Hypersensitivity and chaos signatures in the quantum baker's maps

2006 J. Phys. A: Math. Gen. 39 13405

(http://iopscience.iop.org/0305-4470/39/43/002)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 13405–13433 doi:10.1088/0305-4470/39/43/002

Hypersensitivity and chaos signatures in the quantum baker's maps

A J Scott[1,2], Todd A Brun[3], Carlton M Caves[1] and Rüdiger Schack[4]

[1] Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131-1156, USA
[2] Institute for Quantum Information Science, University of Calgary, Calgary, AB T2N 1N4, Canada
[3] Communication Sciences Institute, University of Southern California, Los Angeles, CA 90089-2565, USA
[4] Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK

E-mail: ascott@qis.ucalgary.ca, tbrun@usc.edu, caves@info.phys.unm.edu and r.schack@rhul.ac.uk

Received 17 June 2006, in final form 22 August 2006
Published 11 October 2006
Online at stacks.iop.org/JPhysA/39/13405

Abstract

Classical chaotic systems are distinguished by their sensitive dependence on initial conditions. The absence of this property in quantum systems has led to a number of proposals for perturbation-based characterizations of quantum chaos, including linear growth of entropy, exponential decay of fidelity, and hypersensitivity to perturbation. All of these accurately predict chaos in the classical limit, but it is not clear that they behave the same far from the classical realm. We investigate the dynamics of a family of quantizations of the baker's map, which range from a highly entangling unitary transformation to an essentially trivial shift map. Linear entropy growth and fidelity decay are exhibited by this entire family of maps, but hypersensitivity distinguishes between the simple dynamics of the trivial shift map and the more complicated dynamics of the other quantizations. This conclusion is supported by an analytical argument for short times and numerical evidence at later times.

PACS number: 05.45.Mt

1. Introduction

A full characterization of quantum chaos is an elusive matter. Classical chaotic systems are distinguished by their exponential sensitivity to initial conditions. Quantified in terms of Lyapunov exponents, this characterization is the key ingredient in any definition of classical chaos. The linearity of quantum mechanics prohibits such sensitivity to initial conditions,

thus obstructing any straightforward extension of the classical definition of chaos to quantum systems. The standard fix is to categorize as 'chaotic' those quantum systems that are chaotic in a classical limit. These systems are not strictly chaotic by the classical definition—they are quasiperiodic—but they have properties, involving the spectrum of energy eigenvalues and the behaviour of energy eigenstates, that are distinctly different from that for quantizations of classically regular systems.

So far, however, little agreement has been reached on a characterization of quantum chaos that does not make reference to a classical limit. The classical approach of looking at small perturbations of the initial state fails due to the unitarity of quantum dynamics. Attempts at a direct dynamical characterization of quantum chaos, which applies even in the hard quantum regime, thus generally look at the effects of small perturbations of the dynamics. A number of perturbation-based criteria (or signatures) have been proposed, including *linear growth of entropy* in the presence of stochastic perturbations [1, 2]; *exponential decay of fidelity* between states that evolve under two close, but distinct unitary transformations [3, 4]; and *hypersensitivity to perturbation*, which considers stochastic perturbations and compares the amount of information known about the perturbation to the resulting reduction in the system entropy [5–8], a relation called the *information–entropy trade-off*.

Typically, a quantum system whose classical limit is chaotic exhibits all three of these signatures. They are clearly inequivalent, however. For instance, hypersensitivity to perturbation can be viewed as a measure of how fast, how widely and how randomly the set of all possible perturbed states spreads through Hilbert space. The other two criteria, though they report on how widely the perturbed vectors are dispersed through Hilbert space, are not sensitive to the exact way in which Hilbert space is explored by the perturbed dynamics. It is therefore conceivable that there are quantum systems that are chaotic with respect to one criterion, but regular with respect to another. In this paper we compare the three perturbation-based criteria for a family of quantizations of the quantum baker's map.

The classical baker's map [9] is a well-known toy mapping whose study has led to many insights in the field of classical chaos by demonstrating essential features of nonlinear dynamics. It maps the unit square, which can be thought of as a toroidal phase space, onto itself in an area-preserving way. Interest in the baker's map stems from its straightforward formulation in terms of a Bernoulli shift on binary sequences. It seems natural to consider a quantum version of the baker's map for the investigation of quantum chaos. There is, however, no unique procedure for quantizing a classical map; hence, different quantum maps correspond to the same classical baker's transformation in the classical limit. The family of quantizations [10] used in the present paper is based on the 2^N-dimensional Hilbert space of N qubits. This qubit structure provides a connection to the binary representation of the classical baker's map.

The paper is organized as follows. In section 2 we introduce our mathematical notation and give precise definitions of the three perturbation-based chaos criteria. Section 3 reviews a family of quantizations for the quantum baker's map. These range from a highly entangling unitary transformation to an essentially trivial shift map. In section 4 we give simple analytical results concerning the three criteria for the trivial shift map; these results show that the trivial shift map exhibits linear entropy growth and exponential fidelity decay and suggest that it does not display hypersensitivity to perturbation. Section 5 presents numerical calculations for the entire family of baker's maps. These calculations show that, unlike the other two criteria, hypersensitivity to perturbation differentiates between the different quantizations. In appendix A, we formulate a simple model for the form of the information–entropy trade-off in the case of vectors distributed randomly on Hilbert space. The model serves as a foil for interpreting the results of our numerical work on hypersensitivity. Appendix B derives the von Neumann entropy of an ensemble of vectors that populate half of Hilbert space uniformly;

this result is used to bound the information–entropy trade-off in the case that the amount of information about the perturbation is 1 bit. Finally, in section 6 we discuss our results.

2. Criteria for quantum chaos

2.1. Hypersensitivity to perturbation

2.1.1. Definition of hypersensitivity. In the most general setting, hypersensitivity to perturbation can be defined as follows [11]. Consider a system with Hilbert space $\mathcal{S}$, evolving under some unitary evolution and, in addition, interacting with an environment with Hilbert space $\mathcal{E}$. Let D and $D_{\mathcal{E}}$ denote the Hilbert-space dimensions of the system and environment, respectively. Initially, the joint state of the system and environment is assumed to be a product state; i.e., initially there is no correlation. After a time t, the joint state of the system and environment is a density operator on $\mathcal{S}\otimes\mathcal{E}$, which we denote by $\hat{\rho}_{\text{total}}$. The state of the system at time t, $\hat{\rho}$, is obtained by tracing out the environment,

$$\hat{\rho} = \text{tr}_{\mathcal{E}}(\hat{\rho}_{\text{total}}). \tag{2.1}$$

The von Neumann entropy of the system at time t is

$$H_{\mathcal{S}} = -\text{tr}(\hat{\rho}\log\hat{\rho}). \tag{2.2}$$

We measure entropy in bits (i.e. we take $\log \equiv \log_2$).

Now assume that an arbitrary measurement is performed on the environment. The most general measurement [12] is described by a positive-operator-valued measure (POVM), $\{\hat{E}_r\}$, where $\hat{E}_r$ are positive operators acting on the environment and satisfying the completeness condition

$$\sum_r \hat{E}_r = \hat{\mathbb{1}}_{\mathcal{E}} = \text{(environment identity operator)}. \tag{2.3}$$

The probability of obtaining the measurement outcome r is given by

$$p_r = \text{tr}(\hat{\rho}_{\text{total}}(\hat{\mathbb{1}}_{\mathcal{S}}\otimes\hat{E}_r)), \tag{2.4}$$

where $\hat{\mathbb{1}}_{\mathcal{S}}$ is the identity operator on the system. The system state after a measurement that yields the outcome r is

$$\hat{\rho}_r = \frac{\text{tr}_{\mathcal{E}}(\hat{\rho}_{\text{total}}(\hat{\mathbb{1}}_{\mathcal{S}}\otimes\hat{E}_r))}{p_r}. \tag{2.5}$$

We define the system entropy conditional on the outcome r,

$$H_r = -\text{tr}(\hat{\rho}_r\log\hat{\rho}_r), \tag{2.6}$$

the average conditional entropy,

$$\bar{H} = \sum_r p_r H_r, \tag{2.7}$$

and the average entropy decrease due to the measurement, $\Delta\bar{H} = H_{\mathcal{S}} - \bar{H}$. Furthermore, we define the *average information*,

$$\bar{I} = -\sum_r p_r\log p_r. \tag{2.8}$$

The quantity $\bar{I}$ is within 1 bit of the minimum average algorithmic information needed to specify the measurement outcome r [13].

Now assume we want to perform a measurement that reduces the average conditional system entropy below some given target value, H. We define the quantity

$$I_{\min}(H) = \inf \bar{I}, \tag{2.9}$$

where the infimum is taken over all POVMs $\{\hat{E}_r\}$ such that $\bar{H} \leqslant H$. The function $I_{\min}(H)$ expresses what we call the *information–entropy trade-off*; it can be interpreted as the minimum information about the perturbing environment needed to keep the average system entropy below the target value H. We say the system is *hypersensitive to perturbation* if this information is large compared to the purchased entropy reduction $\Delta H = H_S - H$, i.e.,

$$\frac{I_{\min}(H)}{\Delta H} \gg 1, \tag{2.10}$$

in the region of small enough entropy reductions that this ratio reports on the system dynamics rather than on the multiplicity of possible perturbations. We characterize this region more precisely in the next subsection.

For the analysis of the present paper, we specialize to the case where initially the system is in a pure state, $|\psi_0\rangle$, and the unitary system evolution is given by a quantum map $\hat{B}$. In the absence of any interaction with the environment, the system state after t iterations (or time steps) is $\hat{B}^t|\psi_0\rangle$. In addition, we assume that the effect of the environment is equivalent to a stochastic perturbation. At each time step, a perturbation is chosen randomly from a set of unitary maps, $\{U_k : k \in \mathcal{K}\}$, where $\mathcal{K}$ is some index set.

The joint system–environment density operator after t iterations is then given by

$$\hat{\rho}_{\text{total}} = \sum_{\mathbf{k}\in\mathcal{K}^t} p_{\mathbf{k}} |\psi_{\mathbf{k}}\rangle\langle\psi_{\mathbf{k}}| \otimes \hat{P}^{\mathcal{E}}_{\mathbf{k}}, \tag{2.11}$$

where

$$|\psi_{\mathbf{k}}\rangle = \hat{U}_{k_t}\hat{B}\hat{U}_{k_{t-1}}\hat{B}\cdots\hat{U}_{k_1}\hat{B}|\psi_0\rangle \tag{2.12}$$

is the endpoint of a stochastic trajectory labelled by $\mathbf{k} = (k_1, \ldots, k_t)$, $p_{\mathbf{k}}$ is the probability of $\mathbf{k}$, and the operators $\hat{P}^{\mathcal{E}}_{\mathbf{k}}$ are one-dimensional, orthogonal environment projectors. The perturbation histories $\mathbf{k} = (k_1, \ldots, k_t)$ are thus recorded in the environment in the form of the orthogonal projectors $\hat{P}^{\mathcal{E}}_{\mathbf{k}}$. The reduced density operator for the system is given by

$$\hat{\rho} = \text{tr}_{\mathcal{E}}(\hat{\rho}_{\text{total}}) = \sum_{\mathbf{k}\in\mathcal{K}^t} p_{\mathbf{k}} |\psi_{\mathbf{k}}\rangle\langle\psi_{\mathbf{k}}|. \tag{2.13}$$

In the numerical analysis in section 5, we let the environment be a series of qubits, which interact sequentially with the system; hence the perturbation at each time step is a binary perturbation, $\mathcal{K} = \{0, 1\}$.

