{(\eta N)^2}{(\log N)^2} \right), \tag{5.11}$$

where $c > 0$ depends only on K_s and T.

Proof. This is a direct implication of Kargin's Theorem 5.3, where we use this Theorem point-wise for each $s \in [0,1]$. □

The two concentration inequalities in Theorem 5.1 and Theorem 5.3 have a rate which is quadratic in N. Consequently, the joint spectral measure $\mu_{A_N} \boxplus \mu_{B_N}$ is $\propto \mathcal{O}(N^{-2})$ far away from the limiting distribution $\mu_a \boxplus \mu_b$, predicted by the subordination machinery for free additive convolution. Our Lemma 5.2 and Lemma 5.4 are just point-wise applications of the work by Kargin [Kar12] and thus naturally coincide with the $\mathcal{O}(N^{-2})$ rate. Subsequent works by Collins et al. (2019) [CGP19] confirm the N^{-2} rate for polynomials in GUE random matrices associated to the operator-valued free convolution. In [Par22], the case for i.i.d unitary matrices is analyzed with results coinciding with the additional $\ln^2(N)$ factor, which is also obtained in Theorem 5.1. For operator-valued polynomials, the linearized free additive convolution deviates with $\mathcal{O}\left(\frac{\ln^2(N)}{N^2}\right)$ from polynomials of random matrices with finite dimension N.

Definition 5.5 (Free Adiabatic Quantum Computation)**.** *Let $A, B \in M_d(\mathbb{C})$ be self-adjoint for some large yet finite $d \in \mathbb{N}$. We call the subsequent Hamilton operator the* ***free adiabatic Hamiltonian*** *H_Φ,*

$$H_\Phi(s) := \alpha(s)A + \beta(s)\Phi_U(B), \qquad \alpha(\cdot), \beta(\cdot) \in C^0([0,1],[0,1]), \qquad s \in [0,1], \tag{5.12}$$

where $\alpha(s) = 1 - \beta(s)$, and we call its corresponding quantum adiabatic algorithm the ***Free Adiabatic Algorithm.***

W.l.o.g., we apply Φ to the problem Hamiltonian B. It is emphasized that it is irrelevant, whether we apply the basis randomization to the initial (A) or the problem Hamiltonian (B): asymptotic freeness between two matrices arises, if one basis is transformed at random relative to the other. Here, we consider N qubits, i.e. the multi-qubit Hilbert space dimension corresponds to $d := \dim(\mathcal{H}) = 2^N$.

In practice, such relative rotation originates for instance from fluctuations in the state preparation procedure or due to external noise, that infects the QAA. The practical realization is further discussed in chapter 6.

5.2. Spectral Gap Analysis with Subordination Iteration

Having established the FAQC model, we proceed to show, how one performs a gap analysis using subordination theory. Let $\sigma(A) = \{a_1, ..., a_d\}$ and $\sigma(B) = \{b_1, ..., b_d\}$ be given[1]. The empirical spectral measures μ_A and μ_B are:

$$\mu_{D_A}(t) = \frac{1}{d}\sum_{i=1}^{d}\delta(t - a_i), \qquad \mu_{D_B}(t) = \frac{1}{d}\sum_{i=1}^{d}\delta(t - b_i). \tag{5.13}$$

Here, a_i, b_i denote the $i-$th eigenvalue of A, resp. B. Again, all eigenvalues are counted with multiplicity. According to Example 3.19, the Cauchy transforms of these two measures are

$$G_{\mu_{D_A}}(z) = \frac{1}{d}\sum_{i=1}^{d}\frac{1}{z - a_i}, \qquad G_{\mu_{D_B}}(z) = \frac{1}{d}\sum_{i=1}^{d}\frac{1}{z - b_i}, \qquad z \in \mathbb{C}^+. \tag{5.14}$$

Consequently, the fixed point equation (4.15),

$$g(z,w) = w + \frac{1}{G_\nu\left(z - w + \frac{1}{G_\mu(w)}\right)} - \frac{1}{G_\mu(w)}, \qquad z, w \in \mathbb{C}^+,$$

becomes

$$\begin{aligned} g(z,w) &= w + \frac{1}{G_{\mu_{D_B}}\left(z - w + \frac{1}{G_{\mu_{D_A}}(w)}\right)} - \frac{1}{G_{\mu_{D_A}}(w)} \\ &= w + \frac{1}{\frac{1}{d}\sum_{i=1}^{d}\frac{1}{\left(z - w + \frac{1}{\frac{1}{d}\sum_{j=1}^{d}\frac{1}{w - a_j}}\right) - b_i}} - \frac{1}{\frac{1}{d}\sum_{i=1}^{d}\frac{1}{w - a_i}} \end{aligned}$$

[1] In case that A, B are diagonal (denoted by D_A and D_B), we can directly read off the spectrum.

We are left to parameterize the empirical spectral measures μ_{D_A} and μ_{D_B} with the interpolation schedule as in (5.2), i.e. we consider for the measures in (5.13):

$$s \mapsto \mu_{D_A,s}(t) = \frac{1}{d}\sum_{i=1}^{d}\delta(t-\alpha(s)a_i), \tag{5.15}$$

$$s \mapsto \mu_{D_B,s}(t) = \frac{1}{d}\sum_{i=1}^{d}\delta(t-\beta(s)b_i). \tag{5.16}$$

Hence the Cauchy transforms are:

$$G_{\mu_{D_A},s}(z) = \frac{1}{d}\sum_{i=1}^{d}\frac{1}{z-\alpha(s)a_i}, \qquad G_{\mu_{D_B},s}(z) = \frac{1}{d}\sum_{i=1}^{d}\frac{1}{z-\beta(s)b_i}, \qquad z\in\mathbb{C}^+. \tag{5.17}$$

We prove this construction in the subsequent Theorem 5.6. In a nutshell, the empirical spectral measures are replaced by a one-parameter family of spectral measures which is also inherited in the above Cauchy transforms.

Consequently, we have to solve the fixed point equation (4.15) with the above Cauchy transforms. We summarize our stochastic spectral analysis for the QAA, which we call **Free Adiabatic Quantum Computing,** as follows.

Theorem 5.6. *Let $A, B \in M_d(\mathbb{C})$ be self-adjoint for some $d \in \mathbb{N}$. Consider the free adiabatic Hamiltonian of the form*

$$H_\Phi(s) := \alpha(s)A + \beta(s)\Phi_U(B), \qquad \alpha(\cdot),\beta(\cdot)\in C^0([0,1],[0,1]), \tag{5.18}$$

with $\alpha(s)+\beta(s)=1$, for all $s\in[0,1]$ and $\Phi_U(\cdot) := U(\cdot)U^*$ for any random unitary $U\in\mathcal{U}(d)$. Then, for all $t\in\mathbb{R}$, the $s-$dependent joint distribution over the spectrum, can be approximated by computing the free additive convolution of the spectral measures and hence by running the following algorithm:

1. *Compute the spectrum of A and B, i.e. $\sigma(A)=\{a_1,...,a_d\}$, $\sigma(B)=\{b_1,...,b_d\}$. Denote their diagonal matrices by D_A and D_B.*
2. *Define the one-parameter family of spectral densities by $\mu_{\alpha(s)D_A}$ and $\mu_{\beta(s)D_B}$ as in (5.15) and (5.16) and compute their Cauchy transforms as in (5.17).*
3. *Determine the unique $z-$ and $s-$ dependent fixed point of the mapping $g:\mathbb{C}^+\times[0,1]\times\mathbb{C}^+\to\mathbb{C}^+$,*

$$g(z,s,w) = w + \frac{1}{\frac{1}{d}\sum_{i=1}^{d}\frac{1}{\left(z-w+\frac{1}{\frac{1}{d}\sum_{j=1}^{d}\frac{1}{w-\alpha(s)a_j}}\right)-\beta(s)b_i}} - \frac{1}{\frac{1}{d}\sum_{i=1}^{d}\frac{1}{w-\alpha(s)a_i}}. \tag{5.19}$$

For fixed $(z,s)\in\mathbb{C}^+\times[0,1]$, g naturally induces a contraction mapping in the $w-$component, i.e.

$$g_{z,s}:\mathbb{C}^+\to\mathbb{C}^+, \qquad w\mapsto g_{z,s}(w) := g(z,s,w), \tag{5.20}$$

which converges for each $(z,s)\in\mathbb{C}^+\times[0,1]$ uniquely. This hyperplane of fixed points, i.e.

$$\omega(z,s) := \lim_{n\to\infty} g_{z,s}^{\circ n}(w_0), \qquad \text{for any arbitrary } w_0\in\mathbb{C}^+, \tag{5.21}$$

is an analytic function in z for all $s\in[0,1]$.

4. *For all* $s \in [0,1]$, *compute the subordinated Cauchy transform:*

$$G_{(\mu_{D_A,s} \boxplus \mu_{D_B,s})}(z) = G_{\mu_{D_A}} \circ \omega(z,s) \tag{5.22}$$

5. *Recover the* $s-$*dependent joint density via Stieltjes inversion formula (3.24), i.e.*

$$\rho(s,t) := \mu_{D_A,s} \boxplus \mu_{D_B,s}(t) = \frac{-1}{\pi} \lim_{\varepsilon \to 0^+} \operatorname{Im} G_{\mu_{D_A}}(\omega(t+i\varepsilon, s)) \tag{5.23}$$

Moreover, there exists some $c > 0$, *such that for every fixed* $s \in [0,1]$,

$$\mathbb{P}\left\{ \sup_{t \in \mathbb{R}} \left| \frac{\mathcal{N}_\eta(t)}{2d\eta} - \rho(s,t) \right| \geq \delta_s \right\} \leq \exp\left(-c\delta_s^2 \frac{(\eta d)^2}{(\log d)^2} \right), \tag{5.24}$$

where $\mathcal{N}_\eta(t)$ *counts the eigenvalues in* $(t-\eta, t+\eta]$ *and* $\rho(s,t)$ *denotes the asymptotic joint probability density.*

Proof. 1. $A, B \in M_d(\mathbb{C})$ are self-adjoint. According to the spectral theorem for normal matrices, there exist unitary matrices V_1, V_2 such that A, B are unitarily equivalent to their diagonal matrices, i.e.

$$D_A = V_1 A V_1^*, \qquad D_B = V_2 B V_2^*, \tag{5.25}$$

where $D_A = \operatorname{diag}(a_1, ..., a_d) \in M_d(\mathbb{R})$ and $D_B = \operatorname{diag}(b_1, ..., b_d) \in M_d(\mathbb{R})$ (eigenvalues counted with multiplicity).

2. We show well-definedness by using the following, well-known fact about Dirac's $\delta-$distribution [MMWW10]

$$\delta(f(x)) = \sum_k \frac{1}{|f'(x_k)|} \delta(x - x_k), \qquad f \in C^1(\mathbb{R}),$$

with $f'(x_k) \neq 0$ for all x_k and x_k being the k'th root. We define a family of C^1 functions

$$\tilde{f}_{a,s}(t) := t - f_a(s) = t - \alpha(s)a.$$

Then, $\frac{\partial}{\partial t} \tilde{f}_{a,s}(t) = 1$ for all $a, t \in \mathbb{R}$ and $\alpha(s) \in \mathbb{R}$. Moreover, $\tilde{f}_{a,s}(t) = 0$, iff $t = \alpha(s)a$, so

$$\delta(\tilde{f}_{a,s}(t)) = \frac{1}{|1|} \delta(t - \alpha(s)a) = \delta(t - \alpha(s)a)$$

for any $a, \alpha(s) \in \mathbb{R}$. Hence, we define

$$\mu_{D_A,s}(t) := \frac{1}{d} \sum_{i=1}^d \delta\left(\tilde{f}_{a_i,s}(t)\right) = \frac{1}{d} \sum_{i=1}^d \delta(t - \alpha(s)a_i)$$
$$\mu_{D_B,s}(t) := \frac{1}{d} \sum_{i=1}^d \delta\left(\tilde{f}_{b_i,s}(t)\right) = \frac{1}{d} \sum_{i=1}^d \delta(t - \beta(s)b_i)$$

to obtain that $\mu_{D_A,s}$ and $\mu_{D_B,s}$ are indeed well-defined. Since $a_i, b_i \in \mathbb{R}$ for all $i \in [d]$ and $\alpha(s), \beta(s) \in \mathbb{R}$ for all $s \in [0,1]$, $a_i'(s) := \alpha(s)a_i \in \mathbb{R}$, and $b_i'(s) := \beta(s)b_i \in \mathbb{R}$. Thus,

$\mu_{D_A,s}(t) = \frac{1}{d}\sum_{i=1}^{d}\delta(t-a_i'(s))$ and $\mu_{D_B,s}(t) = \frac{1}{d}\sum_{i=1}^{d}\delta(t-b_i'(s))$ are again compactly supported, real-valued probability measures. Normalization is seen via

$$\int_{\mathbb{R}}\mu_{D_A,s}(t)\,dt = \int_{\mathbb{R}}\frac{1}{d}\sum_{i=1}^{d}\delta(t-a_i'(s))\,dt \overset{a_i'(s)\in\mathbb{R}}{=} \frac{d}{d} = 1,$$

and similarly for $\mu_{D_B,s}$. Hence, the respective Cauchy transform is by construction well-defined, as well.

3. For $z\in\mathbb{C}^+$, the Cauchy-transforms are:

$$G_{\mu_{D_A},s}(z) = \int_{\mathbb{R}}\frac{1}{z-t}\left(\frac{1}{d}\sum_{i=1}^{d}d\delta(t-\alpha(s)a_i)\right) = \frac{1}{d}\sum_{i=1}^{d}\frac{1}{z-\alpha(s)a_i}, \tag{5.26}$$

$$G_{\mu_{D_B},s}(z) = \int_{\mathbb{R}}\frac{1}{z-t}\left(\frac{1}{d}\sum_{i=1}^{d}d\delta(t-\beta(s)b_i)\right) = \frac{1}{d}\sum_{i=1}^{d}\frac{1}{z-\beta(s)b_i}. \tag{5.27}$$

We choose $G_\mu := G_{\mu_{D_A},s}$ and $G_\nu := G_{\mu_{D_B},s}$ and define

$$g(z,s,w) := w + \frac{1}{G_{\mu_{D_B},s}\left(z-w+\frac{1}{G_{\mu_{D_A},s}(w)}\right)} - \frac{1}{G_{\mu_{D_A},s}(w)}, \qquad z,w\in\mathbb{C}^+,$$

For fixed $z\in\mathbb{C}^+$, the above map induces a one-parameter family of mappings,

$$s\mapsto g_{z,s}(\cdot), \qquad s\in[0,1],$$

such that $g_{z,s}(w) = g(z,s,w)$, which is for any $s\in[0,1]$ and any $z\in\mathbb{C}^+$ a contraction mapping in the $w-$component according to Theorem 4.1. Again, due to Theorem 4.1, this $s-$dependent fixed point is determined by $\omega(z,s) = \lim_{n\to\infty} g_{z,s}^{\circ n}(w_0)$ for arbitrary $w_0\in\mathbb{C}^+$ and $\omega(z,s)$ is an analytic function in z.

4. By running the Subordination Algorithm (Algorithm 1) for any $s\in[0,1]$, we compute both the subordinated Cauchy transform and

5. recover the joint density via Stieltjes inversion formula (3.24) Define the resulting distribution by $\rho(s,t)$.

