

## Assessing dimensions from evolution

Michael M. Wolf<sup>1</sup>, David Perez-Garcia<sup>2</sup>

<sup>1</sup> *Niels Bohr Institute, 2100 Copenhagen, Denmark*

<sup>2</sup> *Dpt. Análisis Matemático,*

*Universidad Complutense de Madrid, 28040 Madrid, Spain*

(Dated: September 3, 2009)

Using tools from classical signal processing, we show how to determine the dimensionality of a quantum system as well as the effective size of the environment's memory from observable dynamics in a model-independent way. We discuss the dependence on the number of conserved quantities, the relation to ergodicity and prove a converse showing that a Hilbert space of dimension  $D + 2$  is sufficient to describe every bounded sequence of measurements originating from any  $D$ -dimensional linear equations of motion. This is in sharp contrast to classical stochastic processes which are subject to more severe restrictions: a simple spectral analysis shows that the gap between the required dimensionality of a quantum and a classical description of an observed evolution can be arbitrary large.

In Quantum Information Science the dimension of the accessible Hilbert space has the character of a resource—larger dimensions mean potentially more powerful protocols. Various implementations deal with huge Hilbert spaces corresponding to ensembles of atoms [1] or molecules [2], or continuous degrees of freedom leading to an infinite dimensional space in the first place. But how many degrees of freedom are effectively used? Can we assess the dimension of the underlying system from observable data without assuming a detailed model description beforehand? These questions, partially motivated by the need of questioning and pinpointing the assumptions of security proofs in quantum key distribution [3], were recently addressed in the context of ‘non-local’ quantum correlations.

In [4] it was shown that a tripartite system can only yield violations of certain Bell inequalities of the order of  $\sqrt{d}$  if each subsystem has dimension at least  $d$ , and in [5, 6] the dimension dependence of correlations has been investigated in detail for bipartite systems. As these approaches are based on static correlations between several parts, the question has been raised [5] whether and how the dimension of a single system can be observed.

Similarly, one might want to have a preferably model-independent way of assessing the effective dimensionality of the systems environment (quantifying non-Markovianity [7]) or the number of preserved, ‘noiseless’ degrees of freedom.

The present work addresses these question from a dynamical point of view. Given a discrete time evolution of an expectation value, we ask what can be inferred about the effective dimension of the systems Hilbert space or the environments memory. We thereby focus on using as little a priori information as possible. When addressing the systems dimensionality, the only assumptions are that the evolution is homogenous in time and Markovian in the sense that it is performed on time scales large compared to the relevant relaxation times of the systems environment. If the latter is not fulfilled we will see the environmental memory degrees of freedom in our dimension count.

For this analysis two standard tools from classical signal processing [8] can be employed: delayed embeddings and

analysis in the frequency domain. These will allow us not only to tackle the above question but also to address the converse: which Hilbert space dimension is sufficient for a given sequence of measurements? and to compare the efficiencies of quantum versus classical descriptions of a given evolution. While every sequence produced by quantum mechanical evolution can in principle be described by a classical stochastic process, we will easily see that quantum mechanics can be arbitrarily more efficient in terms of the required number of discrete degrees of freedom.

Not surprisingly we will see a connection between the effective dimension and the number of conservation laws and get a glimpse on more (in particular spectral) information which can be obtained from the evolution.

### PRELIMINARIES

Our interest lies in the discrete time evolution of expectation values of the form

$$\langle A(t) \rangle = \text{tr} [AT^t(\rho)], \quad (1)$$

where  $\rho$  is a density matrix acting on a  $d$ -dimensional Hilbert space,  $T$  is quantum channel, i.e., a trace-preserving completely positive linear map with equal input and output space and  $A = A^\dagger$  an observable. For simplicity we will in the first part assume that (1) is a half-infinite sequence, i.e.,  $t \in \mathbb{N}_0 = \{0, 1, \dots\}$ ; the extension to finite sequences and noisy data will then be discussed at the end.