It is in general very difficult to determine the function $I_{\min}(H)$ because it is generally impossible to do the required optimization over all POVMs. In the numerical results reported in section 5, we restrict the optimization to POVMs of the form

$$\hat{E}_r = \sum_{\mathbf{k}\in K_r} \hat{P}^{\mathcal{E}}_{\mathbf{k}}, \tag{2.14}$$

where the subsets $K_r \subset \mathcal{K}^t$ are nonoverlapping subsets of the perturbation histories. Such POVMs can be regarded as sampling a coarse-grained version of the perturbation histories. For the special form of $\hat{\rho}_{\text{total}}$ considered here, it seems reasonable that ensembles which are optimal with respect to this class of measurements are also, to a good approximation, optimal with respect to the class of all possible environment measurements. We have not, however, been able to prove this statement rigorously.

The measurements (2.14) correspond to forming groups of system vectors $|\psi_{\mathbf{k}}\rangle$. Assuming that all perturbations are equally likely, i.e., $p_{\mathbf{k}} = 1/\mathcal{N}$ for all $\mathbf{k}$, where $\mathcal{N}$ is the number of perturbation histories or system vectors $|\psi_{\mathbf{k}}\rangle$, the probability of obtaining outcome r is given by

$$p_r = |K_r|/\mathcal{N}, \tag{2.15}$$

and the system state after a measurement that yields outcome r is the average of the grouped vectors,

$$\hat{\rho}_r = \frac{\mathrm{tr}_{\mathcal{E}}(\hat{\rho}_{\mathrm{total}}(\hat{\mathbb{1}}_{\mathcal{S}} \otimes \hat{E}_r))}{p_r} = \frac{1}{|K_r|} \sum_{\mathbf{k} \in K_r} |\psi_{\mathbf{k}}\rangle\langle\psi_{\mathbf{k}}|. \tag{2.16}$$

We use this simplified framework in the discussion in sections 4 and 5 below.

Even this simplified framework is not sufficient to make the problem tractable for numerical purposes, because the number of vectors, $\mathcal{N}$, increases exponentially with the number of time steps, rapidly making it impossible to search over all possible ways of grouping the vectors $|\psi_{\mathbf{k}}\rangle$. To get around this, we employ efficient algorithms for grouping the vectors, which are plausibly able to find optimal or near-optimal groupings. For the numerical results reported in section 5, we first use a particularly simple, but intuitive grouping algorithm devised to clarify the procedure and then take a general approach based on genetic algorithms. We have compared the results obtained using these algorithms with those obtained using other grouping algorithms, some of which have been used previously [7], and found that the current algorithms are generally superior for the vectors generated by the perturbed baker's map.

2.1.2. Quantitative measure of hypersensitivity. Hypersensitivity to perturbation tests how fast and how fully the state of the perturbed system explores the system Hilbert space. To see how this is quantified by the information–entropy trade-off, we consider the trade-off relation for vectors that are distributed randomly in Hilbert space. Such a relation was formulated in [7, 11], using a model that groups the random vectors into spheres of uniform radius (measured by Hilbert-space angle) on projective Hilbert space [14]. We refine this model and its trade-off relation in appendix A. The main result is that for $\mathcal{N}$ vectors distributed randomly in d Hilbert-space dimensions, the information–entropy trade-off, written in inverse form, is approximated by

$$H = \begin{cases} \log\mathcal{N} - I_{\min}, & \log\mathcal{N} \geqslant I_{\min} \geqslant \log\mathcal{N} - \log d, \\ \log d - \dfrac{1}{d}((1 + I_{\min}\ln 2)\log(1 + I_{\min}\ln 2) - I_{\min}), & \log\mathcal{N} - \log d \geqslant I_{\min} \geqslant 1. \end{cases} \tag{2.17}$$

This expression assumes that d is large and that the number of random vectors, though large in the sense that $\mathcal{N} \gg d$, satisfies $\mathcal{N} \ll 2^d$, a situation we refer to as a *sparse* collection of vectors. Examples of the information–entropy trade-off are shown in figure 3 (section 5), in which the upper solid curve closely resembles the exact situation for random vectors in 32 dimensions; note that $I_{\min}$ follows H in a linear fashion before dropping quickly at a 'knee' close to the maximum entropy. This agrees with our approximation (2.17) for random vectors, which is shown as the upper dotted line and the rightmost dotted curve in figure 3.

Several features of the sphere-grouping trade-off (2.17) deserve discussion. The number of spheres, $2^{I_{\min}}$, gives the number of vectors per group $\mathcal{N}_V = \mathcal{N}2^{-I_{\min}}$. The knee at $I_{\min} = \log\mathcal{N} - \log d$ thus corresponds to $\mathcal{N}_V = d$. For $I_{\min} > \log\mathcal{N} - \log d$, i.e., $\mathcal{N}_V < d$, the number of vectors in each sphere is insufficient to explore all the Hilbert-space dimensions. This gives a linear dependence on $I_{\min}$, with slope -1 and intercept $\log\mathcal{N}$. In the context of a

stochastically perturbed map, where $\mathcal{N}$ is the number of perturbation histories, this part of the trade-off relation tells us about the multiplicity of the perturbation instead of about the dynamics of the map. In contrast, for $I_{\min} < \log \mathcal{N} - \log d$, i.e., $\mathcal{N}_V > d$, where the number of vectors in each sphere is large enough to explore all Hilbert-space dimensions, the information–entropy trade-off becomes independent of $\mathcal{N}$. *It is this part of the trade-off relation, beyond the knee in the information–entropy trade-off, that tells us about hypersensitivity to perturbation in the system dynamics*. Note that we need $\mathcal{N} \gg d$ to investigate this region, but we do *not* need $\mathcal{N}$ so large that a random collection of vectors would sample generic vectors, which requires at least $\mathcal{N} \sim 2^d$ vectors, i.e., what we call a dense collection. Our stochastic perturbation need only produce a sparse collection of vectors to see evidence of hypersensitivity; we can say that the vectors in such a sparse collection are pseudo-random instead of random [15].

Projective Hilbert space can never be partitioned exactly into spheres of uniform radius. This has little effect when the spheres are tiny and numerous, but it becomes a problem when there are just a few spheres and prompts us to treat the sphere-grouping trade-off relation with caution in this situation. In particular, for the case of just two groups, i.e., $I_{\min} = 1$, a better method for grouping random vectors is to partition projective Hilbert space into two equal volumes defined by the closeness to two orthogonal subspaces of dimension $d/2$. The resulting trade-off for $I_{\min} = 1$ is analysed in appendix B and summarized below. We abandon the trade-off relation (2.17) entirely for $I_{\min} < 1$ because a grouping into spheres of uniform radius makes no sense when there are fewer than two spheres.

We are now in a position to introduce a quantitative measure of a map's hypersensitivity to perturbation. For this purpose we introduce the quantity

$$s \equiv \frac{1}{H_{\mathcal{S}} - H(I_{\min} = 1)}; \tag{2.18}$$

$1/s$ is the reduction in system entropy purchased by gathering one (optimal) bit of information about the environment. In a Hilbert-space context, s is an indicator of the randomness in a collection of vectors. It is a considerably more informative indicator than the entropy. For example, the members of an orthonormal basis together achieve maximal entropy, yet a grouping of these vectors into two equally sized groups gives $s = 1$, independent of the dimension. In contrast, for vectors distributed randomly across Hilbert space, the sphere-grouping trade-off relation (2.17) gives

$$s = \frac{1}{\log d - H(I_{\min} = 1)} = \frac{d}{(1 + \ln 2)\log(1 + \ln 2) - 1} \approx 3.5\, d. \tag{2.19}$$

While s provides a signature of randomness, it is the change in s with time, as applied to the perturbed system vectors, which indicates the degree to which a system is hypersensitive to perturbation. A rapid increase in s over time has been proposed as a criterion of chaos for both classical and quantum systems [5]. In particular, if s increases exponentially with time, we say that the system exhibits *exponential hypersensitivity to perturbation* [6, 8].

A detailed analysis [6] of stochastic perturbations of classically chaotic maps described by a symbolic dynamics shows that for such systems, s is indeed a measure of the phase-space stretching and folding characteristic of chaotic dynamics. Specifically, s grows as 2^{Kt}, where K is the Kolmogorov–Sinai entropy of the dynamics [16], showing that these systems do display exponential hypersensitivity to perturbation and that exponential hypersensitivity is equivalent to the standard characterization of classical chaos via the Kolmogorov–Sinai entropy, which in turn is equivalent to characterization in terms of Lyapunov exponents.

For quantum systems, s is a measure of how fast and how fully the state of a perturbed system explores the system Hilbert space. An exponential increase in s indicates both that the number of dimensions, d, explored by the perturbed vectors grows exponentially and that

the vectors populate the explored dimensions randomly. Thus s provides a direct dynamical characterization of quantum chaotic dynamics, a characterization that is analogous to the characterization of classical chaos in terms of sensitivity to initial conditions. The reason that hypersensitivity to perturbation goes beyond the Zurek–Paz chaos criterion of linear entropy increase under stochastic perturbations [1, 2] is clear: a linear entropy increase indicates that the perturbed vectors explore an exponentially increasing number of dimensions, but is silent on whether those dimensions are explored randomly.

A related parameter for characterizing hypersensitivity is the slope of the information–entropy trade-off, $|\mathrm{d}I_{\min}/\mathrm{d}H|$, evaluated at $I_{\min} = 0$ (i.e., $H = H_S$) or perhaps at $I_{\min} = 1$. Both the classical analysis in [6] and the analysis of appendix A prompt us to shy away from using the slope evaluated at $I_{\min} = 0$, since there are uncertainties about the behaviour of the slope for very small values of $I_{\min}$. Moreover, the slope evaluated at $I_{\min} = 1$ seems to have no advantages over the parameter s. Thus, in this paper, we calculate numerically information–entropy trade-offs for the perturbed quantum baker's maps, and from these we determine the time evolution of the hypersensitivity parameter s, preferring it to the more problematic use of the slope.