The last estimate follows directly from Lemma 5.4, which concludes the proof. □

Theorem 5.6 shows, how subordination theory for free additive convolution is transferred to the FAQC model. The bounded error rate (5.24) in Theorem 5.6 has one special case. Let g denote the magnitude of a spectral gap for some fixed $s_0\in[0,1]$. That is, $\eta = \frac{g}{2}$ and the number of eigenvalues $\mathcal{N}_\eta(t) = 0$. Then, Theorem 5.6 provides the existence of a constant $c>0$ such that the error rate (for some given $\delta_{s_0}>0$) is bounded by

$$\mathbb{P}\left\{\sup_{t\in\mathbb{R}}\left|\rho(s,t)\right| \geq \delta_{s_0}\right\} \leq \exp\left(\frac{-c\delta_{s_0}^2}{4}\frac{g^2d^2}{(\log d)^2}\right). \tag{5.28}$$

We are left to discuss, what we actually understand as a "spectral gap" and how we can determine those gaps within our FAQC model.

5.2.1. FASGE

Due to the fact that any free convolution of two probability distributions is again a probability distribution, we must define how we can estimate (stochastic) spectral gaps from it. Throughout, the term "gap", "spectral gap" (of a finite-dimensional, positive operator) indicates the distance between the two lowest eigenvalues. The precise definition was given in Chapter 2.2. If we deal with the one-parameter family of operators (5.2), we say that a gap exists, if $g_{\min}(H(s)) > 0$.

By construction, we consider a probability density ρ over positive real numbers, i.e. $\mathrm{supp}(\rho) = \mathbb{R}^+$. Our intuition is clear if, ρ is a sum of Dirac distributions, i.e. $\rho(t) = \frac{1}{d}\sum_{j=1}^{d}\delta(t-\lambda_j)$, $\lambda_j \in \mathbb{R}^+$, for all $j \in [d]$. In this particular case, the gap is just $g = |\lambda_2 - \lambda_1|$.

However, since we calculate ρ via the Subordination Algorithm numerically, we can neither assume a certain shape of ρ, nor impose e.g. a C^1- structure in general. Furthermore, any gap analysis is sorely stochastic, since the main object of spectral information is a probability density. For this reason, the subsequent Definition 5.7 is rather numerical than analytical. In order to make the definition precise, we have to make one assumption. We assume that ρ has a finite number of local maxima, and we denote the set where these maxima are attained by

$$L_\rho := \{t_{\max} \in \mathbb{R}^+ | \rho(s_0, t_{\max}) \text{ is a local maximum}, s_0 \in [0,1]\} \tag{5.29}$$

for fixed s_0. The height of each maximum might even be infinite (in the case of $\delta-$peaks). We define the *stochastic gap* in two ways, as follows:

Definition 5.7. *Let $s_0 \in [0,1]$ be fixed and let $\rho(s_0,t)$ be the joint density computed by Theorem 5.6 over $t \in \mathbb{R}^+$ and let L_ρ as in (5.29). We define the **sharp gap**, denoted by g_{s_0}, where for some small $\zeta > 0$,*

$$t_0 := \inf_t \{t \in \mathbb{R}^+ : |\rho(s_0,t)| \geq \zeta\} \cap L_\rho, \tag{5.30}$$

$$t_1 := \inf_{t \setminus \{t_0\}} \{t \in \mathbb{R}^+ : |\rho(s_0,t)| \geq \zeta\} \cap L_\rho. \tag{5.31}$$

Then,

$$g_{s_0} := |t_1 - t_0| \tag{5.32}$$

*Moreover, we define the **soft gap** g'_{s_0}, as follows. Consider for some small $\eta > 0$ the set*

$$M_\eta := \{t \in \mathbb{R}^+ | \rho(s_0,t) \geq \eta\}.$$

Let $M_1, M_2 \subset M_\eta \subset \mathbb{R}^+$ be disjoint subsets such that $\sup M_1 < \inf M_2$ and $\inf M_2 \leq \inf\{M_\eta \setminus M_1\}$. We denote these by $t'_1 := \sup M_1$ and $t'_2 := \inf M_2$. Then,

$$g'_{s_0} := \mathrm{dist}(M_1, M_2) \equiv |t'_2 - t'_1|. \tag{5.33}$$

A visualization of these gaps is provided in Figure A.1. In all of the plots that we present, the red gap corresponds to the 'soft gap' and the green gap corresponds to the 'sharp gap'. Throughout, the sharp gap is computed by the *Peak Finder* (Algorithm 4) and the *soft gap* by the *Gap Estimator* (Algorithm 3). The two algorithms are located in Appendix A. In words, the sets M_1, M_2 are the pre-images of those two hills of the probability density function, which are closest to zero. Hence determining stochastic spectral gaps is equivalent to isolating two local maxima. Our analysis is purely numerical. The two gaps from Definition 5.7 are distinguished by the sharpness of corresponding local maxima. Just as Figure A.1 demonstrates, the "softer" gap is usually a more pessimistic estimate on the gap, compared to the sharper gap which measures

the distance between two peaks. It might happen that the density function is constant or that it oscillates so fast, that neither peaks nor slopes are numerically detectable by our method. Then, we set the gap to zero.

Lemma 5.8. *Consider an arbitrary probability density function on $\mathbb{R}^+$. Then, there exist algorithms that determine the stochastic spectral gaps from Definition 5.7. The 'soft gap' is computed by the* ***Gap Estimator*** *(Algorithm 3) and the 'sharp gap' is computed by the* ***Peak Finder*** *(Algorithm 4).*

Finally, we are able to combine the above methods to construct the numerical algorithm that follows the recipe from Theorem 5.6. Our spectral gap estimating algorithm for the FAQC model is provided by the FASGE (Algorithm 2), which we formalize in the subsequent Lemma.

Lemma 5.9. *Consider the free adiabatic Hamiltonian (5.18). By combining the Subordination Iteration (Algorithm 1) with the Peak Finder (Algorithm 4) and the Gap Estimator (Algorithm 3) for every $s \in [0,1]$, there is a spectral gap estimating algorithm (Algorithm 2) for the Hamiltonian (5.18), which is called* ***Free Adiabatic Spectral Gap Estimator (FASGE).***

Algorithm 2 `FASGE` $(\mu_A, \mu_B, \mathcal{D}, \mathcal{S})$

Input: μ_A, μ_B **and** $\mathcal{D}$ **and** $\mathcal{S} = \{s_0, ..., s_{S-1}\}$ ▷ $\mathcal{D}$: see Algorithm 1
Input: (optional) η, ζ, ξ **and optional** parameters of 'Sub' (Algorithm 1)
  Initialize parameters according to Algorithm 1
  for $s \in \mathcal{S}$ **do** ▷ iterate through $\mathcal{S}$
    $\mu_s \leftarrow (1-s) \cdot \mu_A$
    $\nu_s \leftarrow s \cdot \mu_B$
    $\mu_s \boxplus \nu_s(\mathcal{D}) \leftarrow$ `Sub` $(\mu_s, \nu_s, \mathcal{D})$ ▷ `Sub`: run Subordination Iteration (Algorithm 1)
    method 1: $g_s \leftarrow$ `Peaks` $(\mu_s \boxplus \nu_s(\mathcal{D}), \zeta, \xi)$ ▷ determine gaps using `Peaks` (Algorithm 4)
    method 2: $g'_s \leftarrow$ `GE` $(\mu_s \boxplus \nu_s(\mathcal{D}), \eta)$ ▷ determine gaps using `GE` (Algorithm 3)
    store g_s, g'_s
  end for
  $g_{\min} \leftarrow \min_{s\in\mathcal{S}}\{g_s\}$
  $g'_{\min} \leftarrow \min_{s\in\mathcal{S}}\{g'_s\}$
Output: $g_{\min}$, $g'_{\min}$ ▷ outputs minimal spectral gaps according to the respective methods

A few remarks on the run time $\mathcal{T}$ of Algorithm 2. Just like the Subordination Algorithm, the main costs are due to iteration quantities. With $|\mathcal{S}| = S, |\mathcal{D}| = D$ and the maximal number of fixed point iterations M, the FASGE converges in $\mathcal{O}(SDM)$. If the number of distinct eigenvalues scales with the dimension (as it is the case for the initial Hamiltonian A in (5.35)), we get an additional factor that inherits this dependence. For an operator with $d = 2^N$ for N qubits, the run time scales with

$$\mathcal{T}_{\text{FASGE}} \propto \mathcal{O}(SDM \log(d)). \tag{5.34}$$

Typically, we choose

$$S = 10^2, \qquad D = 10^4, \qquad M = 10^2, \qquad d = 10^3,$$

which gives $\mathcal{T}_{\text{FASGE}} \propto \mathcal{O}(10^8)$ operations. Comparing this performance with the performance of a singular value decomposition, ($\mathcal{T}_{SVD} \propto \mathcal{O}(d^3)$) gives $\mathcal{T}_{\text{FASGE}} \propto S \cdot 10^6$ vs. $\mathcal{T}_{\text{SVD}} \propto S \cdot 10^9$ for

$d = 10^3$. If we take $N = 22$ qubits, i.e $d = 10^6$, we have for fixed $s \in [0, 1]$ (i.e. leaving S out):

$$\mathcal{T}_{\text{FASGE}} \propto \mathcal{O}(10^6) \qquad \text{vs.} \qquad \mathcal{T}_{\text{SVD}} \propto \mathcal{O}(10^{18}),$$

which shows an **exponential speed-up** in approximating spectral gaps, compared to the deterministic SVD. Anyway, we must not forget that the prize we pay for this speed-up is fusing the QAA with a random unitary transformation! Since we *change* the model of quantum computation here, a direct comparison to classical pre-processing algorithms is to be drawn with caution.

Having described our FAQC model of determining spectra of Hamilton operators related to QAA, we proceed by presenting a generic example of such a Hamiltonian. This example is of a rather theoretical nature, but it is nevertheless a suitable candidate for demonstrating subordination theory applied to adiabatic quantum computation. Our final example is a problem of utmost interest for quantum search algorithms: the *(free) adiabatic Grover problem.* Before we run the FASGE to determine spectral gaps, we present several snapshots of the subordination algorithm for fixed s. Furthermore, we expound numerical challenges and finally provide an outlook, how a *dynamic* version of the FASGE might overcome the current thresholds of the (static) FASGE.

5.3. A Generic Example of the Free QAA

We exemplify the FAQC model by considering a Hamilton operator of the form (5.18), as follows. We take the initial Hamiltonian A to be

$$A := \sum_{j=1}^{N} (\mathbb{1}_2 - \sigma_x)^{(j)}, \tag{5.35}$$

where σ_x is the Pauli-x matrix and $\mathbb{1}_2$ denotes the two-dimensional identity matrix. The notation $(\cdot)^{(j)}$ indicates that $(\mathbb{1}_2 - \sigma_x)$ acts on the j'th qubit, e.g. $(\mathbb{1}_2 - \sigma_x)^{(3)} \equiv \mathbb{1}_2 \otimes \mathbb{1}_2 \otimes (\mathbb{1}_2 - \sigma_x)$. Its ground state is the easy-to-prepare uniform superposition over N qubits. The spectrum of A is

$$\sigma(A) = \{0, ..., N\}, \tag{5.36}$$

where $a_1 := 0$ and $a_d := N$ denote the lowest, resp. highest eigenvalues, which are non-degenerate. The algebraic multiplicity m of the eigenvalues in between is binomially distributed, i.e. $m(a_k) = \binom{N}{k}$, $k \in [N]$. The operator (5.35) is the standard choice for the initial Hamiltonian in QAA [FGGS00]. We choose B' to be the randomized, unitarily rotated operator of A, thus

$$H_\Phi(s) := (1-s)A + s\Phi_U(A), \qquad \Phi_U(\cdot) = U(\cdot)U^*, \qquad U \in \mathcal{U}(d), \tag{5.37}$$

i.e. $d = 2^N$, $\alpha(s) = 1 - s$ and $\beta(s) = s$. This operator is a convex combination of A with itself, despite a random basis rotation on one of the extreme points. Due to the problem being symmetric, we expect some symmetry, both for the spectral probability density and the spectral gap.

In order to compare our algorithm with simulations involving random matrices, we consider 11 qubits, i.e. $N = 11$ and $d = 2^{11} = 2048$. This has two reasons: at first, random matrices of dimension $d \propto 10^3$ can be diagonalized reasonably fast on a classical computer. Second, the spectral distribution of finite-dimensional matrices of this dimension is sufficiently close to their limit distribution (see Fig. 3.1 and Fig. 4.1). Following Theorem 5.6, we run the subordination algorithm for each $s \in [0, 1]$, as follows. There exists a unitary operator $V \in U(d)$, such that $D_A = VAV^*$, where

$$D_A = \mathrm{diag}(a_1, a_2, ..., a_d). \tag{5.38}$$

Counting the eigenvalues with multiplicity, we have: $0 = a_1 \leq a_2 \leq ... \leq a_d = N$, where a_1 and a_d are non-degenerate. Thus, the (un-parameterized) density of the spectral measure for A is

$$\mu_{D_A}(t) = \frac{1}{d} \sum_{i=1}^{d} \delta(t - a_i),$$

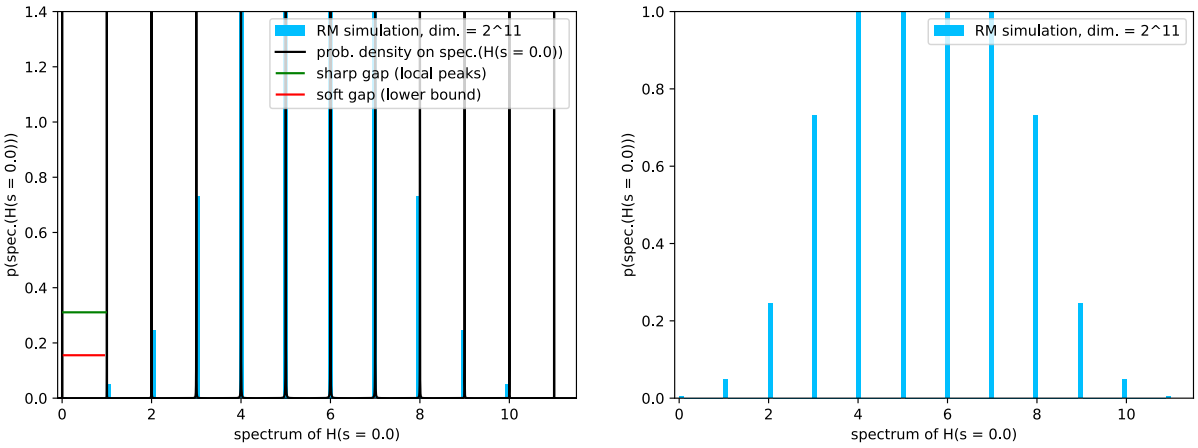
which is the same for μ_{D_B} in this specific example. Therefore, we take the parameterized densities of the spectral measures to be

$$\mu_{D_A,s}(t) = \frac{1}{d} \sum_{i=1}^{d} \delta(t - (1-s)a_i) \tag{5.39}$$

$$\mu_{D_B,s}(t) = \frac{1}{d} \sum_{i=1}^{d} \delta(t - sa_i) \tag{5.40}$$

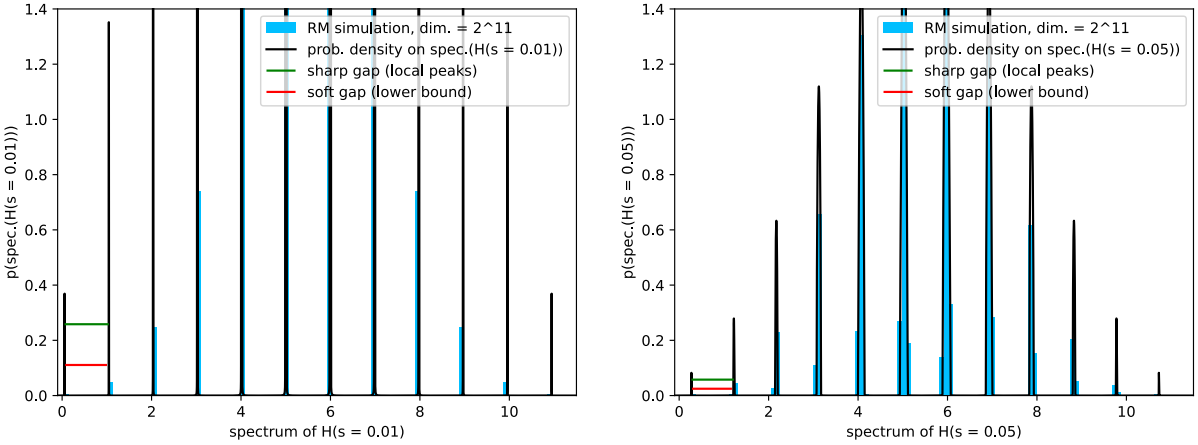
and run Sub $(\mu_{D_A,s}, \mu_{D_B,s}, \mathcal{D})$ (Algorithm 1) for the desired values of $s \in [0, 1]$ on the iteration domain $\mathcal{D}$ which corresponds to an interval partitioning of $[0, N]$. We take $\mathcal{D} := \{0 = t_0 \leq t_1 \leq ... \leq t_D = N\}$ with $D = 10^4$, i.e. a uniform partitioning: $\Delta t = t_k - t_{k-1} = const.$ for all $k \in [D]$.