Note that the description in (1) assumes homogeneity in time and *Markovness* in the sense that future evolution only depends on the state at present and not on its history. This means that given the l.h.s. of (1) the Hilbert space underlying the description of the r.h.s. has to contain all effectively relevant degrees of freedom. Hence, if a system of dimension  $d_S$  undergoes a non-Markovian evolution due to  $d_E$  memory degrees of freedom in the environment, then  $d = d_S + d_E$ .

It will sometimes be advantageous to consider  $\rho$  and  $A$  as elements of a  $d^2$ -dimensional Hilbert space  $\mathcal{H}$  equipped with the *Hilbert-Schmidt* scalar product  $\langle A|\rho \rangle := \text{tr} [A^\dagger \rho]$ . As a

linear map  $T$  has a matrix representation on  $\mathcal{H}$  which we will denote by  $\hat{T}$ . Using matrix units as a basis of  $\mathcal{H}$  we can write  $\hat{T} = \sum_k K_k \otimes \bar{K}_k$ , where  $\{K_k\}$  is a set of *Kraus operators* of  $T(\cdot) = \sum_k K_k \cdot K_k^\dagger$ . While  $T$  refers to time evolution in the Schrödinger picture, we will denote by  $T^*$  the respective map in the Heisenberg picture so that  $\text{tr}[AT(\rho)] = \text{tr}[T^*(A)\rho]$ .

We will denote by

$$\mathcal{C} := \{H = H^\dagger \mid \langle H(t) \rangle \text{ independent of } t \in \mathbb{N}_0\} \quad (2)$$

the space of *conserved quantities* which obviously includes all  $H = T^*(H)$  and in particular  $\mathbb{1} \in \mathcal{C}$  as the evolution is trace preserving. Note that  $\mathcal{C}$  depends on  $T$  and  $\rho$ .

A quantum channel  $T$  will be called *ergodic* w.r.t. a state  $\rho$  (an observable  $A$ ) if the orbit generated by  $T^t$  ( $T^{*t}$ ) spans the entire space of  $d \times d$  matrices.

### ASSESSING THE DIMENSION

The central tool in this section is the space  $\mathcal{V}$  spanned by all *delayed vectors* of the form

$$v_\tau = (\langle A(\tau) \rangle, \langle A(\tau+1) \rangle, \dots), \quad \tau \in \mathbb{N}_0. \quad (3)$$

The employed approach, often called *method of delays*, is particularly widespread in the analysis of chaotic dynamics [9] and it provides the following simple and tight relation:

**Proposition 1 (A bound on the dimensionality)** *Consider the space  $\mathcal{V} = \text{span}\{v_\tau\}_{\tau \in \mathbb{N}_0}$  spanned by the delayed vectors obtained from a sequence of the form (1). Then*

$$\dim \mathcal{C} + \dim \mathcal{V} \leq d^2 + 1, \quad (4)$$

where  $d$  is the dimension of the underlying Hilbert space and  $\dim \mathcal{C}$  the number of linearly independent conserved quantities. Equality holds in (4) if  $T$  is ergodic w.r.t.  $A$  and  $\dim \mathcal{V} = d^2$  iff it is ergodic w.r.t.  $A$  and  $\rho$ .

*Proof.* Consider a basis  $\{H_i\}$ ,  $i = 1, \dots, D := \dim \mathcal{C}$  of the space  $\mathcal{C}$  of conserved quantities. Then  $\{H_i - \text{tr}[\rho H_i] \mathbb{1}\}$  span a  $D - 1$  dimensional subspace of  $\mathcal{H}$  which is orthogonal to the space spanned by  $\hat{T}^t |\rho\rangle$ ,  $t \in \mathbb{N}_0$ . Hence the latter space, denote it by  $\mathfrak{H}$ , has dimension  $d^2 - D + 1$  and

$$\langle A(t) \rangle = \langle A_{\mathfrak{H}} | \hat{T}_{\mathfrak{H}}^t | \rho_{\mathfrak{H}} \rangle, \quad (5)$$

where the index  $\mathfrak{H}$  refers to the restriction onto  $\mathfrak{H}$ . Since the minimal polynomial [10] of  $\hat{T}_{\mathfrak{H}}$  has degree at most  $\dim \mathfrak{H}$  there are complex coefficients  $c_j$  such that

$$\hat{T}_{\mathfrak{H}}^{\dim \mathfrak{H}} = \sum_{j=0}^{d^2-D} c_j \hat{T}_{\mathfrak{H}}^j. \quad (6)$$