Having settled on s as our signature of hypersensitivity, we can formulate a better information–entropy trade-off for random vectors when $I_{\min} = 1$, i.e., for the case of two groups. An optimal way of grouping a sufficiently dense collection of random vectors, analysed in appendix B, is then the following: choose two orthogonal subspaces, each of dimension $d/2$, and partition projective Hilbert space into two equal volumes defined by the distance in Hilbert-space angle to these subspaces. The entropy of each partition is

$$H = \frac{d}{2}(-\lambda_+ \log \lambda_+ - \lambda_- \log \lambda_-) \approx \log d - \frac{1}{\pi d \ln 2} \tag{2.20}$$

(cf equations (B.22) and (B.23) with $n = d/2$), where

$$\lambda_\pm = \frac{1}{d}\left(1 \pm \frac{d!}{2^d[(d/2)!]^2}\right) \approx \frac{1}{d}\left(1 \pm \sqrt{\frac{2}{\pi d}}\right) \tag{2.21}$$

(cf equations (B.20) and (B.21) with $n = d/2$). The approximate expressions on the right-hand side hold for large d and give

$$s = \frac{1}{\log d - H} = \pi d \ln 2 \approx 2.2\, d. \tag{2.22}$$

The coefficient 2.2, smaller than the 3.5 of equation (2.19), indicates that this is a better way to partition random vectors into two groups. This value of s represents an approximate upper bound for any collection of vectors in Hilbert space.

Another scenario that is important for the current study occurs when the perturbed vectors are restricted to product states of N qubits. Random product vectors can be grouped into the two groups corresponding to $I_{\min} = 1$ by partitioning the projective Hilbert space of one of the qubits into two equal volumes, just as above. The entropy of each partition for this qubit is $2 - (3/4)\log 3$ (equation (2.20) with $d = 2$). Thus for all the qubits, the entropy of each partition is

$$H = N - \tfrac{3}{4}\log 3 + 1, \tag{2.23}$$

which gives

$$s = \frac{1}{N - H} = \frac{4}{3\log 3 - 4} \approx 5.3, \tag{2.24}$$

independent of $D = 2^N$. This value represents a rather restrictive approximate upper bound on s for product vectors.

Suppose that for random product vectors, we partition the projective Hilbert spaces of j constituent qubits into two equal volumes, thus using $j = I_{\min}$ bits of information to purchase a reduction of the entropy to

$$H(I_{\min}) = N - \left(\tfrac{3}{4}\log 3 - 1\right) I_{\min} \approx N - I_{\min}/5.3, \tag{2.25}$$

for $N \geqslant I_{\min} \geqslant 0$. This information–entropy trade-off, which, unlike equation (2.17), is linear near the maximal entropy, is plotted as the lower dotted line in figure 3 (section 5). It shows that nonentangling quantum maps are not hypersensitive to perturbation.

2.2. Other perturbation-based criteria for quantum chaos

Fidelity decay as a criterion for quantum chaos was introduced by Peres [3, 4] (see also [17–19] and references therein). One compares the unitary evolution of an initial state $|\psi_0\rangle$ under the action of a quantum map $\hat{B}$ with the evolution of the same initial state under the action of a modified map, $\hat{B}' = \hat{U}\hat{B}$, where the unitary map $\hat{U}$ is close to the identity operator. According to this criterion, a quantum map is chaotic if the fidelity,

$$F(t) = |\langle\psi_0|(\hat{B}^\dagger\hat{U}^\dagger)^t\hat{B}^t|\psi_0\rangle|^2, \tag{2.26}$$

decreases exponentially with the number of iterations at short times. In contrast to the criterion of hypersensitivity to perturbation, where the effects of a stochastic perturbation are analysed, fidelity decay focuses on just two perturbation histories, corresponding to the unperturbed evolution and to a modified evolution where the same perturbation operator $\hat{U}$ is applied at each time step.

Linear entropy increase as a chaos criterion was introduced by Zurek and Paz [1, 2]. According to this criterion, a quantum map is chaotic if the entropy (2.2) of the reduced system density operator (2.1) increases linearly with the number of iterations at short times. As we have already discussed, a linear entropy increase is essential for exponential hypersensitivity to perturbation, but it is not the whole story.

3. Quantum baker's maps

The classical baker's map is a standard example of chaotic dynamics [9]. It is a symplectic map of the unit square onto itself, defined through the equations

$$q_{n+1} = 2q_n - \lfloor 2q_n \rfloor, \tag{3.1}$$

$$p_{n+1} = (p_n + \lfloor 2q_n \rfloor)/2, \tag{3.2}$$

where $q, p \in [0, 1)$, $\lfloor x \rfloor$ is the integer part of x, and n denotes the nth iteration of the map. Geometrically, the map stretches the unit square by a factor of 2 in the q direction, squeezes by a factor of a half in the p direction, and then stacks the right half onto the left.

Interest in the baker's map stems from its straightforward symbolic-dynamical characterization in terms of a Bernoulli shift on binary sequences. If each point of the unit square is identified through its binary representation, $q = 0 \cdot s_1 s_2 \cdots = \sum_{k=1}^{\infty} s_k 2^{-k}$ and $p = 0 \cdot s_0 s_{-1} \ldots = \sum_{k=0}^{\infty} s_{-k} 2^{-k-1}$ ($s_i \in \{0, 1\}$), with a bi-infinite symbolic string

$$s = \underbrace{\ldots s_{-2} s_{-1} s_0}_{p} \bullet \underbrace{s_1 s_2 s_3 \ldots}_{q}, \tag{3.3}$$

then the action of the baker's map is to shift the position of the dot by one digit to the right,

$$s \to s' = \underbrace{\ldots s_{-2} s_{-1} s_0 s_1}_{p'} \bullet \underbrace{s_2 s_3 \ldots}_{q'}. \tag{3.4}$$

It seems natural to consider a quantum version of the baker's map for the investigation of quantum chaos. There is, however, no unique procedure for quantizing a classical map: different quantum maps can lead to the same classical baker's transformation.

To construct a quantum baker's map, we work in a D-dimensional Hilbert space, $\mathcal{H}_D$, spanned by either the position states $|q_j\rangle$, with eigenvalues $q_j = (j+1/2)/D$, or the momentum states $|p_k\rangle$, with eigenvalues $p_k = (k + 1/2)/D$ $(j, k = 0, \ldots, D - 1)$. The constants of $1/2$ determine the type of periodicity assumed for the position and momentum states, in this case, $|q_{j+D}\rangle = -|q_j\rangle, |p_{k+D}\rangle = -|p_k\rangle$, and thus identify $\mathcal{H}_D$ with a toroidal phase space with antiperiodic boundary conditions. The vectors of each basis are orthonormal, $\langle q_j|q_k\rangle = \langle p_j|p_k\rangle = \delta_{jk}$, and the two bases are related via the discrete Fourier transform $\hat{F}_D$,

$$\langle q_j|\hat{F}_D|q_k\rangle \equiv \langle q_j|p_k\rangle = \frac{1}{\sqrt{D}}\,\mathrm{e}^{\mathrm{i}q_j p_k/\hbar}. \tag{3.5}$$

For consistency of units, we must have $2\pi\hbar D = 1$.

The first work on a quantum baker's map was done by Balazs and Voros [20]. Assuming an even-dimensional Hilbert space with periodic boundary conditions, they defined a quantum baker's map in terms of a unitary operator $\hat{B}$ that executes a single iteration of the map. Saraceno [21] later improved certain symmetry characteristics of this quantum baker's map by using antiperiodic boundary conditions as described above. To define the Balazs–Voros–Saraceno unitary operator in our notation, imagine that the even-dimensional Hilbert space is a tensor product of a qubit space and the space of a $(D/2)$-dimensional system. Writing $j = x(D/2) + j', x \in \{0, 1\}$, we can write the position eigenstates as $|q_j\rangle = |x\rangle \otimes |j'\rangle$, where the states $|x\rangle$ make up the standard basis for the qubit, and the states $|j'\rangle$ are a basis for the $(D/2)$-dimensional system. The state of the qubit thus determines whether the position eigenstate lies in the left or right half of the unit square. The Balazs–Voros–Saraceno quantum baker's map is defined by

$$\hat{B} = \hat{F}_D \circ \left(\hat{\mathbb{1}}_2 \otimes \hat{F}_{D/2}^{-1}\right), \tag{3.6}$$

where $\hat{\mathbb{1}}_2$ is the unit operator for the qubit, and $\hat{F}_{D/2}$ is the discrete Fourier transform on the $(D/2)$-dimensional system. The unitary $\hat{B}$ does separate inverse Fourier transforms on the left and right halves of the unit square, followed by a full Fourier transform.

For dimensions $D = 2^N$, an entire class of quantum baker's maps can be defined in analogy with the symbolic dynamics for the classical baker's map [10]. In this case, we can model our Hilbert space as the tensor-product space of N qubits, and the position states can be defined as product states for the qubits in the standard basis, i.e.,

$$|q_j\rangle = |x_1\rangle \otimes |x_2\rangle \otimes \cdots \otimes |x_N\rangle, \tag{3.7}$$

where j has the binary expansion

$$j = x_1 \ldots x_N \cdot 0 = \sum_{l=1}^{N} x_l 2^{N-l} \tag{3.8}$$

and $q_j = (j + 1/2)/D = 0 \cdot x_1 \ldots x_N 1$.

To make the connection with the symbolic dynamics for the classical baker's map, we proceed as follows. The bi-infinite strings (3.3) that specify points in the unit square are replaced by sets of orthogonal quantum states created through the use of a partial Fourier transform

$$\hat{G}_n \equiv \hat{\mathbb{1}}_{2^n} \otimes \hat{F}_{2^{N-n}}, \qquad n = 0, \ldots, N, \tag{3.9}$$

where $\hat{\mathbb{1}}_{2^n}$ is the unit operator on the first n qubits and $\hat{F}_{2^{N-n}}$ is the Fourier transform on the remaining qubits. The partial Fourier transform thus transforms the $N-n$ least significant qubits of a position state,

$$\hat{G}_n|x_1\rangle\otimes\cdots\otimes|x_n\rangle\otimes|a_1\rangle\otimes\cdots\otimes|a_{N-n}\rangle = |x_1\rangle\otimes\cdots\otimes|x_n\rangle\otimes\frac{1}{\sqrt{2^{N-n}}} \times\sum_{x_{n+1},\ldots,x_N}|x_{n+1}\rangle\otimes\cdots\otimes|x_N\rangle\,\mathrm{e}^{2\pi\mathrm{i}ax/2^{N-n}}, \tag{3.10}$$

where a and x are defined through the binary representations $a = a_1\ldots a_{N-n}\cdot 1$ and $x = x_{n+1}\ldots x_N\cdot 1$. In the limiting cases, we have $\hat{G}_0 = \hat{F}_D$ and $\hat{G}_N = i\,\hat{\mathbb{1}}$. The analogy to the classical case is made clear by introducing the following notation for the partially transformed states:

$$|a_{N-n}\ldots a_1\bullet x_1\ldots x_n\rangle \equiv \hat{G}_n|x_1\rangle\otimes\cdots\otimes|x_n\rangle\otimes|a_1\rangle\otimes\cdots\otimes|a_{N-n}\rangle. \tag{3.11}$$

For each value of n, these states form an orthonormal basis and are localized in both position and momentum. The state $|a_{N-n}\ldots a_1\bullet x_1\ldots x_n\rangle$ is strictly localized in a position region of width $1/2^n$ centred at $0\cdot x_1\ldots x_n1$ and is roughly localized in a momentum region of width $1/2^{N-n}$ centred at $0\cdot a_1\ldots a_{N-n}1$. In the notation of equation (3.3), it is localized at the phase-space point $1a_{N-n}\ldots a_1\bullet x_1\ldots x_n1$. Note that $|a_N\ldots a_1\bullet\rangle = \hat{G}_0|a_1\rangle\otimes\cdots\otimes|a_N\rangle$ is a momentum eigenstate and that $|\bullet\, x_1\ldots x_N\rangle = \hat{G}_N|x_1\rangle\otimes\cdots\otimes|x_N\rangle = i|x_1\rangle\otimes\cdots\otimes|x_N\rangle$ is a position eigenstate, the factor of i being a consequence of the antiperiodic boundary conditions.