Moreover, we include visualizations of the stochastic gaps from Definition 5.7. These gaps correspond to the results from Algorithm 3 (red gap) and Algorithm 4 (green gap). Due to symmetry, we show several snapshots up to $s = 0.5$. Then, the density "moves backwards", i.e. $\rho(s', t) = \rho((1 - s'), t)$ for $s' \in [0, 0.5]$ and $t \in \mathcal{D}$. Furthermore, we present the FASGE for the



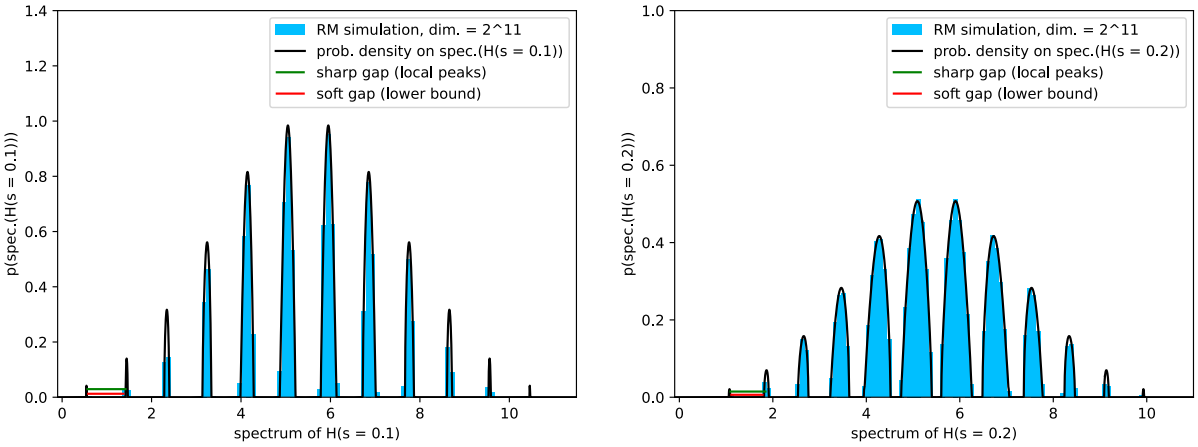
(a) compared to Subordination Iteration (b) without Subordination Iteration

Figure 5.1.: Spectral density of the generic Hamiltonian $H_\Phi(s)$ from (5.37) at $s = 0$: uniform superposition over all qubits. We see how the subordination algorithm Sub $(\mu_{D_A,s}, \mu_{D_B,s}, \mathcal{D})$ determines the positions of the eigenvalues correctly. Since the density of $s = 0$ corresponds to a sum of delta distributions, the "only" ambiguity is obtained in the height of these peaks. However, the stochastic gaps are computed precisely to $g = 1$, as Figure 5.1a shows.



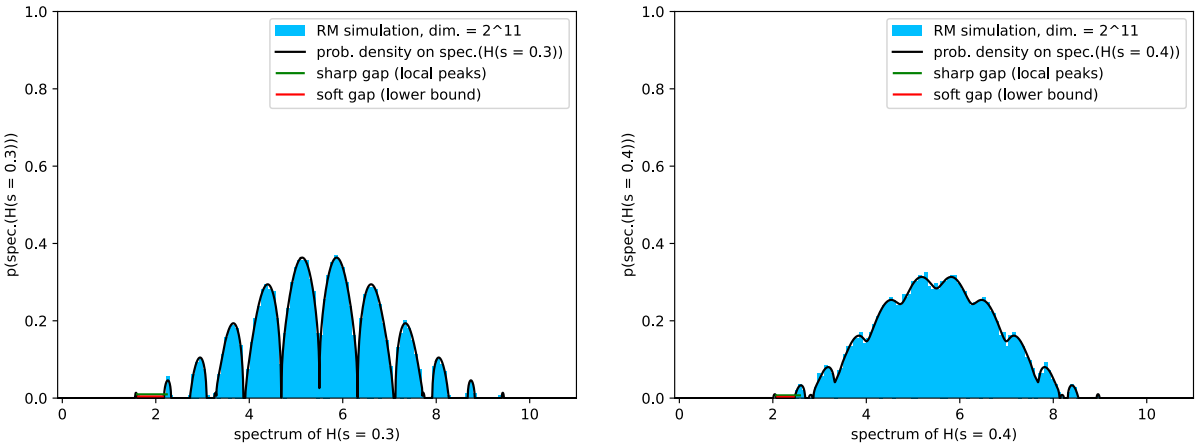
(a) $s = 0.01$ (b) $s = 0.05$

Figure 5.2.: Spectral density of the generic Hamiltonian $H_\Phi(s)$ from (5.37) at $s = 0.01$ and $s = 0.05$. The delta peaks become smaller and broader. Therefore, the predicted spectral density from the subordination algorithm gets closer to the random-matrix simulation. For illustration purposes, we set the scaling of the $y-$axis a bit higher than $y = 1$. It is emphasized that all plots show probability densities.



(a) $s = 0.1$ (b) $s = 0.2$

Figure 5.3.: Spectral density of the generic Hamiltonian $H_\Phi(s)$ from (5.37) at $s = 0.1$ and $s = 0.2$. The delta peaks are smeared out. The predicted spectral density from the subordination algorithm gets closer further.



(a) $s = 0.3$ (b) $s = 0.4$

Figure 5.4.: Spectral density of the generic Hamiltonian $H_\Phi(s)$ from (5.37) at $s = 0.3$ and $s = 0.4$. Determining the gap with bare eyes becomes more difficult. At this point, one can see why we introduced two kinds of a stochastic spectral gap. Since the resulting probability densities can have several distinct "shapes", the precise definition of a gap might be ambiguous. Figure A.1 enlightens this.

generic Hamilton operator in (5.37). The result is shown in Figure 5.6. In order to determine the minimal gaps precisely, we repeat the FASGE for several values of s on an interval, closely around the expected minimal spectral gap, e.g. $s \in [0.4, 0.6]$. The numerical results are

$$s_{\text{soft}} = 0.4904, \quad g'_{\text{min}}(s_{\text{soft}}) \approx 0.004 \quad \text{and} \quad s_{\text{sharp}} = 0.4988, \quad g_{\text{min}}(s_{\text{sharp}}) \approx 0.024. \tag{5.41}$$

This procedure of starting with fewer values for s and then "zooming in" to the expected gap is a suitable candidate for determining minimal gaps, in general.

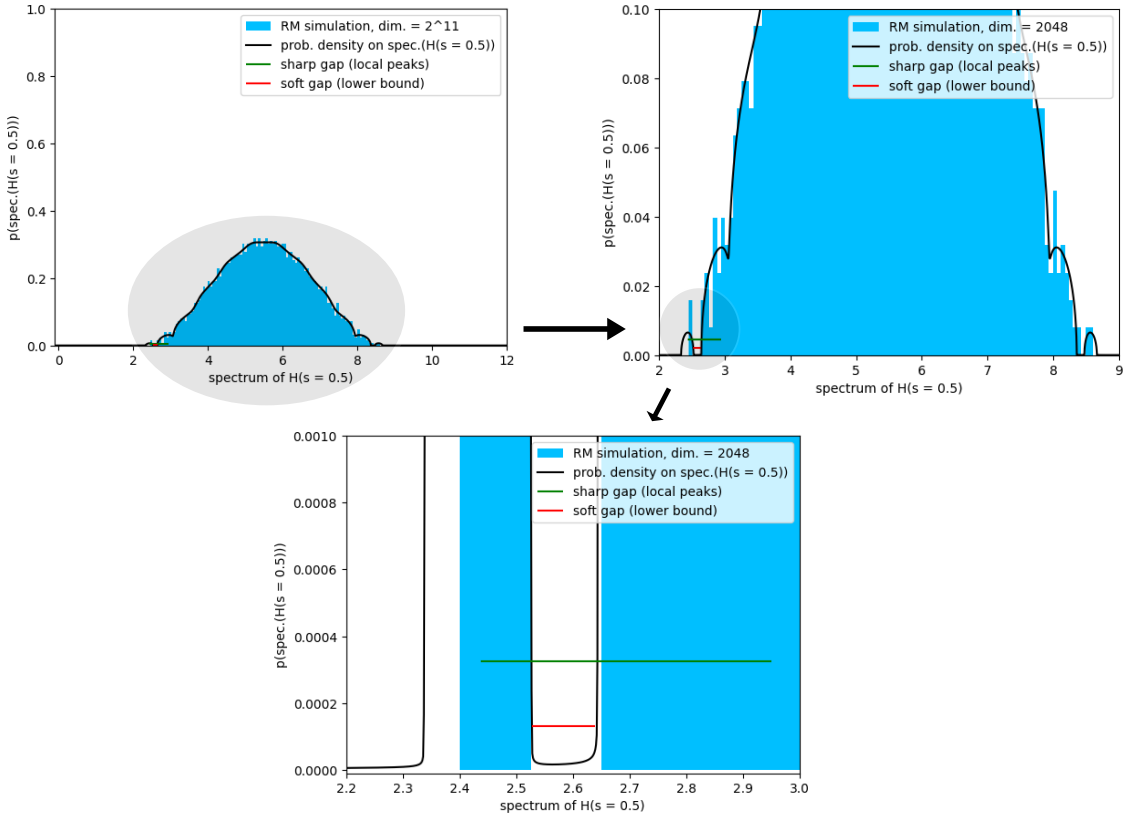


Figure 5.5.: Spectral density of the generic Hamiltonian $H_\Phi(s)$ from (5.37) at $s = 0.5$. At $s \approx 0.5$, the FQAA for $H_\Phi(s)$ reaches its minimal spectral gap. The gap is not visible from a static axis-scaling, thus we "zoom in" to get a closer look at the spectral gap. At first, it is not even clear whether a gap exists. However, zooming in twice answers this question. Even several rounds of generating random matrices never interfere with the gap, which we see here. The soft gap (displayed in red) provides a lower bound on the gap, where the sharp gap (displayed in green; gap between local maxima) is understood as a more optimistic gap in this case. However, as Figure 5.6 shows, the **minimal** gap is not reached at $s = 0.5$, but at $s_{\text{soft}} = 0.4904$ with a soft gap of $g'_{\text{min}} \approx 0.004$, resp. at $s_{\text{sharp}} = 0.4988$ with a sharp gap of $g_{\text{min}} \approx 0.024$. These slight differences in gap determination are expected due to the different definitions of the soft and the sharp gap (see Definition 5.7).

5.4. The Free Adiabatic Grover Problem

In this section we present our free version of the adiabatic Grover problem. We consider the general Grover operator dedicated to search problems with one marked state. Besides Shor's factorization algorithm [Sho94], Grover's search algorithm [Gro96] is perhaps the most famous algorithm based on quantum mechanics. Being given an unsorted database, the task is to find one marked item or eventually several marked items in as few search steps as possible. Let the marked item be denoted by w. Denoting the other elements in that database by x, the problem

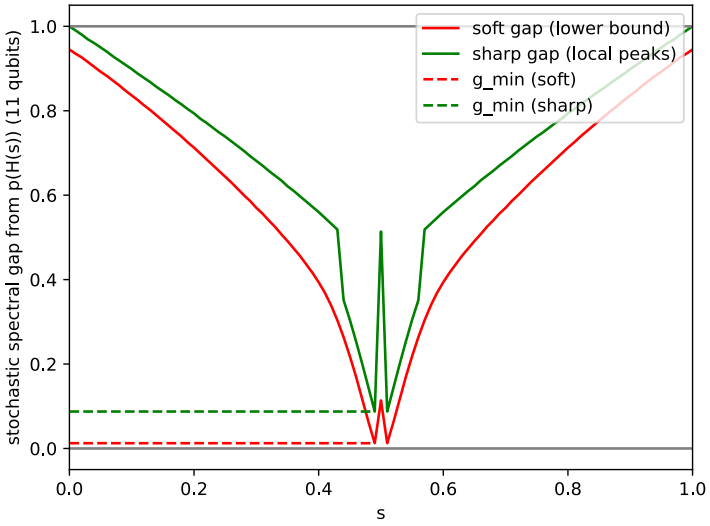


Figure 5.6.: FASGE for the generic Hamiltonian $H_\Phi(s)$ from (5.37). The minimal spectral gaps are obtained at $s \approx 0.49$ (due to symmetry, also at $s \approx 0.51$) with magnitudes as presented in (5.41). In order to make this computation more precise, we did more rounds of the FASGE closer at the expected minimal gap. Both soft and sharp gap exist in all computations.

can be formulated, as follows [AL18, Chapter III.A]. The goal of the Grover search is, for

$$f : \{0,1\}^N \to \{0,1\}, \qquad f(x) = \begin{cases} 1, & x = w \\ 0, & x \neq w, \end{cases} \tag{5.42}$$

(where there are $d = 2^N$ bit strings), to call this function in the smallest number of calls. This (classical) oracle function is translated to a quantum Hamilton operator of N qubits, via bit-to-qubit mapping. Hence, the oracle Hamiltonian H_P for the adiabatic Grover problem is of the form ($z = z_1 z_2 ... z_N$) [AL18]:

$$H_P = \mathbb{1}_{2^N} - |z = w\rangle\langle z = w|. \tag{5.43}$$

Consequently, the marked state $|w\rangle$ corresponds to the ground state of H_P with energy $\lambda_0(H_P) = 0$ and each of the other states $|z\rangle$ ($z \neq w$) has the energy $\lambda_1(H_P) = 1$. Thus, the algebraic multiplicities are $m(\lambda_1) = 2^N - 1$, and $m(\lambda_0) = 1$.