Recalling that  $(v_\tau)_k = \langle A(\tau+k-1) \rangle$  this implies that there are at most  $\dim \mathfrak{H}$  linearly independent vectors in  $\mathcal{V}$  since

$$(v_{d^2-D+x})_k = \sum_{j=0}^{d^2-D} c_j (v_j)_{x+k-1} \quad \forall x, k \in \mathbb{N}. \quad (7)$$

Let us now assume that  $T$  is ergodic w.r.t.  $A$ , which means that  $\{\langle A | T^t \rangle\}_{t \in \mathbb{N}_0}$  spans  $\mathcal{H}$ . Suppose that (4) would not be an equality. Then  $\sum_{j=0}^{d^2-D} c_j v_j = 0$  for some  $c$  which implies, by ergodicity, that  $\sum_{j=0}^{d^2-D} c_j \langle \hat{T}^{j+n} | \rho \rangle = 0$  for all  $n \in \mathbb{N}_0$ . Therefore  $\dim \mathfrak{H} \leq d^2 - D$  so that for the orthogonal complement  $\dim \mathfrak{H}^\perp \geq D$ . However, this contradicts the inequality  $\dim \mathfrak{H}^\perp \leq D - 1$  which comes from the fact that every element in  $\mathfrak{H}^\perp$  is in  $\mathcal{C}$  and in addition  $\mathbb{1} \in \mathcal{C} \setminus \mathfrak{H}^\perp$ . So, ergodicity w.r.t.  $A$  implies equality in (4).

If  $T$  is in addition ergodic w.r.t.  $\rho$  then  $\mathfrak{H} = \mathcal{H}$  implies  $D = 1$  so that indeed  $\dim \mathcal{V} = d^2$ . We finally prove the converse again by contradiction: assume linear dependence of the form  $\sum_{j=0}^{d^2-1} c_j v_j = 0$ . Then for all  $a, b \in \mathbb{N}_0$  we have  $\langle A | \hat{T}^a (\sum_j c_j \hat{T}^j) \hat{T}^b | \rho \rangle = 0$  so that, due to ergodicity,  $\hat{T}$  would have a minimal polynomial of degree less than  $d^2$ . However, this would again imply the existence of a proper subspace  $\mathfrak{H}$  contradicting the assumption  $\dim \mathcal{V} = d^2$ . To see this recall that a minimal polynomial of smaller degree requires an eigenvalue  $\lambda$  with geometric multiplicity larger than one [10]. Denoting by  $|\phi\rangle$  a linear combination of the corresponding left eigenvectors, we get  $\langle \phi | \hat{T}^t | \rho \rangle = \lambda^t \langle \phi | \rho \rangle$  such that there is always a  $\phi \perp \mathfrak{H}$ .  $\square$

*Some remarks.* Depending on the assumptions we may use the above result for different purposes: (i) assuming Markovianity it provides a lower bound for  $d$ , (ii) assuming we know  $d$  in addition it yields an upper bound on  $\dim \mathcal{C}$  and (iii) if we only know the dimension of the system  $d_S = d - d_E$  it gives a lower bound on the number  $d_E$  of effective memory degrees of freedom in the environment. In fact, if  $\langle A(t) \rangle$  exhibits algebraic decay then, as one would expect, Eq.(4) leads to  $d = \infty$ .