Using this notation, a quantum baker's map on N qubits is defined for each value of $n = 1,\ldots,N$ by the single-iteration unitary operator [10]

$$\hat{B}_{N,n} \equiv \hat{G}_{n-1}\circ\hat{S}_n\circ\hat{G}_n^{-1} = \sum_{x_1,\ldots,x_n}\sum_{a_1,\ldots,a_{N-n}}|a_{N-n}\ldots a_1x_1\bullet x_2\ldots x_n\rangle\langle a_{N-n}\ldots a_1\bullet x_1x_2\ldots x_n|, \tag{3.12}$$

where the shift operator $\hat{S}_n$ acts only on the first n qubits, i.e., $\hat{S}_n|x_1\rangle\otimes|x_2\rangle\otimes\cdots\otimes|x_n\rangle\otimes|x_{n+1}\rangle\otimes\cdots\otimes|x_N\rangle = |x_2\rangle\otimes\cdots\otimes|x_n\rangle\otimes|x_1\rangle\otimes|x_{n+1}\rangle\otimes\cdots\otimes|x_N\rangle$. Note that since $\hat{S}_n$ commutes with $\hat{G}_n^{-1}$, we can put $\hat{B}_{N,n}$ in the form

$$\hat{B}_{N,n} = \hat{\mathbb{1}}_{2^{n-1}}\otimes\left(\hat{F}_{2^{N-n+1}}\circ\left(\hat{\mathbb{1}}_2\otimes\hat{F}_{2^{N-n}}^{-1}\right)\right)\circ\hat{S}_n. \tag{3.13}$$

Since $\hat{S}_1$ is the unit operator, it is clear that $\hat{B}_{N,1}$ is the Balazs–Voros–Saraceno quantum baker's map (3.6). It is worth mentioning here that Ermann and Saraceno [22] have recently proposed and investigated an even larger family of quantum baker's maps, which includes all of the above quantizations as members. For the purposes of this paper, however, we need only consider $\hat{B}_{N,n}$.

We can also write

$$\hat{B}_{N,n} = \hat{\mathbb{1}}_{2^{n-1}}\otimes\hat{B}_{N-n+1,1}\circ\hat{S}_n, \tag{3.14}$$

which shows that the action of $\hat{B}_{N,n}$ is a shift of the n leftmost qubits followed by an application of the Balazs–Voros–Saraceno baker's map to the $N-n+1$ rightmost qubits. At each iteration, the shift map $\hat{S}_n$ does two things: it shifts the nth qubit, the most significant qubit in position that was subject to the previous application of $\hat{B}_{N-n+1,1}$, out of the region subject to the next application of $\hat{B}_{N-n+1,1}$, and it shifts the most significant qubit in position (first qubit) into the region of subsequent application of $\hat{B}_{N-n+1,1}$.

The quantum baker's map $\hat{B}_{N,n}$ takes a state localized at $1a_{N-n}\ldots a_1\bullet x_1\ldots x_n1$ to a state localized at $1a_{N-n}\ldots a_1x_1\bullet x_2\ldots x_n1$. The decrease in the number of position bits and

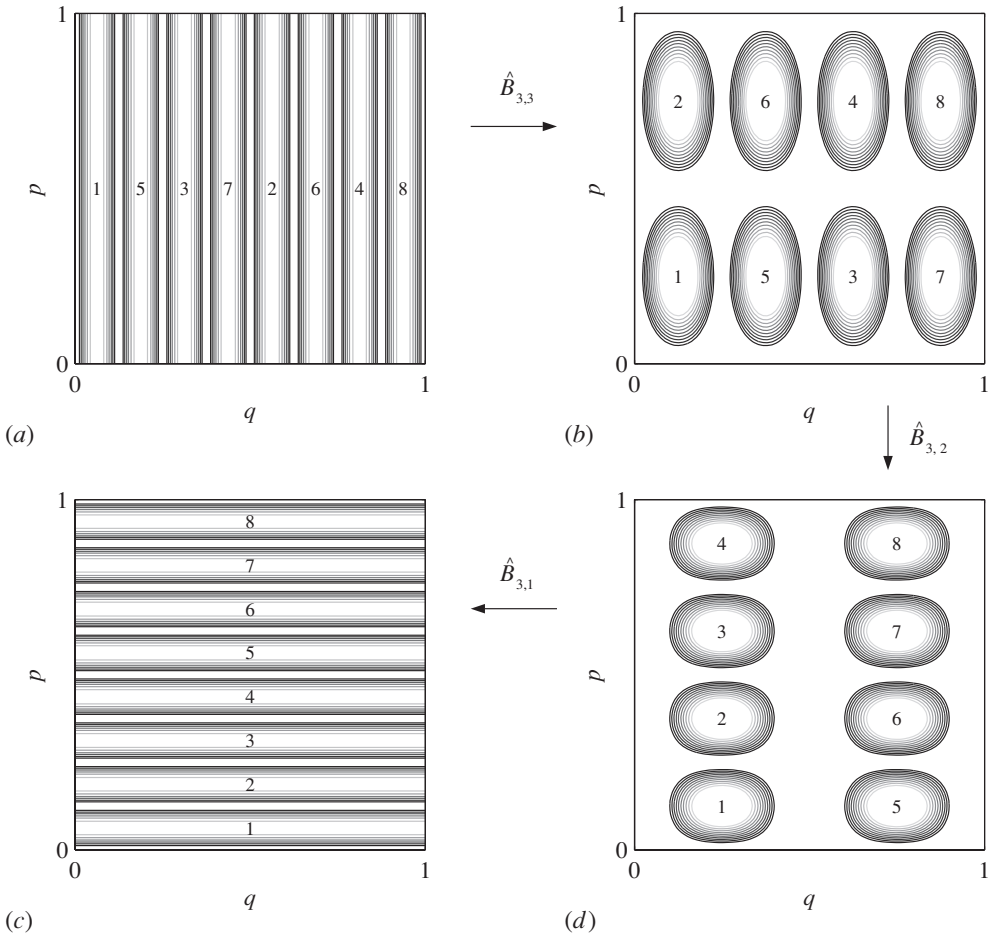

Figure 1. Husimi function for each partially Fourier transformed state (3.11) when $N = 3$: (*a*) $n = 3$, (*b*) $n = 2$, (*c*) $n = 0$ and (*d*) $n = 1$. The action of the quantum baker's map $\hat{B}_{3,3}$ is to map the eight states in (*a*) to the eight states in (*b*), as shown by the numbers labelling the states. Similarly, $\hat{B}_{3,2}$ and $\hat{B}_{3,1}$ map one set of partially Fourier transformed states to another, as indicated by the arrows. The map $\hat{B}_{3,1}$ is the Balazs–Voros–Saraceno quantum baker's map.

increase in momentum bits enforces a stretching and squeezing of phase space in a manner resembling the classical baker's map. In figure 1(*a*)–(*d*), we plot the Husimi function (defined as in [23]) for the partially Fourier transformed states (3.11) when $N = 3$, and $n = 3, 2, 0$ and 1, respectively. The quantum baker's map is a one-to-one mapping of one basis to another, as shown in the figure.

One useful representation of our quantum baker's maps, introduced in [10], starts from using standard techniques [24] to write the partially transformed states (3.10) as product states:

$$|a_{N-n}\ldots a_1 \bullet x_1\ldots x_n\rangle = \mathrm{e}^{\pi \mathrm{i}(0\cdot a_1\ldots a_{N-n}1)}\left(\bigotimes_{k=1}^{n}|x_k\rangle\right) \otimes\left(\bigotimes_{k=n+1}^{N}\frac{1}{\sqrt{2}}(|0\rangle + \mathrm{e}^{2\pi \mathrm{i}(0\cdot a_{N-k+1}\ldots a_{N-n}1)}|1\rangle)\right). \tag{3.15}$$

These input states are mapped by $\hat{B}_{N,n}$ to output states

$$|a_{N-n}\ldots a_1x_1\bullet x_2\ldots x_n\rangle = \mathrm{e}^{\pi\mathrm{i}(0\cdot x_1a_1\ldots a_{N-n}1)}\left(\bigotimes_{k=2}^{n}|x_k\rangle\right)$$
$$\otimes\left(\bigotimes_{k=n+1}^{N}\frac{1}{\sqrt{2}}(|0\rangle+\mathrm{e}^{2\pi\mathrm{i}(0\cdot a_{N-k+1}\ldots a_{N-n}1)}|1\rangle)\right)\otimes\frac{1}{\sqrt{2}}(|0\rangle+\mathrm{e}^{2\pi\mathrm{i}(0\cdot x_1a_1\ldots a_{N-n}1)}|1\rangle). \tag{3.16}$$

These forms show that the quantum baker's map $\hat{B}_{N,n}$ shifts the states of all the qubits to the left, except the state of the leftmost qubit. The state $|x_1\rangle$ of the leftmost qubit can be thought of as being shifted to the rightmost qubit, where it suffers a controlled phase change that is determined by the state parameters $a_1,\ldots,a_{N-n}$ of the original 'momentum qubits'. The quantum baker's map can thus be written as a shift map on a finite string of qubits, followed by a controlled phase change on the least significant qubit in position. In [25] this shift representation was developed into a useful tool. Using an approach based on coarse graining in this representation, the classical limit of the quantum baker's maps was investigated.