Returning to the FQAA, we take for A again the initial Hamiltonian (5.35). Clearly, we take $B := H_P$,

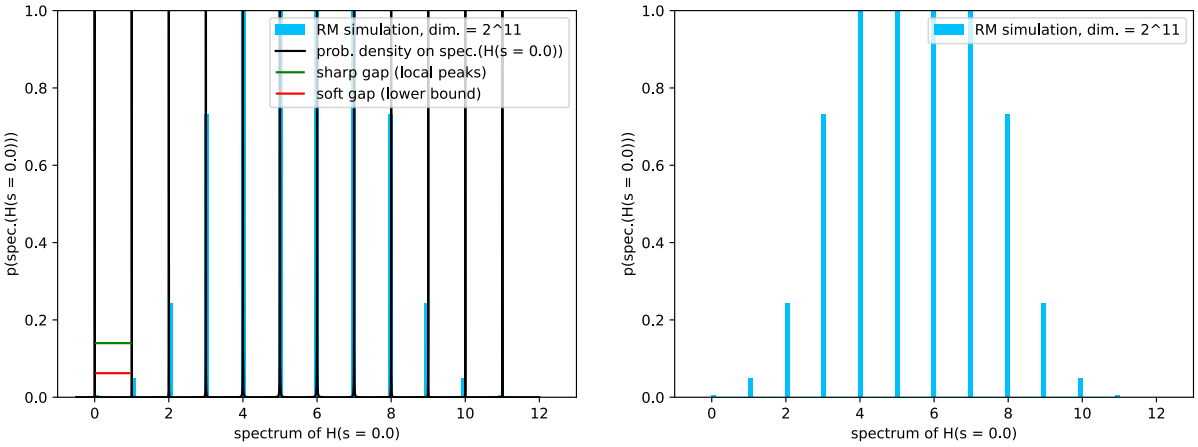
$$H_\Phi(s) = (1-s)A + s\Phi_U(H_P), \qquad \Phi_U(\cdot) = U(\cdot)U^*, \qquad U \in \mathcal{U}(d), \tag{5.44}$$

so that the densities of the parameterized spectral measures become

$$\mu_{D_A,s}(t) = \frac{1}{d}\sum_{i=1}^{d} \delta(t - (1-s)a_i) \tag{5.45}$$

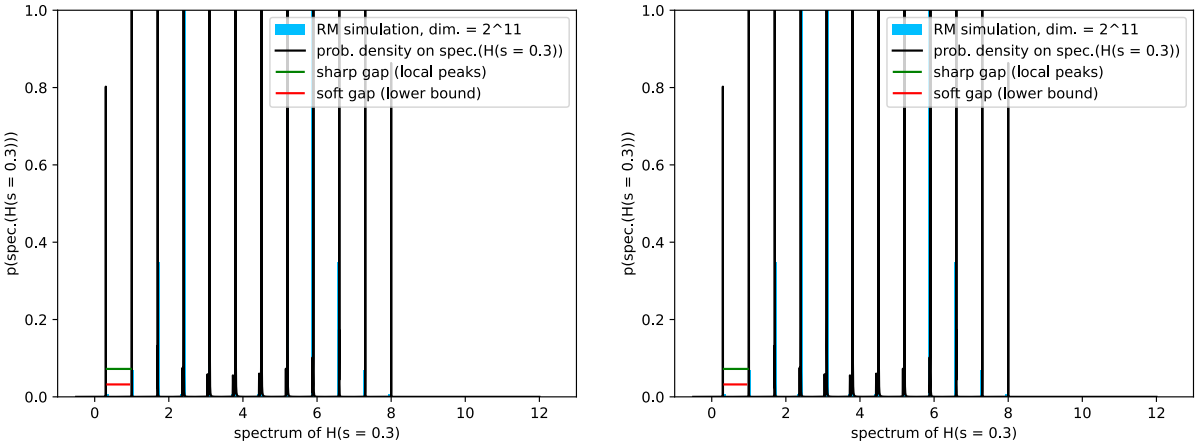
$$\mu_{D_B,s}(t) = \frac{1}{d}\left(\delta(t - s\cdot 0) + (d-1)\cdot\delta(t - s\cdot 1)\right). \tag{5.46}$$

We run Sub $(\mu_{D_A,s}, \mu_{D_B,s}, \mathcal{D})$ for several values of $s \in [0,1]$ with $\mathcal{D}$ being the same domain as in our first example. We are left to demonstrate the FASGE for the free adiabatic Grover.



(a) compared with Subordination Iteration (b) without Subordination Iteration

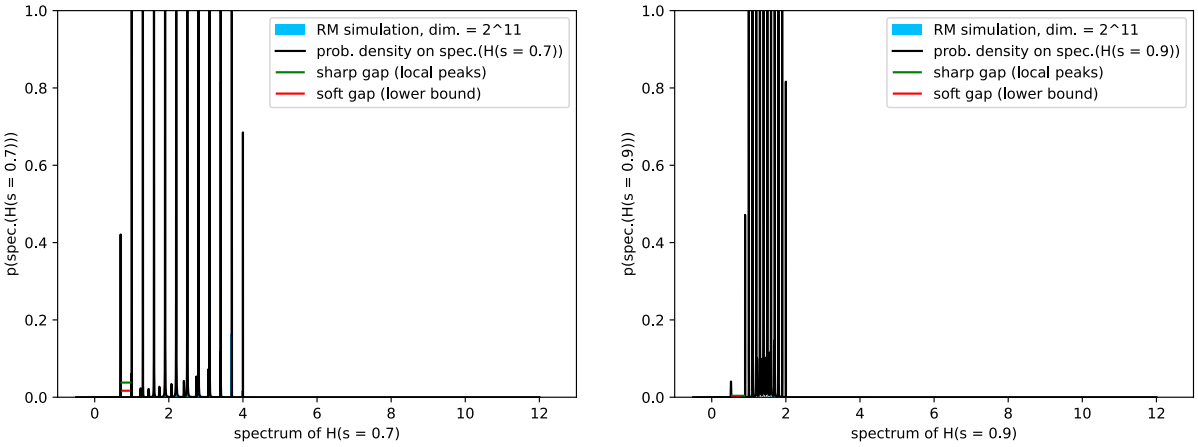
Figure 5.7.: Spectral density of the free adiabatic Grover Hamiltonian $H_\Phi(s)$ from (5.44) at $s = 0$: uniform superposition over all qubits. Results from the Subordination Iteration and simulation with random matrices coincide well. We show again the difference between Sub $(\mu_{D_A,s=0}, \mu_{D_B,s=0}, \mathcal{D})$ covered with random matrix simulations (a), and the random matrix simulation without subordination (b). Similarly to Figure 5.1a, the height of the subordination results does not fit the height of the uniform superposition accurately. However, the location of the spectral values is matched precisely and the gap in Figure 5.7a is correctly determined.



(a) $s = 0.3$ (b) $s = 0.5$

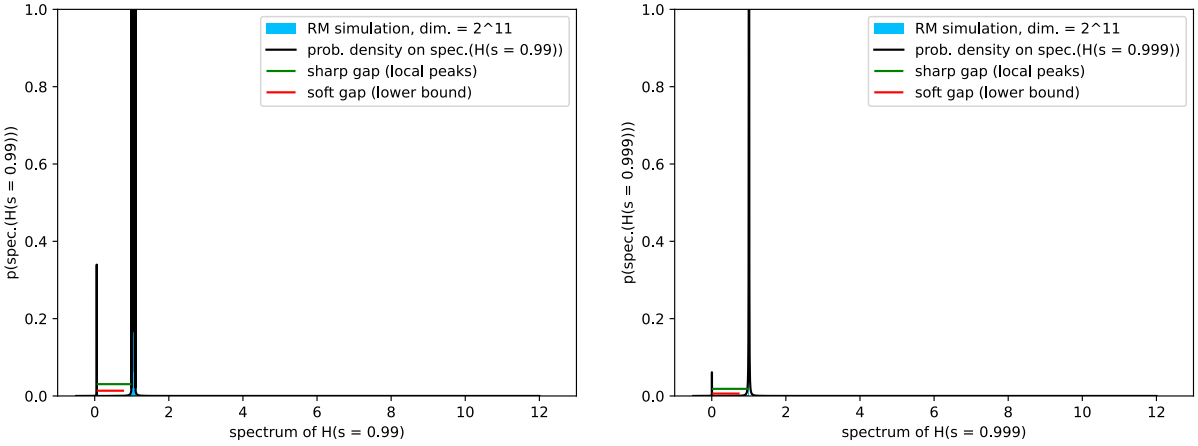
Figure 5.8.: Spectral density of the free adiabatic Grover Hamiltonian $H_\Phi(s)$ from (5.44) at $s = 0.3$, and $s = 0.5$. The spectrum adiabatically evolves towards the spectrum of the final Hamiltonian (5.43). Random matrix simulation and subordination show a perfect coincidence. In contrast to the continuous density displayed in Figure (5.4a), the density remains a discrete sum of delta peaks, here.

Figure 5.12 displays the spectral gap of the free adiabatic Grover using the FASGE. We take only the sharp gap for a reliable spectral gap analysis, since the density over the spectrum is a



(a) $s = 0.7$ (b) $s = 0.9$

Figure 5.9.: Spectral density of the free adiabatic Grover Hamiltonian $H_\Phi(s)$ from (5.44) at $s = 0.7$ and $s = 0.9$. The gap becomes smaller.



(a) $s = 0.99$ (b) $s = 0.999$

Figure 5.10.: Spectral density of the free adiabatic Grover Hamiltonian $H_\Phi(s)$ from (5.44) at $s = 0.99$ and $s = 0.999$. A closer look provides that the blue peaks from the random matrix simulation fit perfectly with the results from the Subordination Algorithm. The stochastic gap evolves towards $g = 1$, between $\lambda_0 = 0$ and $\lambda_1 = 1$. Here, one obtains why for discrete densities, the sharp gap (green gap) is more accurate than the soft gap (red gap).

sum of delta peaks. The minimal gap is reached at $s_{\min} = 0.83$ with $g_{\min}(s_{\min}) = 0.0075$.

Moreover, we run the FASGE for the free adiabatic Grover of 22 qubits. The result is displayed in Figure 5.13. The minimal gap fluctuates shortly before the end of the adiabatic evolution. The minimal gap is reached at $s_{\min} = 0.99$ with $g_{\min}(s_{\min}) = 0.0094$

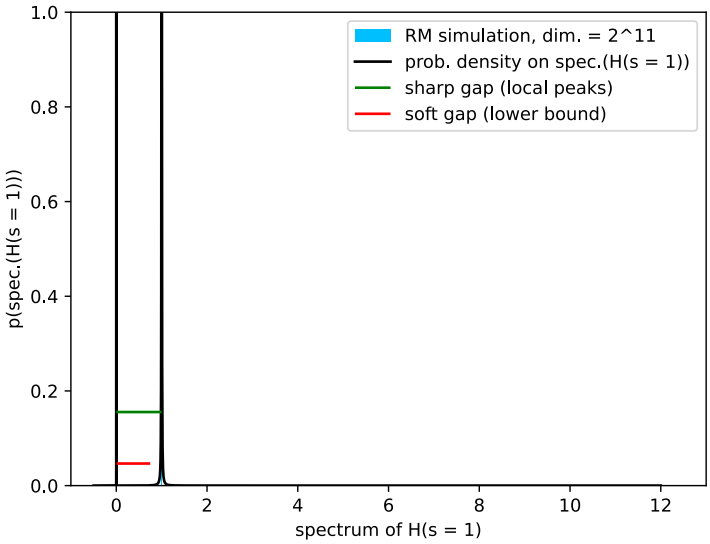


Figure 5.11.: Spectral density of the free adiabatic Grover Hamiltonian $H_\Phi(s)$ from (5.44) at $s = 1$. The adiabatic evolution evolved towards the ground state of H_P. The sharp gap is $g = 1$, where the soft gap is $g' = 0.5$. If the probability density over the spectrum remains discrete, one should rely on the sharp gap, since the soft gap considers changes in the slope.

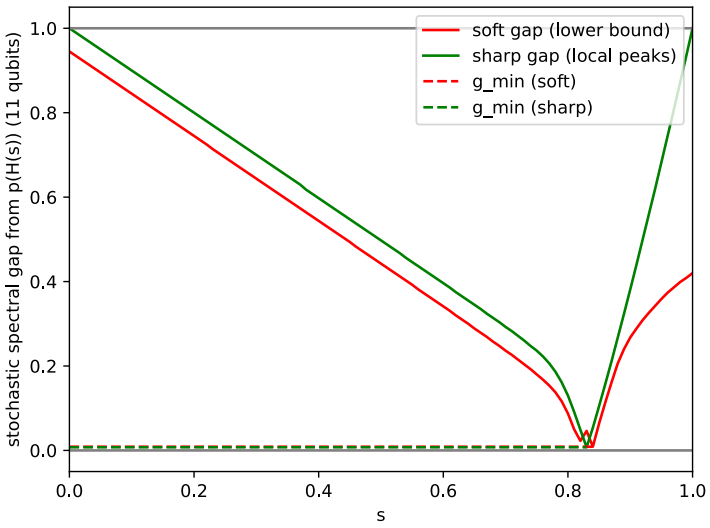


Figure 5.12.: Spectral gap analysis: FASGE for the free adiabatic Grover $H_\Phi(s)$ from (5.44) with the linear schedule $\alpha(s) := 1 - s$, $\beta(s) := s$ with 11 qubits, $s \in [0, 1]$. The interval is uniformly partitioned with 100 steps, i.e. $|\mathcal{S}| = 100$. The stochastic gaps exist. The minimal gap is reached at $s_{\min} = 0.83$ with $g_{\min}(s_{\min}) = 0.0075$.

5.5. Outlook to 100 Qubits: Dynamic FASGE

Both the generic Hamiltonian and the free adiabatic Grover problem show, that we have developed a powerful, classical stochastic pre-processing method to compute adiabatic spectra. The

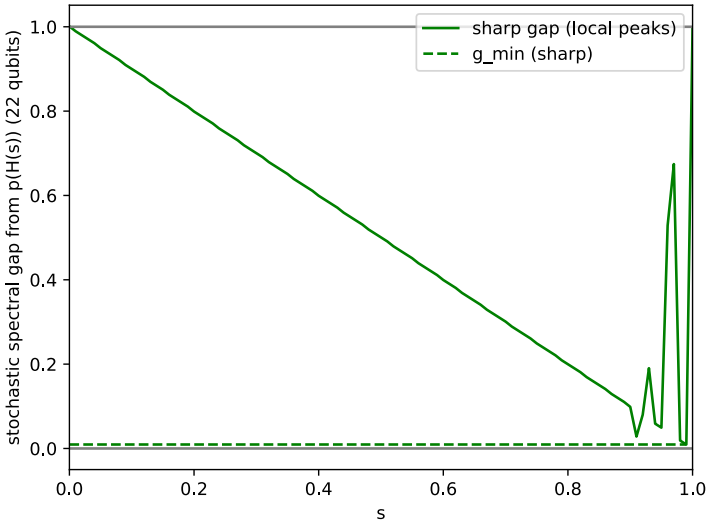


Figure 5.13.: Spectral gap analysis: FASGE for the free adiabatic Grover $H_\Phi(s)$ from (5.44) with the linear schedule $\alpha(s) := 1 - s$, $\beta(s) := s$ with 22 qubits. We restrict to the sharp gap (compare with Definition 5.7). The sharp gap exists. The minimal gap is reached at $s_{\min} = 0.99$ with $g_{\min}(s_{\min}) = 0.0094$. In contrast to the 11-qubit example, we detect more fluctuations of the gap.

Parameters	**Generic Example**	**Grover, 11 qubits**	**Grover, 22 qubits**
η	10^{-4}	$5 \cdot 10^{-4}$	$5 \cdot 10^{-5}$
ζ	10^{-4}	$1 \cdot 10^{-4}$	$5 \cdot 10^{-5}$
ξ	0	0	0
ε [Sub]	10^{-4}	10^{-4}	10^{-6}

Table 5.1.: Numerical Parameters for the `FASGE`

computational costs only scale with increasing numerical accuracy but not with the number of qubits. Taking the initial Hamilton operator A into account, the computational complexity scales with $\log(d)$, which is significant improvement compared to running a classical singular value decomposition.

However, numerical parameters will have to be adjusted more carefully, if we aim to study problems involving several *hundreds of qubits.* The heart of the FASGE is inherited in the title of this entire project: subordination. Running the Subordination Algorithm (Algorithm 1) for several values of the renormalized time s, builds the foundation of computing stochastic gaps. What happens next, boils down to data analysis of the resulting probability density. If the number of qubits increases, this very data analysis (Algorithm 3 and Algorithm 4) needs to be improved as Figure 5.14 and Figure 5.15 suggest: towards the end of the adiabatic evolution, the density oscillates rapidly, which cannot be detected if one considers the static probability density (Figure 5.14). Due to this very problem, a reliable gap analysis must involve dynamic extensions of η, ζ, ξ so that these oscillations are being taken into account. Nevertheless, the problem of determining spectra is not limited by too large Hilbert space dimensions any more! Figures 5.14 and 5.15 were generated within seconds (despite corresponding to probability densities involving 100 qubits).