Prop. 1 is easily generalized to the case where, instead of a single expectation value, we observe a set of observables  $\{A_\alpha\}$  or, equivalently, take higher moments of the observable into account (i.e.,  $A_\alpha = A^\alpha$ ). Then the delayed vectors have to be replaced by ‘delayed matrices’ so that  $\langle A_\alpha(\tau) \rangle$  are the entries of the first column of the matrix  $v_\tau$ . Eq.(4) is then obtained in just the same way where  $\dim \mathcal{V}$  is now the number of linearly independent matrices.

### QUANTUM EVOLUTION FOR GIVEN SEQUENCES

The previous section provided a lower bound on the dimension of the Hilbert space in terms of the dimension of the space  $\mathcal{V}$  of delayed vectors which are in turn solely based on  $\langle A(t) \rangle$ . Clearly, there cannot be an upper bound to the ‘true’ dimension since there might always be redundant degrees of freedom. Nevertheless, we can provide a converse to the above observation in the sense that there always exists a quantum representation in a Hilbert space of dimension not much larger than  $\dim \mathcal{V}$ . Remarkably, such a converse does not exist for classical evolutions (see subsequent section).

#### Proposition 2 (Quantum representation)

*Given any bounded sequence  $\langle A(t) \rangle \in \mathbb{R}$ ,  $t \in \mathbb{N}_0$  there is*

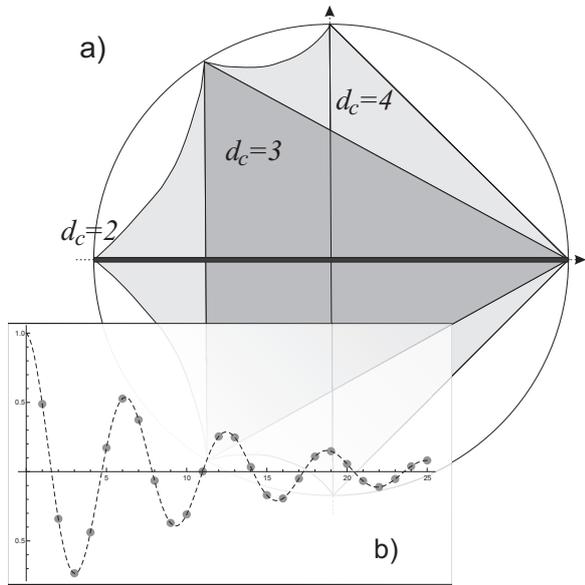


FIG. 1: a) Eigenvalues of the evolution operator lie on the unit disc and can be obtained from the observable time dependence of any expectation value. For a classical description they have to be located in a region depending on the number  $d_c$  of degrees of freedom. For  $d_c = 2$  this is the real line, for  $d_c = 3, 4$  the dark and light grey regions become additionally accessible. b) A simple damped oscillation leads to eigenvalues  $e^{-\gamma \pm i\omega}$ , implying that a quantum mechanical description can access the entire unit disc even for  $d = 2$ .

always a quantum state  $\rho$ , a quantum channel  $T$  and a Hermitian observable  $A$  acting on a Hilbert space of dimension  $\dim \mathcal{V} + 2$  such that (1) holds.

*Proof.* We begin with the fact proven in Lemma 1 in the appendix: there is always a contractive matrix  $M$  and vectors  $R, L$  of dimension  $\dim \mathcal{V}$  such that  $\langle A(t) \rangle = \langle L | M^t | R \rangle$ . We will proceed in two steps: first establish complete positivity by adding one degree of freedom and then impose the trace preserving property by adding another degree. Define a square ‘Kraus-operator’  $C = 1 \oplus M$  of dimension  $\dim \mathcal{V} + 1$  and, referring to the same block structure  $(\mathbb{C} \oplus \mathbb{C}^{\dim \mathcal{V}})$ , a vector  $|\Psi\rangle = |0 \oplus R\rangle + |1 \oplus 0\rangle$  and an ‘observable’  $B = (|1 \oplus 0\rangle\langle 0 \oplus L| + h.c.)/2$ . Then, using that  $\langle A(t) \rangle \in \mathbb{R}$  we have  $\langle L | M^t | R \rangle = \text{tr} [B C^t |\Psi\rangle\langle \Psi | C^{\dagger t}]$ . In order to make this trace preserving we embed it again, now referring to the block structure  $\mathbb{C} \oplus \mathbb{C}^{\dim \mathcal{V} + 1}$ . With  $K := 0 \oplus C, A := (0 \oplus B) / \|\Psi\|^2, \rho := (0 \oplus |\Psi\rangle\langle \Psi |) / \|\Psi\|^2$  we obtain that