The classical limit for the above quantum baker's maps was also investigated in [23], using an analysis based on the limiting behaviour of the coherent-state propagator of $\hat{B}_{N,n}$. When $D=2^N\to\infty$, the total number of qubits N necessarily becomes infinite, but one has a considerable choice in how to take this limit. For example, we could use only one position bit, thus fixing $n=1$, and let the number of momentum bits $N-1$ become large. This is the limiting case of the Balazs–Voros–Saraceno quantization. There is, however, a wide variety of other scenarios to consider, e.g., $n=N/2$ or $n=2N/3-1$ as $N\to\infty$. In [23] it was shown that provided the number of momentum bits $N-n$ approaches infinity, the correct classical behaviour is recovered in the limit. If the number of momentum bits remains constant, i.e., $n=N-k$ (k constant) as $N\to\infty$, a stochastic variant of the classical baker's map is found. In the special case $n=N$ this variant takes the form

$$s=\ldots s_{-2}s_{-1}s_0\bullet s_1s_2s_3\ldots\to s'=\ldots s_{-2}s_{-1}s_0r\bullet s_2s_3\ldots \tag{3.17}$$

when written in the symbolic-dynamical language of equation (3.3). The bit r takes the value s_1 with probability $\cos^2[\pi/2(0\cdot s_0s_{-1}s_{-2}\cdots-1/2)]$ (and $1-s_1$ otherwise). These results are consistent with those obtained previously [25].

The extremal map has other interesting properties. All finite-dimensional unitary operators are quasi-periodic; the quantum baker's map $\hat{B}_{N,N}$, however, is strictly periodic,

$$(\hat{B}_{N,N})^{4N}=\hat{\mathbb{1}}, \tag{3.18}$$

as we show below. All its eigenvalues, therefore, are $4N$th roots of unity, i.e., of the form $\mathrm{e}^{\pi\mathrm{i}k/2N}$, and hence, there are degeneracies when $N>4$. This represents a strong deviation from the predictions of random matrix theory [26]. The eigenstates of the extremal map were recently studied by Anantharaman and Nonnenmacher [27], where $\hat{B}_{N,N}$ (with periodic rather than antiperiodic boundary conditions) was called the 'Walsh-quantized' baker's map. The above degeneracy in the eigenvalues allows constructions of eigenstates that remain partially localized in the semiclassical limit, which means that 'quantum unique ergodicity' [28] fails for this quantization.

The periodicity of the extremal map (3.18) can be easily shown after noting that $\hat{B}_{N,N}=-\mathrm{i}\hat{G}_{N-1}\circ\hat{S}_N=-\mathrm{i}(\hat{\mathbb{1}}_{2^{N-1}}\otimes\hat{F}_2)\circ\hat{S}_N$; i.e., $\hat{B}_{N,N}$ is a shift followed by application of the unitary

$$\hat{C}\equiv-\mathrm{i}\hat{F}_2=\tfrac{1}{\sqrt{2}}(\mathrm{e}^{-\pi\mathrm{i}/4}(|0\rangle\langle0|+|1\rangle\langle1|)+\mathrm{e}^{\pi\mathrm{i}/4}(|0\rangle\langle1|+|1\rangle\langle0|))=\mathrm{e}^{-\pi\mathrm{i}/4}\,\mathrm{e}^{(\pi i/4)\hat{\sigma}_x}, \tag{3.19}$$

which is a rotation by 90° about the x axis, to the least significant position qubit. On product states, the action of $\hat{B}_{N,N}$ can be written explicitly as

$$\hat{B}_{N,N}|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_N\rangle = |\psi_2\rangle \otimes \cdots \otimes |\psi_N\rangle \otimes \hat{C}|\psi_1\rangle. \tag{3.20}$$

Since $\hat{C}^4 = 1$, we get the property (3.18). One can also see that $\hat{B}_{N,N}$ cannot entangle initial product states.

When $n < N$, the action of the quantum baker's map is similar to equation (3.20), but with a crucial difference. After the qubit string is cycled, instead of applying a unitary to the rightmost qubit, a joint unitary is applied to all of the $N-n+1$ rightmost qubits. As discussed above, this joint unitary can be realized as controlled phase change of the rightmost qubit, where the control is by the state parameters $a_1, \ldots, a_{N-n}$ of the original momentum qubits. This controlled phase change means that initial product states become entangled.

Entanglement production under iterations of the quantum baker's maps was the subject of a recent paper [29]. Since the entangling controlled-phase change involves an increasing number of qubits as n decreases from $n = N$ to $n = 1$ (the Balazs–Voros–Saraceno map), one might expect that the entanglement increases as n ranges from N to 1. What was found, however, is that provided n is not too close to N, all the maps are efficient entanglement generators, but the greatest entanglement is produced when n is roughly midway between N and 1. Starting with a uniform distribution of initial product states, the mean entanglement 'quantum-baked' into the distribution was found to saturate at a level near to that expected in random states. The small deviations from the entanglement of random states might be due to hidden symmetries in the quantum baker's maps [30].

Lastly, we mention another difference between the extremal quantum baker's map $\hat{B}_{N,N}$ and other members of the baker's map family of quantizations. Ermann, Paz, and Saraceno [31] have found that when a system with the dynamics of a quantum baker's map is cast in the role of an environment acting on another quantum system, the extremal quantum baker's map $\hat{B}_{N,N}$ is less effective at inducing decoherence than other members of the family. In particular, they showed that while the entropy production rates of the different quantum baker's maps are indistinguishable on a short time scale, which scales linearly with N, $\hat{B}_{N,N}$ saturates much sooner than the other maps, thus displaying the behaviour expected for regular systems.

In view of the above described anomalous behaviour experienced by the extremal map, $\hat{B}_{N,N}$, our curiosity now invites an investigation into the various currently prevailing perturbation-based tests for quantum chaos, as applied to our class of quantum baker's maps. We start, however, by investigating the simplest example, the extremal map $\hat{B}_{N,N}$ itself.

4. Chaos in the extremal quantum baker's map?

4.1. The extremal map and perturbations

In section 3 we considered different quantizations of the baker's map as unitary transformations, $\hat{B}_{N,n}$ ($n = 1, \ldots, N$), on a set of N qubits. When written in the form (3.14), each of these transformations consists of two steps: a cyclic shift, $\hat{S}_n$, in which the n leftmost qubits are shifted without otherwise being altered and a unitary transformation on the rightmost $N-n+1$ qubits. For the extremal quantum baker's map, $n = N$, which we consider in this section, this second transformation is the gate $\hat{C}$ of equation (3.19), which acts only on the single rightmost qubit as in equation (3.20) and rotates it by 90° about the x axis.

We first examine the behaviour of $\hat{B}_{N,N}$ under perturbations after each iteration. In an effort not to affect qualitatively the dynamics of the map itself, we choose our perturbations to be correlated across the smallest possible distances in phase space. One choice might then

be to perturb only the single rightmost, least significant qubit in the position basis. Indeed, such a choice leads to the smallest changes in position. As a consequence of the uncertainty principle, however, perturbing the *least* significant qubit in the position basis causes correlated changes across the *greatest* distances in momentum. In opposition to our classical intuition, no single qubit can be thought of as being 'more significant' than another in an overall phase-space sense. Our particular choice of qubit upon which to perturb does not affect the phase-space area of correlated changes made to a state; however, perturbations affecting the middle qubit(s) give rise to correlated changes across the smallest phase-space distances. In the present analytical study, it is simplest to take the rightmost qubits as being the least significant. These considerations are revisited later in our numerical investigations, where we instead choose to perturb the middle qubit(s).

Suppose that with each step we perturb the m rightmost qubits, where $m \ll N$, by applying an m-qubit unitary transformation $\hat{U}_k^{(m)}$ chosen at random. For the moment, we assume that *any* transformation is allowed, but the arguments still work even if only a finite set of transformations is allowed.

For simplicity, we assume that the system is initially in the tensor-product state $|0\rangle^{\otimes N}$. Suppose that $m = 1$ and that the perturbation affects only the single rightmost qubit. Then the perturbation operators are all of the form $\hat{\mathbb{1}}_{2^{N-1}} \otimes \hat{U}_k^{(1)}$, and the state after a single step becomes

$$|0\rangle^{\otimes N} \to |0\rangle^{\otimes N-1} \otimes \left(\hat{U}_{k_1}^{(1)}\hat{C}|0\rangle\right), \tag{4.1}$$

while the next step transforms it to

$$|0\rangle^{\otimes N-1} \otimes \left(\hat{U}_{k_1}^{(1)}\hat{C}|0\rangle\right) \to |0\rangle^{\otimes N-2} \otimes \left(\hat{U}_{k_1}^{(1)}\hat{C}|0\rangle\right) \otimes \left(\hat{U}_{k_2}^{(1)}\hat{C}|0\rangle\right), \tag{4.2}$$

and so forth. It is clear that the above dynamics does not explore the entire Hilbert space, since the state remains a tensor product as long as the perturbation is restricted to a single qubit.

The situation changes if we let $m = 2$. Since two-qubit gates between nearest neighbours are sufficient for universal quantum computation, any state can be produced by the shift map plus two-qubit perturbations. It does not follow, however, that all states can be reached *quickly*; in general, the number of gates needed to reach a generic state of N qubits increases exponentially with N, which implies that many 'rounds' (complete sets of N steps) are needed to reach most states. On the other hand, as we saw in section 2.1.2, the perturbation need not sample generic vectors to elicit evidence for hypersensitivity, so considerations of universality in quantum computation and the time needed to sample generic states provide little information about hypersensitivity.

4.2. Signatures of chaos for $\hat{B}_{N,N}$

In this section we show that the extremal quantum baker's map is chaotic according to two popular signatures of quantum chaos: it displays a linear entropy increase when coupled to an environment, and the fidelity between two vectors evolving according to the original map and a slightly changed version of the map decreases exponentially. It does not, however, display hypersensitivity to perturbation.

4.2.1. Fidelity decay.

Define a modified baker's map by

$$\hat{B}'_{N,N}|x_1\rangle \otimes \cdots \otimes |x_N\rangle \equiv |x_2\rangle \otimes \cdots \otimes |x_N\rangle \otimes \hat{U}^{(1)}\hat{C}|x_1\rangle, \tag{4.3}$$

where $\hat{U}^{(1)}$ is a single-qubit unitary map satisfying $0 < |\langle 0|\hat{C}^\dagger\hat{U}^{(1)}\hat{C}|0\rangle| < 1$. We can define $\lambda = -2\ln|\langle 0|\hat{C}^\dagger\hat{U}^{(1)}\hat{C}|0\rangle| > 0$. Letting $|\psi(t)\rangle = (\hat{B}_{N,N})^t|\psi_0\rangle$ and $|\psi'(t)\rangle = (\hat{B}'_{N,N})^t|\psi_0\rangle$,

where $|\psi_0\rangle = |0\rangle^{\otimes N}$, we see that the fidelity decreases exponentially with the number of iterations:

$$F(t) = |\langle\psi'(t)|\psi(t)\rangle|^2 = \mathrm{e}^{-\lambda t}, \qquad t = 0, \ldots, N. \tag{4.4}$$

For greater numbers of iterations, the simple exponential decay is modified as the qubits experience more than one application of $\hat{C}$ or $\hat{U}^{(1)}\hat{C}$.