Combining the Subordination Algorithm with dynamic gap estimating routines (i.e. a *dynamic FASGE*) should do the job. Here, one could make use of machine learning methods, which would open the door to another project: FAQC with ML.

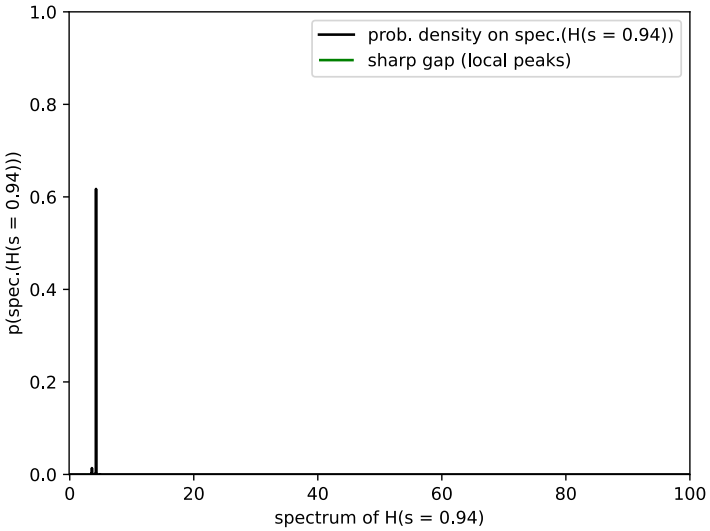
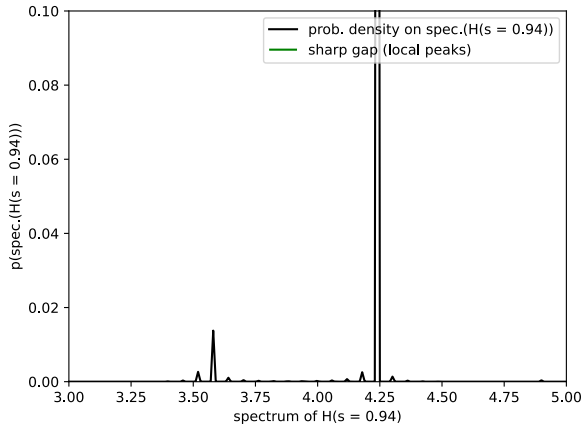
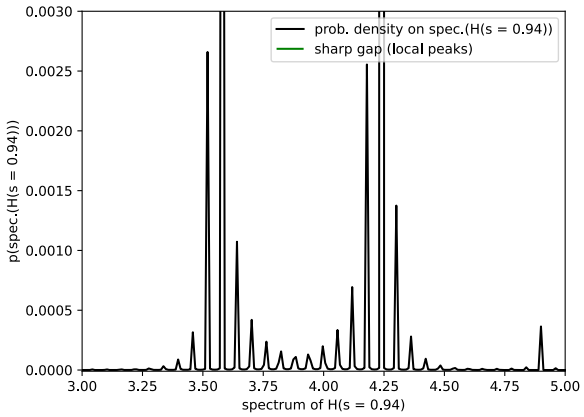


Figure 5.14.: Free adiabatic Grover $H_\Phi(s)$ from (5.44) with 100 qubits at $s = 0.94$. The gap becomes practically invisible, if the plot is generated over the entire spectrum. A closer look towards the gap must be taken, in order to determine the gap numerically.



(a) $s = 0.94$, a closer look towards gaps



(b) $s = 0.94$, rapid oscillations

Figure 5.15.: Free adiabatic Grover $H_\Phi(s)$ from (5.44) with 100 qubits at $s = 0.94$: rapid oscillations towards the gap. With respect to static parameters, the FASGE cannot detect the gap correctly. The plot in (a) would suggest a gap of $g \approx 0.6$, whereas the magnified plot in (b) is amibiguous, if we also consider the oscillations. At this point, dynamic parameters could solve the problem of rapidly oscillating probability densities.

6. Conclusion and Outlook

We started our research with focusing purely on random matrix theory. The quest (and *deterministic* barrier) of determining spectra corresponding to functions of operators, assuming that we know the (sub-) spectra of the involved operators, led us to the idea of considering a *randomized scenario:* can we add spectra of random matrices? It turned out that the answer to this question is embedded into a field of research on its own: free probability theory.

6.1. Summary

In this project, we have combined two fields of research: free probability theory and adiabatic quantum computing. Along this way, we focused on the special role that free probability plays, if one studies spectra of large random matrices.

Free probability is a non-commutative probability theory where random variables respect a modified notion of statistical independence, which is called *free independence.* For asymptotically large dimension, free probability is the limiting probability theory of (non-commuting) operators on Hilbert spaces: random matrices. For self-adjoint random matrices (Wigner matrices), there is a free version of the central limit theorem: Wigner's semi-circle law. It states that the distribution of the eigenvalues of random Hermitian matrices converges almost surely to the semicircular distribution. Due to the fact that this limiting spectral distribution is not random at all, one can develop a *deterministic* spectral theory for large *random* matrices. Motivated by the problem of adding spectra of random matrices, we studied a generalized additive convolution: *free convolution.* It turns out that the free convolution of spectral measures admits an analytic subordination formalism [BB07], which leads to a numerical fixed point algorithm: the *Subordination Algorithm.* This algorithm enables to compute the free additive convolution of spectral measures corresponding to asymptotically large random Hermitian matrices. In order to compute spectral densities of *functions* that take random matrices as *variables,* one must extend scalar-valued subordination theory [BMS17], if the required functions are polynomials instead of sums or linear combinations. Polynomials of random-matrix variables are treatable within operator-valued free probability theory. Starting from a block matrix of random matrices, we sketched a generalized setting of non-commutative probability spaces [Spe20a].

Moreover, we transferred the free probabilistic *Subordination Algorithm* to a spectral analysis of Hamilton operators suited to problems treated in adiabatic quantum computing. Since the QAA is a convex combination of two Hermitian operators, one can apply the (scalar-valued) *Subordination Algorithm*, provided that the operators are asymptotically freely independent. To generate this asymptotic freeness between the initial and final Hamilton operator, we applied a random unitary basis transformation to the problem Hamiltonian. We call this model the *Free QAA* (FQAA), resp. *Free Adiabatic Quantum Computation* (FAQC). By relying on subordination for free additive convolution, we could study spectral distributions of quantum adiabatic Hamiltonians without diagonalization.

Finally, we developed the FASGE (Algorithm 2) to compute stochastic spectral gaps, based on the *Subordination Algorithm* for linear combinations of random matrices. We demonstrated our

spectral gap analysis both for a generic example, and for the FQAA formulation of the Grover problem. Our results verify, how asymptotic freeness between the initial and final Hamiltonian activates a powerful probabilistic method for determining spectral gaps of multi-qubit Hamilton operators. Determining those gaps with a classical diagonalization routine, such as the singular value decomposition, is computationally infeasible for 20 qubits or more.

6.2. Applicability to Quantum Computation

Determining minimal spectral gaps is of utmost importance for the QAA [FGGS00]. Since the QAOA [FGG14] makes use of a Trotterization of the adiabatic evolution, any performance indicating quantity on QAA has direct implications on the QAOA. However, if our findings are used for experimental realization, we must discuss how we *changed* the QAA to our FQAA. The problem specific initial (A) and final operator B are fixed. It is emphasized that we did not change the spectrum of these operators. What we did change, is the relative orientation between their *eigenbases.* If we do this unitarily at random, this modification will generate asymptotic freeness between A and B. This means, that if we take A and B to be infinitely large (but nevertheless converging to their empirical spectral measures), free convolution can compute the spectral distribution of the convex combination *accurately.* However, Hamilton operators that are used for the QAA, are clearly *not infinitely large.* Therefore, we will always make an error, if we sorely rely on free convolution. The fact that this error scales with $\mathcal{O}(d^{-2})$ (Chapter 5) justifies our approach anyhow. Many plots in Chapter 4 and 5 support that this ambiguity does not cause problems in predicting spectral densities. Moreover, the error that one makes when stochastic spectral gaps are determined by our method instead of simulating random matrices, is of $\mathcal{O}(d^{-2})$. For Hilbert spaces of 20 qubits or more, this error is of $\mathcal{O}(10^{-13})$ or less, which should be enough for most applications.

6.2.1. Strengths

Spectral theory for asymptotically large operators is a helpful numerical "way out", if one needs to avoid computationally expensive diagonalization, which takes $\propto \mathcal{O}(d^3)$ operations. If the system under consideration admits to add more randomness artificially, or is already assumed to be randomly correlated, then free probability is *the* classical pre-/ post processing method to compute spectra. This enables a novel, classical approach of benchmarking the quantum system under investigation. That might be a canonical way to deal with noise. Moreover, the free QAA might be generalizable to free versions of the quantum alternating operator ansatz, and entire classes of variational quantum algorithms, including the variational quantum eigensolver [TCC+22]. Both classical and quantum (search) algorithms contain heuristics that inherently use Monte-Carlo methods, i.e. involve random steps. If spectral information is relevant, a combination with free probability could improve pre-and post processing.

6.2.2. Weaknesses

If we use the partial random unitary transformation to model noise, we encounter one problem. If noise is only obtained on minor parts of the system, this would mean that there are other perfectly prepared, noise-free parts, which is an odd assumption in practice. Moreover, it might not be clear whether the noise model is compatible with e.g. a random unitary transformation. On the other hand, inserting additional randomness could destroy the system entirely. If too much randomness (e.g. via noise) is assumed, the quantum system might lose its ability to

operate completely. In this case, the mathematical model might get too chaotic. In our FAQC model, we added a random unitary basis transformation. Random unitaries are experimentally hard to realize, but can nevertheless be approximated [EWS+03].

6.3. Outlook

In theory, the FASGE can estimate spectral gaps for hundreds, possibly thousands of qubits. This is achieved due to the logarithmic dependence on the matrix dimension. However, a practical implementation needs a dynamic adjustment of the numerical parameters. This is discussed in Chapter 5.5. If the minimal spectral gap is conjectured to be achieved in a small interval, one can run Algorithm 1 iteratively for those guess values and adjust parameters by hand. Such a brute force method enables to estimate the magnitude of the gap anyways. The spectral analysis in this project opens the door to analyze several more adiabatic Hamiltonians. The first choice is adiabatic Grover with more than one marked state. This corresponds to an unstructured database search with more than one marked item. Moreover, the adiabatic Travelling Salesman Problem (TSP) [Kie19] might be of interest. For complex problem Hamiltonians, such as for TSP, it remains to discuss whether the spectrum of the problem Hamiltonian is accessible beforehand. If the spectra of those Hamiltonians corresponding to combinatorial optimization problems are classically calculable, we might have developed a strong formalism to revive QAA once more.

6.3.1. Mathematical Research

Hermitian operators on Hilbert spaces are axiomatically contained in quantum physics. Naturally, our investigations were restricted to self-adjoint random matrices. However, RMT has emerged rapidly throughout the last two decades and is now a well-known and active field of research [Spe19c]. Besides Wigner's semi-circle law (for self-adjoint matrices), there are more central limiting distributions, such as the Marcenko-Pastur law for Wishart matrices. There also exist bounds on how much a finite-dimensional Wishart matrix deviates from its asymptotic limit [GT04]. Moreover, subordination theory is not restricted to real-valued probability measures. Therefore, it is possible to compute the free additive convolution also for complex measures. Furthermore, there are investigations if the topological assumption on the C^*-NCPS is *weakened* from a C^* to a W^* algebra. The generalization of the eigenvalue distribution for non-normal operators in a finite von Neumann algebra is the Brown measure [MS17, Chapter 11]. Furthermore, there more classes of functions, for which the (operator-valued) subordination algorithm [BMS17] can compute the spectral densities: higher-dimensional, self-adjoint polynomials and non-commutative *rational functions* [HMS18]. Consequently, computing spectra of several inversions of (self-adjoint) operators is also feasible within subordination theory, assuming asymptotic freeness is established. Whether subordination theory covers self-adjoint, *analytic* functions in free variables, is widely conjectured to be true, yet remains an open problem [MS17, Chapter 11].

A. Algorithms

The algorithms in this chapter complete the FASGE. In order to realize a reliable gap estimation, we propose two slightly different approaches, which aim to coincide qualitatively. Both algorithms were implemented in Python 3.11. Algorithm 3 sorely makes use of the numerical library `numpy`. Motivated by the problem specific gap analysis, we have developed this algorithm without further libraries. Algorithm 4 consists of a data evaluation routine from the scientific library `scipy`.

A.1. Gap Estimator

The algorithm 3 determines the gaps of a 1D input array of some data 'D'. It is assumed that D corresponds to the y-values of some tuple $(x, y(x))$, where $D \equiv y(x)$. We assume that $y(x) \geq 0$ for all x. However, it is not assumed that x is known. The data D correspond the probability distribution which is determined by the subordination algorithm 1. The 'Gap Estimator' computes the gap between the edges of the first two hills (from left to right). Figure A.1 demonstrates this. The parameters are:

- float $\eta > 0$: minimal required y-magnitude of the hill
- integer $L \in \mathbb{N}$: length of the data array 'D'

The magnitude η is not (!) the peak height. It relates to the minimal threshold that is allowed to be considered as "greater than zero". Therefore, the red gap is usually the more pessimistic gap estimate compared to the green gap.

Algorithm 3 `GE` (D)

Input: D (1D array of Data) ▷ L: length of this array
Input: **(optional)** η ▷ $\eta > 0$: minimal magnitude of non-zero values
  $p_1 \leftarrow$ empty list ▷ store the values corresponding to the first peak in a list
  for $i \in \{0, ..., L-1\}$ **do**
    if $D[i] \geq \eta$ **then**
      append $D[i]$ to p_1
    end if
    if $D[i] < \eta$ **and** $\#p_1 \geq 1$ **then** ▷ # denotes the length of the list p_1
      $i_1 = i$ ▷ store position
      break for
    end if
  end for
  $p_2 \leftarrow$ empty list ▷ store the values corresponding to the second peak in a list
  for $j \in \{i_1, ..., L-1\}$ **do**
    if $D[j] \geq \eta$ **then**
      append $D[j]$ to p_2
    end if
    if $D[j] < \eta$ **and** $\#p_2 \geq 1$ **then** ▷ # denotes the length of the list p_2
      $j_1 = j$ ▷ store position
      break for
    end if
  end for
  m $\leftarrow \frac{D[-1]-D[0]}{L-1}$ ▷ scaling factor; $[-1]$ denotes last value of the array
  $\rho_1 \leftarrow m \cdot i_1, \qquad \rho_2 \leftarrow m \cdot (j_1 - \#p_2)$
  $g \leftarrow |\rho_2 - \rho_1|$
Output: g, ρ_1, ρ_2 ▷ outputs the positions and the gap.

A.2. Peak Finder

This algorithm is based on the function `find_peaks` from the `signal` class in `scipy`. `find_peaks` is usually used for data analytic problems, such as determining peaks in electrocardiography [Doc23b]. This function determines local maxima by comparing neighboring elements. It also takes a 1D-array 'D' as input.

The optional parameters are:

- float $\zeta > 0$: minimal required peak height (10^{-4}, by default)
- float $\xi \geq 0$: prominence (0, by default)

In contrast to Algorithm 3, the minimal height 'ζ' refers to the local maximum (purple peaks in Figure A.1). The gap which is computed by Algorithm 4 is visualized in green in Figure A.1.