$$T(\rho) = K \rho K^\dagger + |1 \oplus 0\rangle\langle 1 \oplus 0| \text{tr} [(1 - K^\dagger K) \rho] \quad (8)$$

indeed satisfies (1) for the chosen  $\rho$  and  $A$ , and since  $\|M\|_\infty \leq 1$  implies  $K^\dagger K \leq \mathbb{1}$ ,  $T$  is a valid quantum channel.

□

We do not know whether there are more efficient constructions. Clearly, from Prop.1 we know that one needs at least  $d \geq \sqrt{\dim \mathcal{V}}$ , but this will most likely not be achievable in general.

## SPECTRAL ANALYSIS AND SEPARATION FROM CLASSICAL DESCRIPTIONS

So far we investigated the dimensionality of the system based on the sequence  $\langle A(t) \rangle$ . The information obtained can be refined when going to the *frequency domain* by considering the function  $\mathcal{L} : \mathbb{C} \rightarrow \mathbb{C}$  (the ‘ $z$ -transform’ [8]) defined by the series

$$\mathcal{L}(z) := \frac{1}{z} \sum_{t \in \mathbb{N}_0} \frac{\langle A(t) \rangle}{z^t}. \quad (9)$$

This converges outside the unit circle and can be defined inside by analytic continuation. In this way we obtain for a sequence of the form (1)

$$\mathcal{L}(z) = \text{tr} [A(z \text{id} - T)^{-1}(\rho)], \quad (10)$$

so that poles of  $\mathcal{L}$  correspond to eigenvalues of  $T$ . Note that the latter lie always inside (or on) the unit circle, there is an eigenvalue 1, and complex eigenvalues come in conjugate pairs. While there are restrictions [11] for instance for the determinant, i.e., the product of eigenvalues, quantum mechanics does not impose any further constraint on the location of eigenvalues: any point on the unit disc is possible even for  $d = 2$ . The simplest example for this is a damped Rabi oscillation leading to  $\langle A(t) \rangle = e^{-\gamma t} \cos \omega t$  with poles of  $\mathcal{L}$  at  $e^{\pm i\omega - \gamma}$ .

It is instructive to compare this with a potential classical description of the sequence  $\langle A(t) \rangle$ . So assume there are  $d_c$  states to each of which we assign an initial probability  $p_k$ ,  $k = 1, \dots, d_c$ . The evolution of these probabilities for a single time-step is governed by a stochastic matrix  $S$  and in the end a measurement outcome  $a_k \in \mathbb{R}$  is assigned to the  $k$ ’th state. In this way we arrive at

$$\langle A(t) \rangle = \langle a | S^t | p \rangle. \quad (11)$$

Yet, the poles of  $\mathcal{L}$  correspond to eigenvalues of  $S$ , which share the basic properties mentioned above. However, the classical description imposes additional constraints on the location of the eigenvalues depending on the dimension  $d_c$ . In particular, they have to be located inside the convex hull of all roots of unity up to order  $d_c$ . That is, the unit disc will not be entirely covered for any finite  $d_c$ . A more complete characterization of the location of eigenvalues is given in [12, 13] and shown in Fig.1 for  $d_c = 2, 3, 4$ .