4.2.2. Linear increase of entropy. Let the environment be a collection of qubits in the maximally mixed state. After each iteration of the map, the register interacts with a fresh environment qubit. The interaction is given by a controlled $\hat{\sigma}_x$ operation, with the environment qubit acting as control and the target being the rightmost system qubit. In the notation of section 2.1.1, this binary perturbation amounts to an application of one of $\hat{U}_0 = \hat{\mathbb{1}}_{2^N}$ or $\hat{U}_1 = \hat{\mathbb{1}}_{2^{N-1}} \otimes \hat{\sigma}_x$ (chosen with equal probability) at each time step (this stochastic perturbation is the $m = 1$ model of section 4.1, with the single-qubit perturbation unitaries restricted to the identity and $\hat{\sigma}_x$). After tracing out the environment, one iteration of the perturbed map is described by the quantum operation

$$\mathcal{B}(\hat{\rho}) = \tfrac{1}{2}\hat{B}_{N,N}\hat{\rho}\hat{B}_{N,N}^{\dagger} + \tfrac{1}{2}(\hat{\mathbb{1}}_{2^{N-1}} \otimes \hat{\sigma}_x)\hat{B}_{N,N}\hat{\rho}\hat{B}_{N,N}^{\dagger}(\hat{\mathbb{1}}_{2^{N-1}} \otimes \hat{\sigma}_x). \tag{4.5}$$

For the initial state $\hat{\rho}_0 = (|0\rangle\langle 0|)^{\otimes N}$, and denoting by $\mathcal{B}^t$ the t-th iterate of $\mathcal{B}$, we have

$$\mathcal{B}^t(\hat{\rho}_0) = (|0\rangle\langle 0|)^{\otimes(N-t)} \otimes (\hat{\mathbb{1}}_2/2)^{\otimes t}, \tag{4.6}$$

since $\hat{\sigma}_x$ commutes with $\hat{C}$. The entropy of $\mathcal{B}^t(\hat{\rho}_0)$ is t bits. The entropy thus increases at a rate of 1 bit per iteration until it saturates at N bits after N steps. Under single-qubit perturbations, it is clear that the perturbed vectors explore an exponentially increasing number of Hilbert-space dimensions, but it is equally clear that they do not explore these dimensions randomly.

4.2.3. Hypersensitivity to perturbation. From the discussions in sections 2.1.2 and 4.1, it can be seen that $\hat{B}_{N,N}$ is not hypersensitive to perturbations that affect only the single rightmost qubit. For an extreme example of this, consider the binary perturbation in section 4.2.2 immediately above. After $t \leqslant N$ map iterations, there are 2^t perturbation histories, which, with the initial state $|\psi_0\rangle = |0\rangle^{\otimes N}$, correspond to orthogonal system vectors:

$$|\psi_{\mathbf{k}}\rangle = |0\rangle^{\otimes(N-t)}\hat{C}|k_1\rangle \otimes \cdots \otimes \hat{C}|k_t\rangle, \tag{4.7}$$

where $k_i \in \{0, 1\}$, in the notation of section 2.1.1. A measurement on the environment that groups these vectors according to the values of $k_1, \ldots, k_j$, with $0 \leqslant j \leqslant t$, reduces the average system entropy from $H_S = t$ to $H(I_{\min} = j) = t - I_{\min}$ bits. When $t = N$, the perturbed system vectors make up an orthonormal basis, and for $t \geqslant N$, the perturbation produces 2^{t-N} copies of an orthonormal basis. Thus, for $t \geqslant N$, the information–entropy trade-off relation is $H(I_{\min}) = N - I_{\min}$. For all t, our hypersensitivity parameter takes the value $s = 1$. In this extreme example, each bit of information purchases a bit of entropy reduction, as is always true when the perturbed vectors are drawn from an orthonormal basis with each vector in the basis having the same overall probability.

In general, stochastic perturbations that affect only a single qubit of the extremal quantum baker's map are expected to produce an information–entropy trade-off that is linear (to good approximation) near the maximal entropy. Although the hypersensitivity parameter generally varies with both the choice of perturbation and number of map iterations, its magnitude should not exceed 5.3, the bound on s for product states (equation (2.24)). The extremal map $\hat{B}_{N,N}$, therefore, does not exhibit exponential hypersensitivity to perturbation under single-qubit perturbations. By contrast, we have seen above that it does exhibit linear growth of entropy

and exponential decay of fidelity. Hypersensitivity to perturbation is evidently a finer sieve than the other two perturbation-based criteria.

The reason the perturbed extremal map does not explore Hilbert space efficiently is that the map itself produces no entanglement. In contrast, the nontrivial quantizations of the baker's map are efficient entanglement generators [29], producing entanglement that saturates after several iterations at a level close to that expected in random states. For these nontrivial quantizations, even a single-qubit perturbation, together with the entangling transformation of the unperturbed map, generically gives rise to a universal set of unitary gates, so in time the system can approach any state in the Hilbert space. Although the speed at which this happens remains unknown, our numerical results for hypersensitivity to perturbation, presented in the next section, suggest that if n is not too close to N, the perturbed nontrivial quantizations do efficiently explore all of Hilbert space.

Both the simple analytical argument above and the numerical results in the next section are for single-qubit perturbations. A systematic study of hypersensitivity to perturbations acting on two or more qubits is beyond our current numerical capabilities. In the remainder of this section, we present an analytical argument that suggests that, for a small number of time steps and for maps close to the extremal map $\hat{B}_{N,N}$, the information–entropy trade-off is linearly bounded even for entangling perturbations acting on two qubits.

We choose a perturbation that affects the two rightmost qubits, i.e., $k = 2$. Given some reasonable assumptions about the stochastic perturbation, if we average over all perturbations, the state after t steps is approximately equal to

$$\hat{\rho}^{(N)} \approx (|0\rangle\langle 0|)^{\otimes N-t-1} \otimes (\hat{\mathbb{1}}_2/2)^{\otimes t+1}. \tag{4.8}$$

This state has von Neumann entropy $H_S = t + 1$. We would like to acquire some information $I_{\min}$ about the perturbations which enables us to reduce this entropy by a small amount $\Delta H = H_S - H$.

We now show that the ratio $I_{\min}/\Delta H$ is bounded above by a quantity that is independent of t for all $t < N$. Suppose that after t steps, our system is in state (4.8). Now let us apply $\hat{B}_{N,N}$, but *not* the perturbation. If we trace out all but the two least significant qubits, these two qubits are in the state

$$\hat{\rho}^{(2)} = \hat{\mathbb{1}}_2/2 \otimes \hat{C}|0\rangle\langle 0|\hat{C}^\dagger, \tag{4.9}$$

which has 1 bit of entropy. The perturbation affects only these 2 bits, so the state of the other $N - 2$ qubits is irrelevant to the entropy increase. Now we apply the perturbation and get

$$\hat{\rho}^{(2)} \to \sum_\xi p_\xi \hat{U}_\xi \hat{\rho}^{(2)} \hat{U}_\xi^\dagger = (\hat{\mathbb{1}}_2/2)^{\otimes 2}, \tag{4.10}$$

where ξ labels which perturbation is performed, p_ξ is the probability of that perturbation, and $\hat{U}_\xi$ is the corresponding two-qubit unitary transformation. The entropy of the new state is 2 bits, giving an entropy increase of 1 bit.

Clearly, we can reduce the entropy by 1 bit if we can determine which $\hat{U}_\xi$ was actually performed. If the perturbations are drawn from a discrete set, the number of bits needed to determine this is given by the entropy of the distribution p_ξ, i.e.,

$$I_{\min} \leqslant -\sum_\xi p_\xi \log p_\xi. \tag{4.11}$$

If ξ is continuous, then *fully* determining $\hat{U}_\xi$ would require an infinite amount of information. Since the space of two-qubit operators is not very large, however, it does not take that much information to know $\hat{U}_\xi$ to a good approximation; e.g., we could achieve an entropy reduction

of nearly a bit at a cost of approximately 45 bits by knowing each of the 15 relevant parameters of an arbitrary two-qubit unitary with 3 bits of precision.

This procedure, while not necessarily optimal, places a rather low bound on the ratio $I_{\min}/\Delta H$, a bound independent both of the number of iterations, t, and the number of qubits, N. This argument changes little if we use $\hat{B}_{N,N-1}$ instead of $\hat{B}_{N,N}$, or $\hat{B}_{N,N-k}$ for k small compared to N. Nor does it change much if the perturbation affects k bits, so long as k is small compared to N. If the perturbation affects *many* bits, however, or if a quantization $\hat{B}_{N,N-k}$ is used for *large* k, the upper bound on $I_{\min}/\Delta H$ becomes so large that it gives little restriction.

The above provides some evidence for the conjecture that maps close to the extremal map, $\hat{B}_{N,N}$, do not exhibit exponential hypersensitivity to entangling perturbations. Since these results are valid only as long as $t \leqslant N$, i.e., as long as the number of iterations does not exceed the number of qubits, this evidence must be regarded as suggestive, but inconclusive.

5. Numerical results

We now investigate numerically the entire class of quantum baker's maps, $\hat{B}_{N,n}$ ($n = 1, \ldots, N$), in the context of the three perturbation-based criteria for quantum chaos. As remarked in section 4.1, perturbations affecting the middle qubit(s) cause correlated changes to the state in phase space across the smallest possible distances. Up until now we have applied all perturbations to the rightmost, least significant qubit in position. Since the application of $\hat{C}^\dagger$ as a perturbation to this qubit would undo the dynamics in momentum in the case of the extremal quantum baker's map, $\hat{B}_{N,N}$, one might judge this perturbation to be atypical, upsetting the crucial momentum dynamics of the map. To avoid this, we choose henceforth the total number of qubits N to be odd, and we perturb the middle qubit.

The perturbation we choose for this qubit is a simple binary perturbation, a rotation by angle $\pm 2\pi\alpha$ about the y axis,

$$\hat{U}_k(\alpha) \equiv \hat{\mathbb{1}}_{2^{(N-1)/2}} \otimes \mathrm{e}^{\pi \mathrm{i}(-1)^k \alpha \hat{\sigma}_y} \otimes \hat{\mathbb{1}}_{2^{(N-1)/2}}. \tag{5.1}$$

The perturbation is conditioned on the binary environmental states $|k\rangle_E$, $k = 0$ or 1. To be precise, after each iteration of the map, the system couples to its environment through a joint conditional evolution with end result

$$\hat{\rho}_{\text{total}} = \tfrac{1}{2}\big[\big(\hat{U}_0 \hat{\rho} \hat{U}_0^\dagger\big) \otimes |0\rangle_E\langle 0| + \big(\hat{U}_1 \hat{\rho} \hat{U}_1^\dagger\big) \otimes |1\rangle_E\langle 1|\big]. \tag{5.2}$$

To avoid inapt comparisons, we use $\alpha = 0.2$ (rotation angle 0.4π) and initial system state $|\psi_0\rangle = |0\rangle^{\otimes N}$ throughout this section.