Algorithm 4 `Peaks` (D)

Input: D (1D array of Data); **use** signal analysis library (e.g. from SciPy)
Input: **(optional)** ζ, ξ ▷ $\zeta > 0$: peak height, $\xi \geq 0$: prominence
peaks $\leftarrow$ scipy.signal.find_peaks(D, prominence $= \xi$, ζ) ▷ compute local peaks using SciPy
compute m ▷ m: scaling factor (see Algorithm 3)
scaled_peaks $\leftarrow m \cdot$ peaks ▷ determine positions of local peaks
$x_0 \equiv \arg(\lambda_0) \leftarrow$ scaled_peaks[0]
$x_1 \equiv \arg(\lambda_1) \leftarrow$ scaled_peaks[1]
$g \leftarrow |x_1 - x_0|$ ▷ compute the gap
Output: g, x_0, x_1 ▷ outputs positions of the peaks of λ_0 and λ_1 and their gap

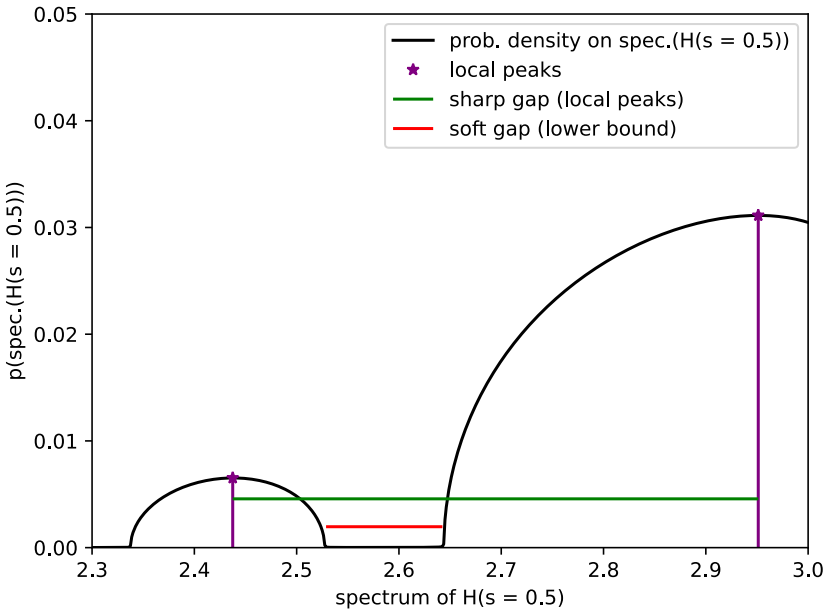


Figure A.1.: Visualization of the algorithms that determine the stochastic gaps. This example is taken from the generic Hamiltonian (5.37) at $s = 0.5$. To demonstrate, how the algorithms operate, the location around the minimal gap is magnified. Red gap (soft gap): the gap, where the probability of determining eigenvalues is zero; calculated by Algorithm 3. Green gap (sharp gap): the gap between local maxima corresponding to the two lowest eigenvalues; calculated by Algorithm 4. Compare with Definition 5.7

B. Supplementary Calculations

This calculation supports understanding the joint probability density function of a GUE matrix.

Corollary B.1. *Let $X \in M_N(\mathbb{C})$ and let X be Hermitian symmetric. Then,*

$$Tr(X^2) = \sum_{i=1}^{N} {x_{ii}}^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} |x_{ij}|^2$$

Proof. Direct computation: we write the matrix multiplication in index notation and we suppress the summation (up to N) by using Einstein summation convention. With $X = (x_{ij})_{i,j=1}^{N}$ for $x_{ij} \in \mathbb{C}$, the assertion reads:

$$\begin{aligned} \mathrm{Tr}(X^2) = \delta_i^k x_j^i x_k^j = x_j^i x_i^j &= (x_1^1)^2 + x_2^1 x_1^2 + x_3^1 x_1^3 + ... + x_N^1 x_1^N \\ &+ (x_2^2)^2 + x_1^2 x_2^1 + x_3^2 x_2^3 + ... + x_N^2 x_2^N \\ &+ ... \\ &+ (x_N^N)^2 + x_1^N x_N^1 + x_2^N x_N^2 + ... + x_{N-1}^N x_N^{N-1} \end{aligned}$$

By using $x_{ij} = \overline{x}_{ji}$ $\forall i, j \in [N]$ and $i \neq j$, one can write this as

$$\mathrm{Tr}(X^2) = \sum_{i=1}^{N} {x_{ii}}^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} |x_{ij}|^2,$$

which proves the assertion. □

Proof of Theorem 3.10. By definition,

$$\mathbb{E}[\mathrm{tr}\{A_N^m\}] = \frac{1}{N^{\frac{m}{2}+1}} \sum_{i(1),i(2),...,i(m)=1}^{N} \mathbb{E}[a_{i(1)i(2)} a_{i(2)i(3)} \cdot ... \cdot a_{i(m)i(1)}],$$

where the factor $\frac{1}{N^{\frac{m}{2}+1}}$ comes from the normalized trace together with m times applying the normalization constant $\frac{1}{\sqrt{N}}$ of the matrix A_N as defined for the Gaussian unitary ensemble. Recall that taking the trace is encoded in the last index being $i(1)$ whereas the middle indices encode the m times matrix multiplication. According to equation (3.9), it is left to determine, which terms in this product are (mutually) dependent and which are independent. Since independent random variables factorize under the expectation functional, all those terms vanish due to the initial assumption of centered random variables, i.e. $\mathbb{E}[a_{ij}] = 0$. Therefore, we apply the Wick-Isserlis formula (Theorem 3.9):

$$\mathbb{E}[a_{i(1)i(2)} a_{i(2)i(3)} \cdot ... \cdot a_{i(m)i(1)}] = \sum_{\pi \in \mathcal{P}_2(m)} \prod_{(k,l) \in \pi} \mathbb{E}[a_{i(k)i(k+1)} a_{i(l)i(l+1)}] \tag{B.1}$$

Let $(k_0, l_0) \in \pi$. Note that π is a permutation, i.e. $\pi(k_0) = l_0$ and $\pi(l_0) = k_0$. Applying equation (3.9), we get the following condition: $i(k_0) = i(l_0 + 1)$ and $i(l_0) = i(k_0 + 1)$. Only these terms have a non-zero contribution in equation (B.1):

$$\mathbb{E}[a_{i(1)i(2)}a_{i(2)i(3)} \cdot \ldots \cdot a_{i(m)i(1)}] = \sum_{\pi \in \mathcal{P}_2(m)} \prod_{(k,l)\in\pi} \delta_{i(k)i(l+1)}\delta_{i(k+1)i(l)}.$$

Writing $l_0 = \pi(k_0)$, this implies $i(k_0) = i(\pi(k_0) + 1)$. Thus, any contributing $k_0 \in [m]$ must be a fixed point of the mapping $s \in S_m, k \mapsto s(k) := (\pi(k) + 1) \mathrm{mod}\, m$. We define this mapping rigorously. For $\gamma, \pi \in S_m$, define $s \in S_m$ as $s := \gamma \circ \pi$, where π is any arbitrary permutation associated to a pair partition of the set $[m]$ and $\gamma \in S_m$ is the "shift by 1 operation" modulo m, i.e. $\gamma : [m] \to [m], k \mapsto \gamma(k) := (k + 1) \mathrm{mod}\, m$. In words, the above construction means that i (as a function for the summation indices $i : [m] \to [N]$) is constant on the cyclic permutations of $\gamma\pi$, i.e.:

$$\prod_{(k,l)\in\pi} \delta_{i(k)i(l+1)}\delta_{i(k+1)i(l)} = \prod_{k=1}^{m} \delta_{i(k)i(\gamma\pi(k))} = \begin{cases} 1, & i|_{\gamma\pi(k)} = \text{const.} \\ 0, & \text{else.} \end{cases}$$

Hence, only the number of these cycles has a non-zero contribution in the product. Then,

$$\begin{aligned}
\mathbb{E}[\mathrm{tr}\{A_N^m\}] &= \frac{1}{N^{\frac{m}{2}+1}} \sum_{i(1),i(2),\ldots,i(m)=1}^{N} \sum_{\pi\in\mathcal{P}_2(m)} \prod_{(k,l)\in\pi} \delta_{i(k)i(l+1)}\delta_{i(k+1)i(l)} \\
&= \frac{1}{N^{\frac{m}{2}+1}} \sum_{\pi\in\mathcal{P}_2(m)} \sum_{i(1),i(2),\ldots,i(m)=1}^{N} \prod_{(k,l)\in\pi} \delta_{i(k)i(l+1)}\delta_{i(k+1)i(l)} \\
&= \frac{1}{N^{\frac{m}{2}+1}} \sum_{\pi\in\mathcal{P}_2(m)} \sum_{i(1),i(2),\ldots,i(m)=1}^{N} \prod_{k=1}^{m} \delta_{i(k)i(\gamma\pi(k))}
\end{aligned}$$

Since $\sum_{i(1),i(2),\ldots,i(m)=1}^{N} \prod_{k=1}^{m} \delta_{i(k)i(\gamma\pi(k))} = N^{\#(\gamma\pi)}$, we finally get

$$\mathbb{E}[\mathrm{tr}\{A_N^m\}] = \frac{1}{N^{\frac{m}{2}+1}} \sum_{\pi\in\mathcal{P}_2(m)} N^{\#(\gamma\pi)}$$

□

C. Basic Notions from Group Theory

Moments of several random matrix ensembles require to take sums over elements of the permutation group. Moreover, the theory of unitary random matrices includes integration over the compact, unitary group. Therefore, we review some elementary definitions from group theory. This section follows the group theoretic introduction presented by Benjamin Sambale [Sam17] and part II, "advanced mathematical theory and applications of special relativity" of D. Giulini's lecture notes on "Foundations and Applications of Special Relativity" [Giu21].

We start by introducing the notion of a (finite) group.

Definition C.1 (Group)**.** *A **group** is pair $(G, \cdot)$ consisting of a finite set G and an operation*

$$\cdot : G \times G \to G, \qquad (g, h) \mapsto g \cdot h$$

which satisfies the following:

- $\forall\ g, h, k \in G : (g \cdot h) \cdot k = g \cdot (h \cdot k)$ *(associativity)*
- $\exists e \in G : \forall g \in G : e \cdot g = g \cdot e = g$ *(neutral element)*
- $\forall g \in G : \exists g^{-1} \in G : g^{-1} \cdot g = g \cdot g^{-1} = e$ *(inverse element)*

If

- $\forall g, h \in G : g \cdot h = h \cdot g$ *(commutativity)*

*holds as well, then we call G an **Abelian group.***

The combinatorial description of joint moments of random matrices involves counting elements of the symmetric group $S_m := Sym(m)$. Since permutation groups are mainly dedicated to finding symmetries of a certain problem, that is formulated for instance over a set M, we define the action of a group G on a (non-empty) set M as follows:

Definition C.2. *The (left) **action** of a group G on a set M is a homomorphism $f : G \to Sym(M)$, i.e.*

$$f : G \times M \to M, \qquad (g, m) \mapsto f(g, m) := gm \in M \quad \forall g \in G, m \in M. \tag{C.1}$$

More precisely, the group homomorphism property means

1. $f(e, m) = m$
2. $f(g^{-1}, m) = f^{-1}(g, m)$
3. $f(g \cdot h, m) = f(g, \cdot) \circ f(h, \cdot)(m) = f(g, f(h, m))$.

$\forall g, h \in G,$ *and* $m \in M$.

The standard example of a permutation group is the symmetric group over $n \in \mathbb{N}$ elements, denoted by S_n. Although group actions are studied more carefully in the context of representation theory of groups, we characterize fixed points of group actions, that lead to the notions of **orbits** of group actions and the **stabilizer subgroup.** Specifically, the genus expansion of GUE(N) matrices includes counting the orbits of a permutation.

Definition C.3. *The set of elements that one can reach by applying $f(g,\cdot)$ to m_0 is called the **orbit** of m_0, denoted by O_{m_0},*

$$O_{m_0} := \{f(g, m_0) : g \in G\} \subset M \tag{C.2}$$

*A fixed point $m_0 \in M$ of a group action $f : G \to Sym(M)$ fulfills $f(g, m_0) = m_0$ for all $g \in G$. The subset of G which relates to this fixed point equation, i.e. the group elements that act as the identity on m_0 is called the **stabilizer subgroup** $G_{m_0} \subset G$,*

$$G_{m_0} := \{g \in G : f(g, m_0) = m_0\} \subset G \tag{C.3}$$

In the theory of unitary groups, integrals are evaluated with respect to the Haar measure. Evaluating these integrals regards to so called Weingarten formulae. In general, Weingarten calculus is used to evaluate integrals over **compact groups.** We introduce these groups closely following [Far08].

Definition C.4. *A **topological group** is a group equipped with a topology such that the map*

$$G \times G \to G, \qquad (g, h) \mapsto gh^{-1} \tag{C.4}$$

is continuous. If the respective topology is locally compact, then G is called a locally compact group.

The topological locally compactness is needed to make a unitary representation well-defined. If this was not assumed, the group representing homomorphism would map to "different" identities on the Hilbert space.

Definition C.5. *A Radon measure μ is said to be **left invariant,** if for all $f \in C_C(G)$, we have*

$$\int_G f(gx)\,\mu(dx) = \int_G f(x)\,\mu(dx). \tag{C.5}$$

For a Borel set $B \subset G$, this means $\mu(gB) = \mu(B)$ for all $g \in G$, where $gB := \{gb | b \in B\}$.

We combine the above notion to provide the existence of a **Haar measure.**

Theorem C.6. *[Far08, Theorem 5.1.1] Let G be a locally compact group. There exists a non-zero, left-invariant measure on G. It is unique up to a positive factor. This measure is called (left) Haar measure.*

As usual, let $\mathcal{H}$ be a finite-dimensional Hilbert space over $\mathbb{C}$.

Definition C.7. *A (finite-dimensional) **unitary representation** of a locally compact group G is a continuous group homomorphism $\rho : G \to GL(\mathcal{H})$.*

D. Complex Analysis in a Nutshell

Complex integrals are omnipresent when evaluating Cauchy-Stieltjes transforms. In RMT, many integral expressions were abstracted away from basic complex analysis, where the transition is about generalizing the complex Lebesgue measure $dz \equiv dx + i\,dy$ to the empirical spectral measure $d\mu$. In light of this construction, the Cauchy-Stieltjes transform boils down to a complex contour integral alongside the real line, that is computed using the residue theorem with a path over the upper (or lower) complex half plane. We review the required notions from complex analysis, following [FB00, Chapter 3].