A simple consequence of this analysis is that in terms of the required degrees of freedom a quantum mechanical description of a sequence  $\langle A(t) \rangle$  can be far more efficient than a classical one. In the above discussion the separation between a quantum and a classical description comes from the simple fact that oscillations are easier to describe in terms of probability amplitudes than by using probabilities. Certainly more sophisticated separations can be found, however, a complete determination of  $d$  and  $d_c$  from a given sequence  $\langle A(t) \rangle$  seems to be a daunting task (despite considerable results on the classical side, cf. [14, 15]).

## FINITE AND NOISY DATA

So far we addressed the ideal case of a half-infinite and noiseless sequence—also noiseless in the sense that the expectation values are known precisely which requires infinite statistics. It is, however, straight forward to analyze finite and noisy data along the same lines. Let us begin with a finite sequence  $\langle A(t) \rangle$ ,  $t \leq 2(N-1)$  and consider the  $N \times N$  matrix  $V_{kl} := \langle A(k+l-2) \rangle$ . As the rows of  $V$  are a truncation of the vectors  $v_\tau$ ,  $\tau = 0, \dots, N-1$  we have

$$\dim \mathcal{V} \geq \text{rank} V, \quad (12)$$

with equality if  $N \geq \dim \mathcal{V}$ . If the data are noisy or suffering a significant statistical error, then  $V$  will typically be of full rank. However, if an error estimate is available we may consider an effective rank of  $V$  by disregarding all singular values below a certain noise threshold which is set by the estimated amount of errors. More precisely, assume that  $V$  is perturbed by some  $V_\epsilon$  (i.e., we actually observe  $V' = V + V_\epsilon$ ) where  $\|V_\epsilon\| \leq \epsilon$ . Then, by application of the singular value inequality [16], we get

$$\text{rank} V \geq \min\{k \mid s_{k+1}(V') \leq \epsilon\}, \quad (13)$$

where  $s_l(V')$  is the  $l$ 'th largest singular value of  $V'$ .

In Fig.2 the behavior of these singular values is depicted graphically for an example of a unitary evolution  $\text{tr}[AU^t \rho U^{\dagger t}]$  with  $d=3$  and randomly chosen  $\rho, A, U$ . As  $U$  preserves its eigenstates and has (due to the random choice) no other conserved quantities, we have  $\dim \mathcal{V} = 7$  which is well reflected in the singular values of  $V$  for small enough errors. In fact, in such unitary examples the dimension estimates appear to be surprisingly stable up to errors which make up a considerable fraction of the signal.

For the spectral analysis finite and noisy data seem to be more involved to handle. It is not difficult to show [10] that (9) is a rational function with exactly  $\dim \mathcal{V}$  poles and hence we are facing the problems of (i) giving a good (the best) rational approximation (with a fixed number of poles) of a function given by its Taylor coefficients and (ii) obtaining the poles of a rational function given by its (noisy) Taylor coefficients. The abstract solution to (i) is given by the Adamyan-Arov-Krein Theorem [17]. In practice (i.e., for finite data) a possible approach is the use of Padé approximants [18]. For (ii), one can use for instance the QD-algorithm of Rutishauer [19].

## DISCUSSION

In this letter, using tools from classical signal processing, we have discussed how to determine the Hilbert space dimension needed to explain observed data of an evolving quantum system. This introduces a new paradigm, beyond Bell inequalities and the analysis of correlations, to obtain such estimates with minimal assumptions in the model, in our case homogeneity and Markovianity. In particular, the method can