We can cope with the unwanted perturbation in many different ways. One possibility is to accept an increase in entropy and average over all perturbation histories by tracing out the environment. The system entropy, $H_S = -\mathrm{tr}(\hat{\rho} \log \hat{\rho})$, then increases at an initially constant linear rate for all quantum baker's maps. This is shown for the quantizations using $N = 5$ qubits in figure 2(*a*) and $N = 11$ qubits in figure 2(*b*). The rate of entropy production for the different quantizations is nearly the same for the first $t = (N-1)/2$ iterations. In contrast to the other quantizations, the entropy produced by the extremal map, $\hat{B}_{N,N}$, remains constant for times beyond $t = (N-1)/2$ before resuming its climb towards the maximal entropy of 5 or 11 bits. Thus, although there is a quantitative change in entropy production at later times, the different baker's maps behave qualitatively the same. These numerical results support the simple analysis of section 4.2.2.

Alternatively, if the above entropy production proves unacceptable, we could instead perform a measurement on the environment at each time step to record which perturbation actually occurs. Consider the perturbation history of all 1's. The fidelity decay between two

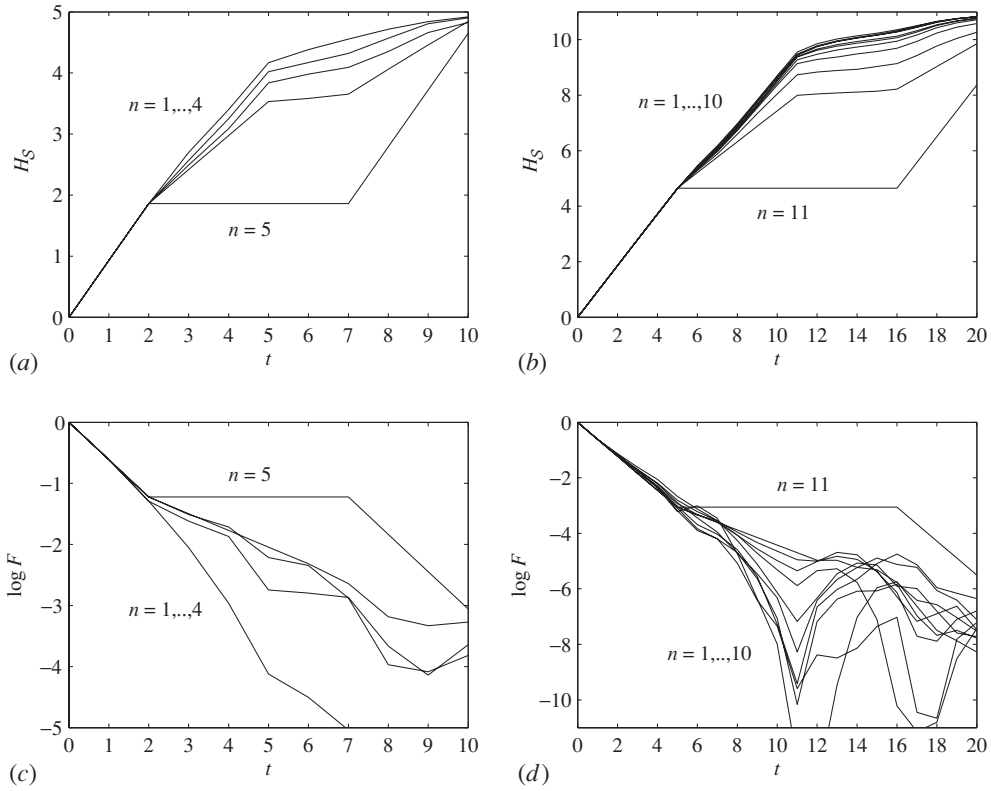

Figure 2. The rate of increase in entropy is initially linear for all (*a*) five-qubit and (*b*) eleven-qubit quantizations of the baker's map. The rate of decrease in fidelity is initially exponential for all (*c*) five-qubit and (*d*) eleven-qubit quantizations.

initially equal quantum states that evolve either according to this extremal perturbation history or the unperturbed map,

$$F(t) = |\langle\psi_0|\big(\hat{B}^{\dagger}_{N,n}\hat{U}^{\dagger}_{1}\big)^{t}(\hat{B}_{N,n})^{t}|\psi_0\rangle|^2, \tag{5.3}$$

might be used as an indicator of the underlying dynamics of the map. The rate of fidelity decay for all quantum baker's maps is initially exponential. This is shown in figures 2(*c*) and 2(*d*) for the five-qubit and eleven-qubit quantizations. Although the fidelity corresponding to the extremal map, $\hat{B}_{N,N}$, stalls at approximately $t = (N+1)/2$ iterations, all quantizations are found to exhibit decay rates which are initially exponential. Again, these numerical results support the simple analysis of section 4.2.1.

The iteration at which the entropies and fidelity decays first become appreciably different for the various quantizations remains at $t = (N+1)/2$ as N increases, and thus our conclusions become stronger in the limit of large N. To keep our analysis strictly in the quantum regime, however, we focus on the five-qubit quantizations for the remainder of this section.

To investigate hypersensitivity to perturbation, we first consider a particularly intuitive algorithm for grouping vectors, which is based on finding structure produced by the temporal order of the perturbations. Each grouping corresponds to measuring, after a fixed number of iterations t, the environment states—and, hence, the applied perturbation—at $l \leqslant t$ times. The

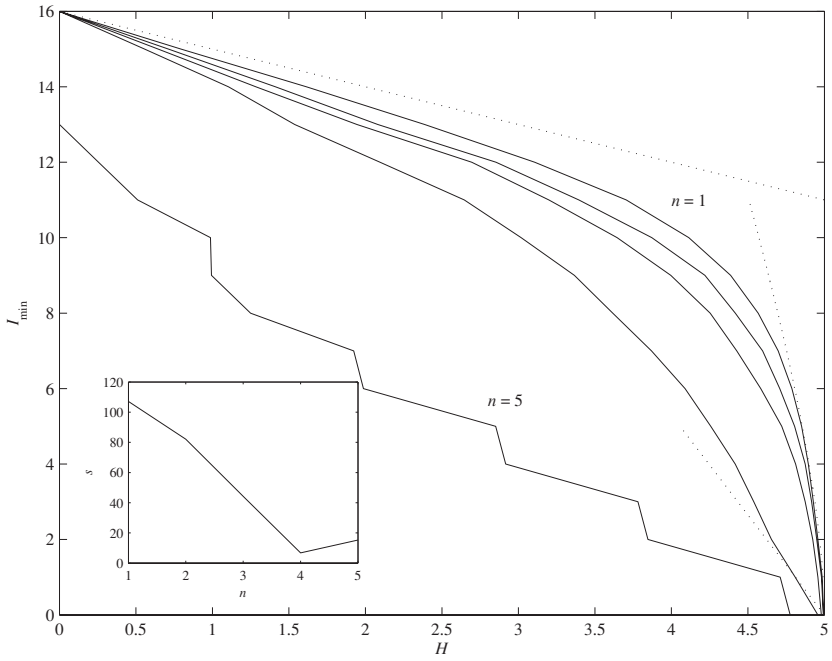

Figure 3. The minimum information $I_{\min}$ needed to reduce the entropy to H after 16 map iterations, using the temporal grouping algorithm, for all perturbed five-qubit quantizations of the baker's map ($n = 1, \ldots, 5$). The upper and rightmost dotted curves are the approximate sphere-grouping trade-off for random vectors in $d = 32$ dimensions (equation (2.17)), while the lower dotted line is the linear trade-off for random product vectors (equation (2.25)). The inset shows the hypersensitivity parameter s for each quantization.

2^t perturbation histories—and their final states—are thus grouped into 2^l sets, each containing 2^{t-l} states. It takes $\bar{I} = l$ bits to specify a group.

As an example of this procedure, suppose we have $t = 4$ iterations and we choose to measure the first and last states of the environment. Thus $l = 2$, and all histories are grouped into $2^l = 4$ sets of $2^{t-l} = 4$ binary strings in the form $0**0$, $1**0$, $0**1$ and $1**1$, where * denotes an arbitrary entry. Defining

$$|k_1 k_2 \cdots k_t\rangle \equiv \hat{U}_{k_1} \hat{B}_{N,n} \hat{U}_{k_2} \hat{B}_{N,n} \cdots \hat{U}_{k_t} \hat{B}_{N,n} |\psi_0\rangle, \tag{5.4}$$

the final state of the system, conditioned on measurement results i and j for the first and last environment qubits, is

$$\hat{\rho}_{ij} = 2^{l-t} \sum_{k_2, k_3 \in \{0,1\}} |i k_2 k_3 j\rangle\langle i k_2 k_3 j|. \tag{5.5}$$

Consequently, at the expense of storing $l = 2$ bits of information, we can, on average, reduce the entropy to

$$\bar{H} = -\frac{1}{2^l} \sum_{i,j \in \{0,1\}} \mathrm{tr}(\hat{\rho}_{ij} \log \hat{\rho}_{ij}). \tag{5.6}$$

The particular 2 bits stored in this example might not be the optimal choices. There are countless other measurements to consider, some of which no doubt lead to lower average

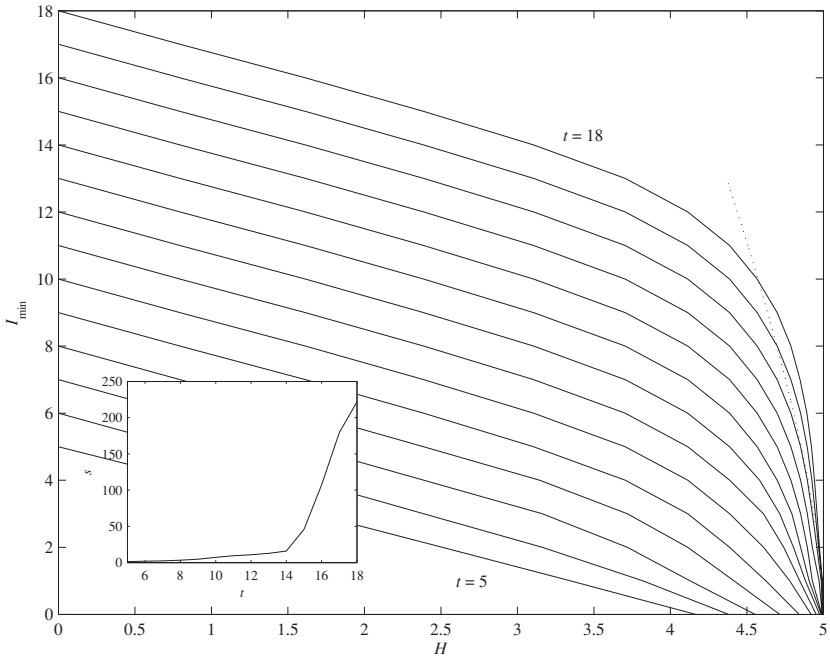

Figure 4. The minimum information $I_{\min}$ needed to reduce the entropy to H after t map iterations, where $5 \leqslant t \leqslant 18$, for the perturbed five-qubit Balazs–Voros–Saraceno quantization of the baker's map ($n = 1$), using the temporal grouping algorithm. The dotted curve is the sphere-grouping trade-off for random vectors in 32 dimensions. The inset shows the hypersensitivity parameter s at each iteration.

entropies. For the moment, however, we restrict our measurements to the above type and minimize $\bar{H}$ over the $\binom{t}{l}$ possible choices for the measurement times. Denoting this minimum entropy by H, the minimum information needed to reduce the average system entropy to H is then $I_{\min} = l$ bits. Although there is no guarantee that l is in fact the overall minimum, this simple scheme, which we call the *temporal* grouping algorithm, proved superior to previously used schemes (e.g., those discussed in [7]) for the maps and perturbations considered here.