D.1. The Cauchy Integral Formula

Throughout, let $D \subset \mathbb{C}$ be an open (and simply connected) domain and $z_0 \in D$. Let $B_r(z_0) \subset D$, with $B_r(z_0) := \{z \in D : |z - z_0| < r\}$. Moreover, let $\mathcal{C}$ denote a (simply) closed contour in D, which is not necessarily of a circular shape. Cauchy's integral theorem yields the foundation for the integral formula and the residue theorem. It states that complex integrals of analytic (i.e. holomorphic) functions f alongside a simply closed contour vanish:

Theorem D.1 (Cauchy Integral Theorem)**.** *Let f be analytic and $\mathcal{C} \subset D$. Then,*

$$\oint_{\mathcal{C}} f(z)\,dz = 0. \tag{D.1}$$

Lemma D.2. *For $z \in B_r(z_0)$, it holds*

$$\oint_{\gamma} \frac{d\zeta}{\zeta - z} = 2\pi i, \qquad \gamma(t) := z_0 + re^{it}, \quad t \in [0, 2\pi], \quad r > 0. \tag{D.2}$$

Theorem D.3 (Cauchy Integral Formula)**.** *Let $f : D \to \mathbb{C}$ be analytic, and let $\overline{B}_r(z_0) \subset D$. Then, for any $z \in B_r(z_0)$:*

$$f(z) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(\zeta)}{\zeta - z}\,d\zeta, \tag{D.3}$$

where γ is again a closed curve, as in Lemma D.2

The Cauchy integral formula enables to compute the values of an analytic function within a disk by evaluating the function on the edge of that disk. Moreover, the Cauchy Integral Formula can be generlalized to evaluate not only the function f, but also all its derivatives. The analyticity of f implies that the n'th derivative exists, and is again analytic.

$$f^{(n)}(z) = \frac{n!}{2\pi i} \oint_{\gamma} \frac{f(\zeta)}{(\zeta - z)^{n+1}}\,d\zeta, \qquad n \in \mathbb{N}. \tag{D.4}$$

D.2. The Residue Theorem

The residue theorem might be the (!) fundamental theorem of complex analysis. It allows to evaluate integrals alongside the real line analytically by constructing a contour through a complex domain. However, this construction certainly would not work while restricting purely to functions on the real line without extension to $\mathbb{C}$. As an implication from the Cauchy Integral Formula, the **winding number** $\chi(\gamma, z)$ is defined by

$$\chi(\gamma, z) := \frac{1}{2\pi i} \oint_{\gamma} \frac{1}{\zeta - z}\, d\zeta, \tag{D.5}$$

which counts how often γ circulates around some $z \in \text{int}(\gamma)$.

Definition D.4. *Let $z_0 \in \mathbb{C}$ be an isolated singularity of the analytic function f and let*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \tag{D.6}$$

*the corresponding Laurent series expansion around z_0. Then, the **residue** of f at position z_0 is defined by the coefficient a_{-1} of the Laurent series.*

Several properties of the Cauchy transform follow by an application of the Laurent decomposition around the support of the spectral measure. The residue theorem justifies the existence of G_μ. Hence, the residue can be calculated via

$$\text{Res}\,(f; z_0) = \frac{1}{2\pi i} \oint_{|\zeta - z_0| = r_0} f(\zeta)\, d\zeta, \tag{D.7}$$

for some small $r_0 > 0$. We present Cauchy's residue theorem, following [FB00, Theorem 6.3]

Theorem D.5 (Residue Theorem)**.** *Let $D \subset \mathbb{C}$ be a simply connected and open domain and let $z_1, ..., z_k \in D$ be finitely many pairwise distinct points. Moreover, let $f : D \setminus \{z_1, ..., z_k\} \to \mathbb{C}$ be an analytic function and $\gamma : [a, b] \to D \setminus \{z_1, ..., z_k\}$ a closed, piece-wise smooth curve. Then,*

$$\oint_{\gamma} f(\zeta)\, d\zeta = 2\pi i \sum_{j=1}^{k} Res\,(f; z_j)\chi(\gamma; z_j). \tag{D.8}$$

If there exists an analytic continuation for f within the points z_j, then $\oint_\gamma f(\zeta)\, d\zeta = 0$, thus the residue theorem is a generalization of the Cauchy integral theorem.

The residue theorem is applied to computing integrals over fractions of rational functions and - as already mentioned - evaluating "improper integrals" of the form $\int_{\mathbb{R}} f(x)\, dx \equiv \int_{-\infty}^{\infty} f(x)\, dx$. The existence of those integrals is established by the existence of the Cauchy principal value, so the above integral reads

$$\int_{-\infty}^{\infty} f(x)\, dx = \lim_{R \to \infty} \int_{-R}^{R} f(x)\, dx.$$

Thus, computing improper integrals follows via shifting residues (w.l.o.g. choose the upper half plane) into $\mathbb{C}^+$, and computing

$$\lim_{R \to \infty} \int_{-R}^{R} f(x)\, dx = 2\pi i \sum_{i=1}^{k} \text{Res}\,(f; a_i). \tag{D.9}$$

That is why the Cauchy transform G_μ takes values in $\mathbb{C}^+$. Note that - due to the compact support of the spectral measures on $\mathbb{R}$ - this indeed corresponds to evaluating an improper integral, as described above.

E. Failed Attempts to Unite QAA with Random Matrices

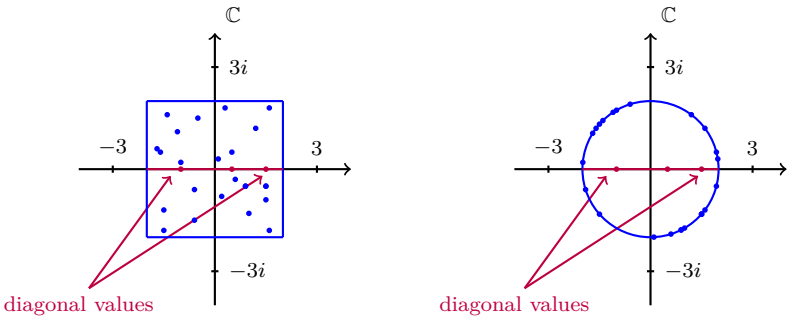
We have constructed a new approach to the QAA by unitarily rotating one operator at random. However, this project started without any idea of free probability and its powerful tools but sorely on trying to combine random matrices with QAA. These initial (failed) attempts are sketched in the following.

E.1. Empirical Deformed Matrix Entries on Geometric Shapes

This project started with the idea of inserting randomness (somehow) into the Quantum Adiabatic Evolution. The level repulsion in the spectral density function of the Gaussian ensembles is a well-known feature in RMT. The fact that distinct energy levels repel each other could make random matrices suitable candidates to be used for Hamilton operators in QAA. At that point, we started a series of simulations, where we inserted randomly generated, deformed matrices into the QAA and analyzed their spectrum. Some plots are presented below.

However, the objects that we inserted into the QAA were not random matrices at all. We accidentally studied some deformed, randomized matrices, where - despite Hermitian symmetry - we drew a square, respectively a circle in $\mathbb{C}$ and generated the matrix entries uniformly at random along these shapes. Since diagonal values are necessarily real, we considered operators where the diagonal values were positioned on the real line (i.e. the middle of the respective shape).

The central limit theorem in RMT is Wigner's semicircle law (see Figure 3.1). If we consider a convex combination of two non-normalized GUE(d) matrices with dimension $d = 2^7$ (Figure E.3), we only consider the convex combination of two semicircle distributions which is maximally deformed for $s = 0.5$. From spectral theory's point of view, nothing interesting happens.



(a) (b)

Figure E.1.: Two types of deformed randomized matrices. The off-diagonal matrix entries (blue) are drawn uniformly at random according to (a) a rectangle and (b) a circle. Due to Hermitian symmetry, the diagonal values are self-adjoint and real (purple).

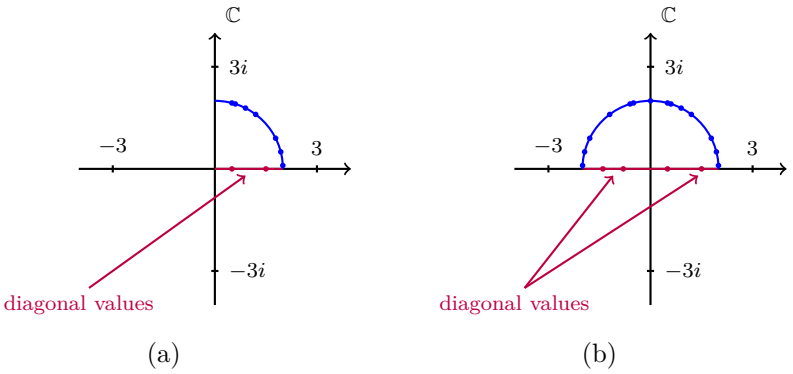


Figure E.2.: Deformed, random matrices generated from circle segments: uniform random values constrained on circle segments: (a) upper right quarter circle and (b) upper half circle.

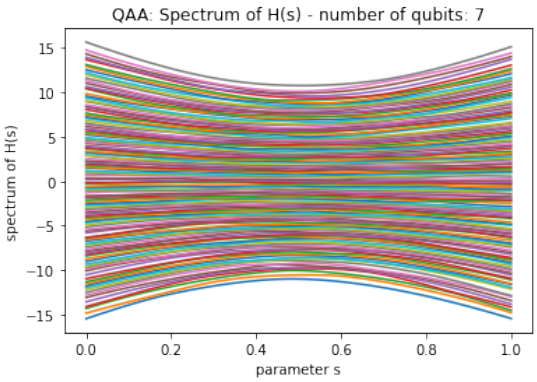


Figure E.3.: Convex combination of two non-normalized "GUE-like" random matrices. Due to a high density of states (for growing dimension), we cannot hope to see notable spectral gaps or anything worth mentioning.

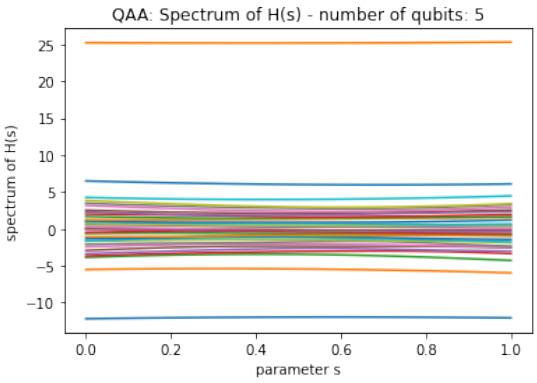


Figure E.4.: Convex combination of two deformed, random matrices with values generated from circle segments. The matrix entries are uniformly distributed along a circle segment, i.e. complex random variables with absolute value $r = 2$. Somehow surprisingly, we get a very streamlined spectrum.

Bibliography

[AL18] Tameem Albash and Daniel A Lidar. Adiabatic quantum computation. *Reviews of Modern Physics*, 90(1):015002, 2018.

[Bau22] Wolfram Bauer. *Lecture Notes on "Functional Analysis"*. Leibniz University Hannover, 2022.

[BB07] Serban T. Belinschi and Hari Bercovici. A new approach to subordination results in free probability. *Journal d'Analyse Mathématique*, 101(1):357–365, 2007.

[BES20] Zhigang Bao, László Erdős, and Kevin Schnelli. On the support of the free additive convolution. *Journal d'Analyse Mathematique*, 142(1):323–348, 2020.

[Bia98] Philippe Biane. Processes with free increments. *Mathematische Zeitschrift*, 227:143–174, 1998.

[BMS17] Serban T. Belinschi, Tobias Mai, and Roland Speicher. Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2017(732):21–53, 2017.

[CDF+02] Andrew M. Childs, Enrico Deotto, Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Andrew J. Landahl. Quantum search by measurement. *Physical Review A*, 66(3):032314, 2002.

[CGP19] Benoît Collins, Alice Guionnet, and Félix Parraud. On the operator norm of non-commutative polynomials in deterministic matrices and i.i.d. gue matrices. *arXiv preprint* https://arxiv.org/abs/1912.04588, 2019.

[CMN22] Benoit Collins, Sho Matsumoto, and Jonathan Novak. The weingarten calculus. *arXiv preprint* https://arxiv.org/abs/2109.14890, 2022.

[CN16] Benoît Collins and Ion Nechita. Random matrix techniques in quantum information theory. *Journal of Mathematical Physics*, 57(1), 2016.

[Doc23a] SciPy Documentation. API Reference, class for the unitary group. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.unitary_group.html, 2023. [accessed 06-October-2023].

[Doc23b] SciPy Documentation. API Reference, signal class. https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html, 2023. [accessed 24-October-2023].

[EW13] Alan Edelman and Yuyang Wang. Random matrix theory and its innovative applications. In *Advances in Applied Mathematics, Modeling, and Computational Science*, pages 91–116. Springer, 2013.

[EWS+03] Joseph Emerson, Yaakov S. Weinstein, Marcos Saraceno, Seth Lloyd, and David G. Cory. Pseudo-random unitary operators for quantum information processing. *science*, 302(5653):2098–2100, 2003.

[Far08] Jacques Faraut. *Analysis on Lie Groups: An Introduction*. Cambridge University Press, New York, 2008.

[FB00] Eberhard Freitag and Rolf Busam. *Funktionentheorie 1*. Springer-Verlag, Berlin, Heidelberg, 2000.

[FGG14] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint* https://arxiv.org/abs/1411.4028, 2014.

[FGGS00] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Quantum computation by adiabatic evolution. *arXiv preprint* https://arxiv.org/abs/quant-ph/0001106, 2000.

[Geo15] Hans-Otto Georgii. *Stochastik: Einführung in die Wahrscheinlichkeitstheorie und Statistik*. Walter de Gruyter GmbH & Co KG, 2015.

[Giu21] Domenico Giulini. *Lecture Notes on "Foundations and Applications of Special Relativity"*. Leibniz University Hannover, 2021.

[GM20] Alice Guionnet and Mylène Maïda. Large deviations for the largest eigenvalue of the sum of two random matrices. *arXiv preprint* https://arxiv.org/abs/1810.02538, 2020.

[Gro96] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 212–219, 1996.

[GT04] Friedrich Götze and Alexander Tikhomirov. Rate of convergence in probability to the marchenko-pastur law. *Bernoulli*, 10(3):503–548, 2004.

[Hal13] Brian C. Hall. *Quantum Theory for Mathematicians*. Springer Science & Business Media, New York, 2013.

[Hei20] Tim F. Heine. Bachelor thesis on "the resource theory of steering". *Leibniz University Hannover, Quantum Information Group*, 2020.

[HKL+20] Philipp Hauke, Helmut G. Katzgraber, Wolfgang Lechner, Hidetoshi Nishimori, and William D. Oliver. Perspectives of quantum annealing: Methods and implementations. *Reports on Progress in Physics*, 83(5):054401, 2020.

[HM23] Jonas Hofmann and Tobias Mai. Non Commutative Distributions, GitHub Repository. https://github.com/johannes-hoffmann/NCDist.jl, 2023. [accessed 25-October-2023].

[HMS18] J. William Helton, Tobias Mai, and Roland Speicher. Applications of realizations (aka linearizations) to free probability. *Journal of Functional Analysis*, 274(1):1–79, 2018.

[HT05] Uffe Haagerup and Steen Thorbjørnsen. A new application of random matrices: Ext(c(f2)) is not a group. *Annals of Mathematics*, pages 711–775, 2005.

[HWO+19] Stuart Hadfield, Zhihui Wang, Bryan O'gorman, Eleanor G. Rieffel, Davide Venturelli, and Rupak Biswas. From the quantum approximate optimization algorithm to a quantum alternating operator ansatz. *Algorithms*, 12(2):34, 2019.

[Ize08] Alan J. Izenman. Introduction to random-matrix theory. *The Ohio State Unviersity,* https://www.asc.ohio-state.edu/statistics/dmsl/Introduction_to_Random_Matrix_Theory.PDF, 2008.

[Jek18] David Jekel. Operator-valued non-commutative probability. *Preprint, available at* https://davidjekel.com/wp-content/uploads/2019/07/OVNCP_06.pdf, 2018.

[Jia21] Tianchong Jiang. Wigner's semicircle law for gaussian random matrices. *Uniersity of Chicago, Department of Mathematics, available at* https://math.uchicago.edu/~may/REU2021/REUPapers/Jiang,Tianchong.pdf, 2021.

[Kar12] Vladislav Kargin. A concentration inequality and a local law for the sum of two random matrices. *Probability Theory and Related Fields*, 154(3-4):677–702, 2012.

[Kie19] Tien D Kieu. The travelling salesman problem and adiabatic quantum computation: an algorithm. *Quantum Information Processing*, 18(3):90, 2019.

[Liu21] Weihua Liu. Relations between convolutions and transforms in operator-valued free probability. *Advances in Mathematics*, 390:107949, 2021.