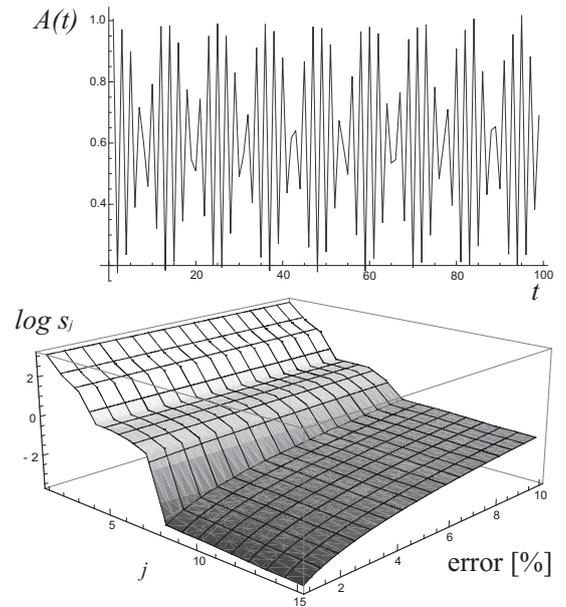


FIG. 2: *top*: a finite sequence  $\langle A(t) \rangle$ ,  $t \leq 100$  obtained from a randomly chosen unitary dynamics of a spin-1 quantum system ( $d=3$ ). As there are 3 conserved quantities (the eigenstates of the evolution) we have  $\dim \mathcal{V} \leq 7$ . *bottom*: log-plot of the 15 largest singular values  $s_j$  of the corresponding  $50 \times 50$  matrix  $V$  as a function of the standard deviation of added Gaussian noise (1%-10% of the signal). While for small noise the largest 7 singular values are clearly separated by a threshold, this washes out as the error increases—for too much noise the data could as well be explained by smaller  $\dim \mathcal{V}$  as expressed quantitatively by Eq.(13).

be used for single systems, which answers a question posed in [5] (for a different static approach, based on several input states, see the recent work [20]). We have also seen that we can use the method to quantify the non-Markovianity of an evolution via the effective dimension of the environments memory—complementing the static approach of [7]. Finally, the analysis of spectral information revealed a dramatic difference between the dimensions needed to give a quantum resp. classical explanation of the data. Clearly, this does not outlaw a classical description (as the Garg-Leggett inequalities [21] do, albeit based on an extra arguable assumption) but it shows that quantum mechanics can be much more efficient in terms of number of degrees of freedom. For the future, it would be nice to combine the two approaches, evolutions and correlations, to obtain more information.

This work has been funded by Spanish grants I-MATH and MTM2008-01366, by QUANTOP and the Ole Roemer grant of the Danish Natural Science Research Council (FNU). M.M.W. acknowledges discussion with M. Christandl and J.Appel.

## Appendix A

**Lemma 1 (Linear representation)** *For every bounded sequence  $\langle A(t) \rangle \in \mathbb{C}$ ,  $t \in \mathbb{N}_0$  there are vectors  $|L\rangle, |R\rangle$  of dimension  $\dim \mathcal{V}$  and a respective matrix  $M$  with  $\|M\|_\infty \leq 1$  such that*

$$\langle A(t) \rangle = \langle L|M^t|R\rangle. \quad (14)$$

*Proof.* Consider  $N \times N$  matrices  $V_{kl} := \langle A(k+l-2) \rangle$  and  $V'_{kl} := \langle A(k+l-1) \rangle$  with  $N$  large enough so that  $\text{rank} V = \text{rank} V' = \dim \mathcal{V} < N$ . Let  $V = V_L V_R$  be a singular value decomposition where  $V_R, V_L^T$  have  $\dim \mathcal{V}$  rows. Then  $|R\rangle := V_R|1\rangle$ ,  $\langle L| := \langle 1|V_L$  and (using pseudo-inverses)  $M := V_L^{-1} V' V_R^{-1}$  satisfy Eq.(14) obviously for  $t \leq 2N - 2$ . For larger values of  $t$ , however, the same relation has to hold by linear dependence, i.e., the fact that by construction  $\dim \mathcal{V} = \text{rank} V$ . boundedness of the sequence ( $|\langle A(t) \rangle| \leq \text{const.}$ ) implies that the spectral radius of  $M$  is at most 1 and that eigenvalues of magnitude one have one-dimensional Jordan blocks (i.e. their geometric multiplicity equals the algebraic multiplicity). To see this first note that powers of a Jordan block  $J(\lambda)$  are Toeplitz matrices with first row  $J(\lambda)_{1k}^t = \lambda^{t-k+1} \binom{t}{k-1}$ . Hence for  $|\lambda| = 1$   $\langle l|J(\lambda)^t|r\rangle$  is a polynomial in  $t$  which remains only bounded for growing  $t$  if it is zero.