[Mai17] Tobias Mai. *On the analytic theory of non-commutative distributions in free probability*. PhD thesis, Saarland University, 2017.

[MALW05] David R. Mitchell, Christoph Adami, Waynn Lue, and Colin P. Williams. Random matrix model of adiabatic quantum computing. *Physical Review A*, 71(5):052324, 2005.

[Meh04] Madan Lal Mehta. *Random Matrices*. Elsevier, Amsterdam, 2004.

[MMWW10] Gerhard Merziger, Günter Mühlbach, Detlef Wille, and Thomas Wirth. *Formeln und Hilfen zur höheren Mathematik*, volume 2. Binomi Verlag, 2010.

[Mor17] Valter Moretti. *Spectral Theory and Quantum Mechanics*. Springer International Publishing AG, 2017.

[MS17] James A. Mingo and Roland Speicher. *Free probability and random matrices*, volume 35. Springer Science & Business Media, New York, 2017.

[MŚS07] James A. Mingo, Piotr Śniady, and Roland Speicher. Second order freeness and fluctuations of random matrices: Ii. unitary random matrices. *Advances in Mathematics*, 209(1):212–240, 2007.

[Nec18] Ion Nechita. Weingarten Calculus and Applications to Quantum Information Theory, Seminar's Talk, Tübingen. https://www.math.uni-tuebingen.de/user/keilen/download/Conferences/SFBTRR195/TRRmeetingNechita.pdf, 2018.

[NS06] Alexandru Nica and Roland Speicher. *Lectures on the combinatorics of free probability*, volume 13. Cambridge University Press, 2006.

[Par21] Félix Parraud. *Random Matrices and Quantum Information Theory*. PhD thesis, University of Lyon; Kyoto University, 2021.

[Par22] Félix Parraud. On the operator norm of non-commutative polynomials in deterministic matrices and iid haar unitary matrices. *Probability Theory and Related Fields*, 182(3-4):751–806, 2022.

[RSDC23] Atanu Rajak, Sei Suzuki, Amit Dutta, and Bikas K. Chakrabarti. Quantum annealing: An overview. *Philosophical Transactions of the Royal Society A*, 381(2241):20210417, 2023.

[Sam17] Benjamin Sambale. *Endliche permutationsgruppen*. Springer Fachmedien, Wiesbaden, 2017.

[Sch07] Franz Schwabl. *Quantenmechanik (QM I): Eine Einführung*. Springer-Verlag, Berlin, Heidelberg, 2007.

[Sch08] Franz Schwabl. *Quantenmechanik für Fortgeschrittene (QM II)*. Springer-Verlag, Berlin, Heidelberg, 2008.

[Sch23] Roland Schnaubelt. *Lecture Notes on "Spectral Theory"*. Karlsruher Institut für Technologie, https://www.math.kit.edu/iana3/~schnaubelt/media/st-skript.pdf, 2023.

[Sho94] Peter W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee, 1994.

[Sim17] Barry Simon. Tosio kato's work on non–relativistic quantum mechanics. *arXiv preprint* *https://arxiv.org/abs/1711.00528*, 2017.

[Spe09] Roland Speicher. Free probability theory. *arXiv preprint* *https://arxiv.org/abs/0911.0087*, 2009.

[Spe15] Roland Speicher. Polynomials in asymptotically free random matrices. *arXiv preprint* *https://arxiv.org/abs/1505.04337*, 2015.

[Spe19a] Roland Speicher. Free Probability Theory, online lecture videos. https://www.youtube.com/watch?v=6MN3gTnJHBw&list=PLY11JnnnTUCYZni2Q7QNVa9hPGu77GK4M, 2019. [accessed March 2023].

[Spe19b] Roland Speicher. *Lecture Notes on "Free Probability Theory"*. Saarland University, arXiv preprint https://arxiv.org/abs/1908.08125, 2019.

[Spe19c] Roland Speicher. *Lecture Notes on "Random Matrices"*. Saarland University, arXiv preprint https://arxiv.org/abs/2009.05157, 2019.

[Spe19d] Roland Speicher. Random Matrices, online lecture videos. https://www.youtube.com/watch?v=fQQimr0rZ5k&list=PLY11JnnnTUCZSUk2xnombjwfg6z3-okP5, 2019. [accessed December 2022].

[Spe20a] Roland Speicher. *Lecture Notes on "Non-Commutative Distributions"*. Saarland University, arXiv preprint https://arxiv.org/abs/2009.03589, 2020.

[Spe20b] Roland Speicher. Non-Commutative Distributions, online lecture videos. https://www.youtube.com/watch?v=MvL1B3RNmOY&list=PLY11JnnnTUCaSoevQTJL8YAf4xQ7buyv8, 2020. [accessed May 2023].

[Tao23] Terence Tao. *Topics in random matrix theory*, volume 132. American Mathematical Society, 2023.

[TCC+22] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth, et al. The variational quantum eigensolver: a review of methods and best practices. *Physics Reports*, 986:1–128, 2022.

[TV+04] Antonia M. Tulino, Sergio Verdú, et al. Random matrix theory and wireless communications. *Foundations and Trends in Communications and Information Theory*, 1(1):1–182, 2004.

[TV08] Terence Tao and Van Vu. Random matrices: the circular law. *Communications in Contemporary Mathematics*, 10(02):261–307, 2008.

[Voi91] Dan-Virgil Voiculescu. Limit laws for random matrices and free products. *Inventiones mathematicae*, 104(1):201–220, 1991.

[Voi94] Dan-Virgil Voiculescu. The analogues of entropy and of fisher's information measure in free probability theory, ii. *Inventiones mathematicae*, 118(1):411–440, 1994.

[Voi95] Dan-Virgil Voiculescu. Free probability theory: random matrices and von neumann algebras. In *Proceedings of the International Congress of Mathematicians: August 3–11, 1994 Zürich, Switzerland*, pages 227–242. Springer, 1995.

[Voi97] Dan-Virgil Voiculescu. *Free Probability Theory*, volume 12. American Mathematical Society, 1997.

[Voi10] Dan-Virgil Voiculescu. *Free analysis questions II: the Grassmannian completion and the series expansions at the origin*. Walter de Gruyter GmbH & Co. KG, 2010.

[Wer18] Dirk Werner. *Funktionalanalysis*. Springer-Verlag, Berlin, Heidelberg, 2018.

[Wig67] Eugene P. Wigner. Random matrices in physics. *SIAM review*, 9(1):1–23, 1967.

[Wil17a] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2017.

[Wil17b] John D. Williams. Analytic function theory for operator-valued free probability. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2017(729):119–149, 2017.

List of Figures

List of Symbols and Abbreviations

Mathematical Symbols

$(\mathcal{A}, \mathcal{B}, E)$ operator-valued non-commutative probability space

$(\mathcal{A}, \varphi)$ non-commutative probability space

$(\Omega, \mathcal{F}, \mathbb{P})$ (classical) probability space

$(M_d(L^{\infty-}(\Omega, \mathbb{P})), \mathrm{tr} \otimes \mathbb{E})$ *-probability space of random matrices with state $\varphi_d \equiv \mathrm{tr} \otimes \mathbb{E}$

0^+ values slightly bigger than 0

$\#(\cdot)$ number of cycles of a permutation

$\alpha(\cdot), \beta(\cdot)$ adiabatic schedule that tunes the adiabatic evolution

α convergence accuracy in `Sub`

$\boxplus$ free additive convolution

δ_0 Dirac delta distribution at 0

δ_{ij} Kronecker delta symbol

diag diagonal matrix

η peak height from `GE`

γ piecemeal differentiable curve in $\mathbb{C}$

$\hat{P}$ Linearization of a self-adjoint polynomial P

κ_n n'th free cumulant

$\lambda(s)$ eigenvalue curve depending on s

λ, λ_i eigenvalues

$\lambda_i^{\times m_i}$ eigenvalue λ_i with algebraic multiplicity m_i

$\Lambda_\varepsilon(z)$ shift in the operator-valued upper half plane

$\mathbb{C}\langle x_1, ..., x_k \rangle$ $\mathbb{C}-$ valued polynomials in non-commuting variables

$\mathbb{C}^+$ upper complex half plane

$\mathbb{E}[\cdot]$ classical expectation value

$\mathbb{K} = \{\mathbb{R}, \mathbb{C}\}$ number field: real, resp. complex numbers

$\mathbb{V}_{\varphi}(\cdot)$ variance of the random variable $(\cdot)$ w.r.t. φ.

$\mathcal{A}$ unital algebra over $\mathbb{C}$

$\mathcal{A}^{+}$ algebra of positive operators

$\mathcal{B}(\cdot)$ bounded linear operators over $(\cdot)$

$\mathcal{B}^{+}$ operator-valued upper half plane of the (sub-) algebra $\mathcal{B}$

$\mathcal{CN}(0,1)$ complex normal distribution with mean 0 and variance 1

$\mathcal{D}$ iteration domain

$\mathcal{F}_{(\cdot)}$ CDF of $(\cdot)$

$\mathcal{H}$ finite-dimensional Hilbert space

$\mathcal{L}(\mathcal{A},\mathcal{B})$ linear operators from $\mathcal{A}$ to $\mathcal{B}$

$\mathcal{M}(\cdot)$ measurable functions on $(\cdot)$

$\mathcal{M}_b(\cdot)$ bounded measurable functions on $(\cdot)$

$\mathcal{NC}_2$ non-crossing pairings

$\mathcal{N}(0,1)$ normal distribution with mean 0 and variance 1

$\mathcal{N}(\cdot)$ number of eigenvalues in the interval $(\cdot)$

$\mathcal{O}(\cdot)$ Landau notation

$\mathcal{P}([n])$ power set of the set $[n] := \{1, 2, ..., n\}$

$\mathcal{P}_2(m)$ non-crossing pairings of $[m]$

$\mathcal{P}_2$ pairings

$\mathcal{S}$ numerical partitioning of the adiabatic parameter (in `FASGE`)

$\mathcal{T}$ run time

$\mathcal{U}(d)$ set of random unitary matrices of dimension d

$\mathbb{1}_{(\cdot)}$ identity operator on $(\cdot)$

$\mathfrak{B}(\cdot)$ Borel σ-algebra over $(\cdot)$

$\mu(\cdot)$ Radon measure of $(\cdot)$

μ, ν probability measures

μ_A, μ_B empirical spectral measures of the operators A, B

μ_s, ν_s spectral measures depending on s

ω (analytic) subordination function

$\Phi_U(\cdot)$ random unitary transformation of $(\cdot)$

π pair partition

$\rho(s,t)$ (asymptotic) joint probability density function

$\sigma(A)$ spectrum of operator A

$\mathrm{supp}(\mu)$ support of measure μ

$\mathrm{GUE}\,(d)$ Gaussian Unitary Ensemble of dimension d

$\tilde{\mathcal{D}}$ shifted iteration domain in `Sub`

tr $\frac{1}{d}\,\mathrm{Tr}$, normalized trace functional

ε shift parameter in `Sub`

$\varphi(\cdot)$ linear functional, typically: state on C^* algebra

ξ prominence from `Peaks`

ζ peak height from `Peaks`

A, B finite-dim. operators on $\mathcal{H}$ (typically deterministic)

a, b non-commutative random variables

$C(\cdot)$ continuous functions from $(\cdot)$ (to $\mathbb{K}$)

$C^1(\cdot)$ continuously differentiable functions from $(\cdot)$ (to $\mathbb{K}$)

C_b bounded, continuous functions

C_m m'th Catalan number

d operator dimension; typical choice: $d = 2^N$ for N qubits

$D_{(\cdot)}$ diagonal matrix corresponding to $(\cdot)$

dg Haar measure

$E[\cdot]$ conditional expectation

$g(z,w)$ fixed point mapping of z, w

$g(z,w,s)$ subordination fixed point mapping, also depending on s

g, g' stochastic gaps

G compact group

$G_X(\cdot)$ operator-valued Cauchy transform

$g_{\min}$ minimal spectral gap

$G_\mu(\cdot)$ Cauchy transform at $(\cdot)$ w.r.t. measure μ

$H(s)$ adiabatic Hamilton operator depending on $s \in [0,1]$

$H_\Phi(s)$ free adiabatic Hamilton operator

$H_{N,s}$ sequence of deterministic/ random matrices, parameterized with the adiabatic parameter

H_N sequence of deterministic matrices, where one is unitarily rotated at random

$K, K_N, T, T_N, c, c_1, \varepsilon, \delta, \delta_s, \eta$ numerical constants, Chapter 5

L^a Lebesgue spaces, for the respective $a \in \{\infty-, p\}$

L_ρ set of local maxima of the probability density ρ

M maximal number of iterations in `Sub`

$M_d(\mathbb{C})$ Algebra of complex, d-dimensional square matrices

$M_d(\mathcal{A})$ $d-$dim. (matrix-) algebra over $\mathcal{A}$

$M_d(L^{\infty-}(\Omega, \mathbb{P}))$ Algebra of random matrices with distribution $\mathbb{P}$

$m_n(\mu)$ n'th moment of the (probability) measure μ

N number of equidistant steps for the interval partitioning in `Sub`

$P(X_1, ..., X_n)$ (typically self-adjoint) polynomial in n random variables

R Voiculescu's R-transform

$S_\mu(\cdot)$ Stieltjes transform at $(\cdot)$ w.r.t. measure

U random unitary matrix

u Haar unitary element

$U_T(s)$ unitary operator for the adiabatic evolution

V_1, V_2 deterministic unitary matrices

w_0 initial guess value for Algorithm 1

$w_{k_{\text{thresh}}}$ numerical fixed point in`Sub`

X Banach space

$X_1, ..., X_n$ n random variables

Z complex random variable

Abbreviations

AQC Adiabatic Quantum Computation

CDF Cumulative Distribution Function

FAQC Free Adiabatic Quantum Computation

FASGE Free Adiabatic Spectral Gap Estimator: `FASGE` (Algorithm 2)

FQAA Free Quantum Adiabatic Algorithm

GE Gap Estimator: `GE` (Algorithm 3)

GOE Gaussian Orthogonal Ensemble

GSE Gaussian Symplectic Ensemble

GUE Gaussian Unitary Ensemble

NCPS Non-Commutative Probability Space

PDF Probability Density Function

Peaks Peak Finder: `Peaks` (Algorithm 4)

QA Quantum Annealing

QAA Quantum Adiabatic Algorithm

QAO Quantum Alternating Operator Ansatz

QAOA Quantum Approximate Optimization Algorithm

QC Quantum Computers/ Quantum Computation

RM Random Matrix

RMT Random Matrix Theory

SU Special Unitary Group

SVD Singular Value Decomposition

Acknowledgments

First and foremost, I would like to thank my supervisors, Tobias J. Osborne and Reinhard F. Werner for helpful guidance and the opportunity, to do my master's research within Hannover's quantum information group. Moreover, I am thankful for many constructive discussions with QIG members, especially Lennart Binkowski, Gereon Koßmann (RWTH Aachen), Debora Ramaciotti, René Schwonnek, Shawn Skelton, and Timo Ziegler. Furthermore, I thank Lennart Binkowski, Leonhard Richter, Felix Wagner, and Timo Ziegler for critical reading.

Helpful advice and literature references from Roland Speicher, Tobias Mai (Saarland University) and Félix Parraud (KTH Royal Institute of Technology) are gratefully acknowledged.

Finally, I am deeply indebted to my entire family, especially to both Claudia and Rolf[1] Heine, and Erika[2] and Karl[3] Heine for great encouragement and (financial) support throughout my studies and beyond.

In this context, supportive conversations with Tim Obrock, Felix Wagner, Dirk Gronwald, Heiko Aumann and personal support by Petra Hahne, Anneliese[4] Hahne, and Elfriede Höltkemeyer are to be gratefully mentioned.

[1] † 03/2014
[2] † 07/2020
[3] † 09/2020
[4] † 09/2022