Hence  $M$  can only have Jordan blocks for  $|\lambda| < 1$ . Let us finally show that  $M$  is therefore similar to a contraction: write  $J(\lambda) = \lambda \mathbb{1} + N$  with  $N$  a nilpotent matrix with ones on the first upper diagonal. Then  $J(\lambda)$  is similar to  $J' := \lambda \mathbb{1} + (1 - |\lambda|)N$  and using the triangle inequality for the norm  $\|J'\|_\infty \leq 1$ . As the representation of the sequence is preserved under similarity transformations of  $M$  (and respective transformations of  $R, L$ ) we can always w.l.o.g. chose  $M$  to be a contraction.  $\square$

- [10] F.R. Gantmacher, The Theory of Matrices I and II (Chelsea Publishing Company, 1959).
- [11] M.M. Wolf, J.I. Cirac, Commun. Math. Phys. **279**, 147 (2008).
- [12] F.I. Karpelevich, Izvestia Acad. Nauk SSSR Seria Mathem. **15**, 361 (1951); *Eleven Papers Translated from the Russian*, **140** American Mathematical Society (1988).
- [13] N. Dimitriev, E. Dynkin, Izvestia Acad. Nauk SSSR, Seria Mathem. **10**, 167 (1946); *Eleven Papers Translated from the Russian*, **140** American Mathematical Society (1988) [see Fig. 11 p.74].
- [14] A. Heller, Ann. Math. Stat. **36**, 1286 (1965).
- [15] Y. Dai, Acta Math. Sinica **10**, 99 (1994).
- [16] Problem III.6.4. in R.Bhatia, *Matrix Analysis*, Springer 1997.
- [17] V.V. Peller, *Hankel Operators and their Applications*, Springer 2003.
- [18] G.A. Baker, P.R. Graves-Morris, *Pade approximants*, Cambridge University Press, 1995.
- [19] P. Herici, *Applied and Computational Complex Analysis*, Wiley, 1974.
- [20] S. Wehner, M. Christandl, A. C. Doherty, arXiv:0808.3960. To appear in Phys. Rev. A.
- [21] A. J. Leggett and A. Garg, Phys. Rev. Lett. **54**, 857 (1985).

- 
- [1] K. Hammerer, A.S. Sorensen, E.S. Polzik, arXiv:0807.3358 (2008).
  - [2] P. Rabl, D. DeMille, J.M. Doyle, M.D. Lukin, R.J. Schoelkopf, P. Zoller, Phys. Rev. Lett. **97**, 033003 (2006).
  - [3] A. Acin, N. Gisin, L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006);
  - [4] D. Perez-Garcia, M.M Wolf, C. Palazuelos, I. Villanueva, M. Junge, Comm. Math. Phys. **279**, 455 (2008).
  - [5] N. Brunner, S. Pironio, A. Acin, N. Gisin, A.A. Methot, V. Scarani, Phys. Rev. Lett. **100**, 210503 (2008).
  - [6] T. Vértesi, K.F. Pál, arXiv:0812.1572; Jop Briët, Harry Buhrman, Ben Toner, arXiv:0901.2009.
  - [7] M.M. Wolf, J. Eisert, T.S. Cubitt, J.I. Cirac, Phys. Rev. Lett. **101**, 150402 (2008).
  - [8] D. Sundararajan, *Digital Signal Processing*, World Scientific, 2003.
  - [9] F. Takens, Detecting strange attractors in turbulence, in *Lecture notes in Mathematics*, No.898 (Springer-Verlag, 1981).