Using this procedure, in figure 3 we plot $I_{\min}$ versus H for all perturbed five-qubit quantum baker's maps after $t = 16$ map iterations (solid lines). The perturbing parameter remains at $\alpha = 0.2$ and initial state at $|\psi_0\rangle = |0\rangle^{\otimes N}$. The region of interest regarding the question of quantum chaos lies to the right where $I_{\min}$ is small. Here we see that, except for the quantizations with n close to $N = 5$, a very large amount of information is required to reduce the system entropy by a small amount. This is a distinguishing characteristic of chaos, which is absent for the extremal quantization ($n = N$). When $I_{\min}$ is small, the information–entropy trade-off is characterized by the hypersensitivity parameter s (equation (2.18)). The inset in figure 3 shows this quantity for all five quantizations. Recall that $1/s$ is the reduction in system entropy purchased by gathering 1 bit of information about the environment. Although entropy reduction is affordable for the extremal quantization, 1 bit buys very little when n approaches 1. The dotted lines show our theoretical trade-offs for random vectors (upper and rightmost) and random product vectors (lower), given by equations (2.17) and (2.25), respectively. When n approaches 1, the information–entropy trade-off approaches that expected for random vectors, while for $n = N$, it is bounded by the trade-off for random product vectors.

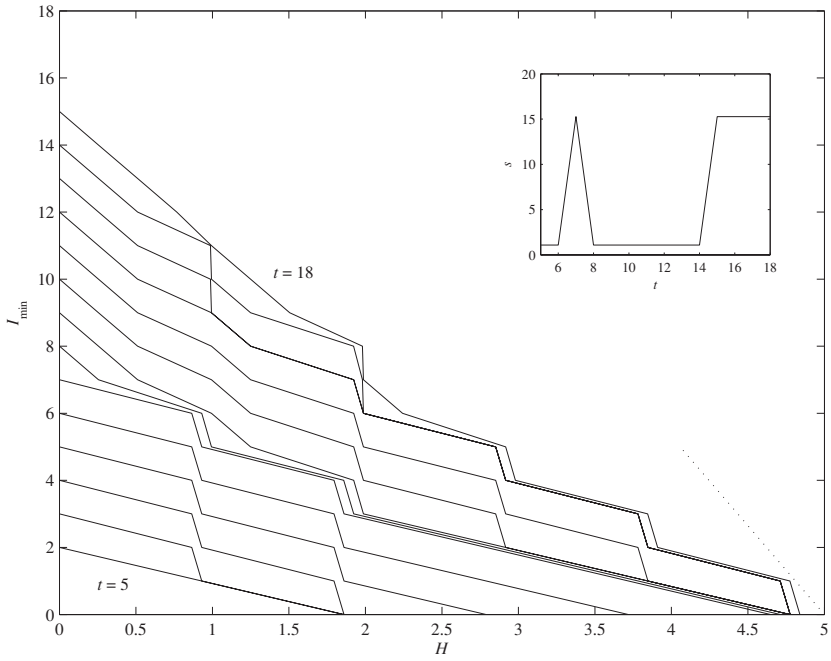

Figure 5. The minimum information $I_{\min}$ needed to reduce the entropy to H after t map iterations, where $5 \leqslant t \leqslant 18$, for the perturbed five-qubit extremal quantization of the baker's map ($n = N$), using the temporal grouping algorithm. The dotted line is the trade-off for random product vectors (equation (2.25)). The inset shows the hypersensitivity parameter s at each iteration.

Using the same grouping algorithm, we plot in figure 4 the information–entropy trade-off for a growing number of iterations of the Balazs–Voros–Saraceno quantization ($n = 1$). The figure shows $I_{\min}$ versus H for 5–18 iterations of $\hat{B}_{5,1}$, and in the inset, the corresponding value of s. To a rough approximation, our hypersensitivity signature s appears to grow exponentially with the number of iterations. This map thus exhibits numerical evidence of exponential hypersensitivity to perturbation. Note, however, that the trade-off violates the sphere-grouping bound (2.17) derived from random states for $t \gtrsim 16$ (dotted line), which means the current method for gathering information about the environment is not optimal, a situation we discuss further below. The extremal quantization displays a strikingly different behaviour. The information–entropy trade-off and the parameter s for 5–18 iterations of $\hat{B}_{5,5}$ are shown in figure 5. In this case the information–entropy trade-off remains approximately linear for all levels of iteration, with a very roughly constant s. There is no evidence of hypersensitivity to perturbation.

We now investigate the hypersensitivity parameter in greater detail for the Balazs–Voros–Saraceno and extremal quantizations. The graphs of s in the insets of figures 4 and 5 are redrawn, now on a logarithmic scale, as the dashed lines in figures 6(a) and (b), respectively. The horizontal dotted lines in each of these figures are the upper bounds (2.22) and (2.24), respectively, corresponding to the values of s for random vectors and random product vectors. Note that in both cases the dashed lines cross these bounds. This indicates that the temporal groupings used up until now are not optimal. Indeed, in the case of the Balazs–Voros–Saraceno quantization, by considering groupings which correspond to partitions of projective Hilbert space into two equal volumes, we find that 1 bit of information can buy larger entropy

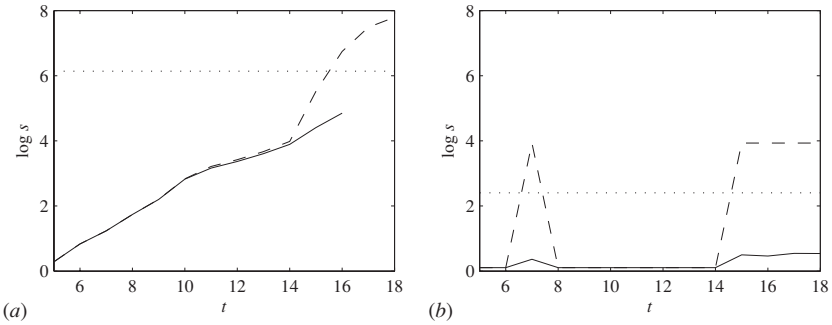

Figure 6. The hypersensitivity parameter s after t map iterations, where $5 \leqslant t \leqslant 18$, for the perturbed five-qubit (*a*) Balazs–Voros–Saraceno ($n = 1$) and (*b*) extremal ($n = N$) quantizations of the baker's map. The spurious higher values of s arising from the temporal groupings (dashed lines) are significantly reduced using a genetic-algorithm approach (solid lines). The dotted lines are the approximate upper bounds on s corresponding to (*a*) random vectors (equation (2.22)) and (*b*) random product vectors (equation (2.24)).

reductions when $t \geqslant 16$. This method of grouping vectors, however, works well only for distributions that are close to random.

We now consider grouping algorithms that are not constrained by a supposed temporal structure of the vector distribution. Although optimal groupings can always be found by simply testing every possibility, the size of the search space is doubly exponential in the number of map iterations. We thus turn to the theory of combinatorial optimization. Specifically, a simple genetic algorithm [32] was used to partition vectors into two groups with the goal of minimizing the average conditional entropy $\bar{H}$. Although these groups were not constrained to be of equal size, the returned solution always corresponded to $\bar{I} = 1 \pm 0.003$, and thus we can take $H(I_{\min} = 1) = \bar{H}$ to a very good approximation. The corresponding value of s for this method is plotted as the solid line in figure 6. In many cases the genetic algorithm located precisely the same vector grouping that was found previously by the temporal grouping algorithm. The spurious higher values of s, however, are now significantly reduced for both quantizations. Although figure 6(*a*) remains incomplete due to computational constraints, for the data points calculated, $\log s$ has regained its linear approach to the upper bound, where it eventually will saturate. The difference between the two quantizations under single-qubit perturbations is now difficult to dispute. The criterion of hypersensitivity to perturbation thus unmistakeably distinguishes the dynamics of the extremal quantum baker's map ($n = N$) as qualitatively different from the Balazs–Voros–Saraceno quantization ($n = 1$).

6. Conclusion

This paper addresses the difficult question of how to characterize quantum chaos dynamically in the hard quantum regime, far from the classical limit where signatures of the classical sensitivity to initial conditions can be identified in the quantum properties of a system. In this hard quantum regime, criteria for quantum chaos rely on studying the effects of perturbing the quantum dynamics. In this paper we study the three perturbation-based criteria that have been proposed: linear increase of entropy when a system is coupled to a perturbing environment; exponential decay of the fidelity between the unperturbed dynamics and a modified dynamics; and hypersensitivity to perturbation under stochastic perturbations of the dynamics. Hypersensitivity to perturbation is formulated in terms of the entropy reduction

achieved by acquiring information about the perturbation, which we call the information–entropy trade-off. Of these three criteria, hypersensitivity to perturbation is by far the most difficult to define rigorously and to investigate analytically and numerically.

We apply these three criteria to a set of qubit-based quantizations of the baker's map. These quantizations range from a map that has a trivial unentangling shift dynamics to the original Balazs–Voros–Saraceno quantized map, which is highly entangling. We find, through a combination of analytical arguments and numerical results, that all the quantizations exhibit a linear entropy increase and an exponential fidelity decay. In contrast, we show that the criterion of hypersensitivity to perturbation distinguishes the entangling quantizations from the shift map. In particular, by focusing on the trivial shift map and the Balazs–Voros–Saraceno map, our numerical work on hypersensitivity provides compelling evidence that these maps behave quite differently under stochastic perturbations, as revealed by studying the information–entropy trade-off for these maps.

The reason that hypersensitivity to perturbation is different from the other two perturbation-based criteria is not hard to identify. Linear entropy increase and exponential fidelity decay both tell one about how widely a perturbation disperses vectors in Hilbert space, but they provide no information about how randomly the perturbed vectors populate Hilbert space. In contrast, the randomness of the distribution of perturbed vectors is precisely what the information–entropy trade-off is sensitive to.

One way to quantify the trade-off in a single number is provided by the hypersensitivity parameter s: s^{-1} is *defined* in terms of the information–entropy trade-off as the entropy reduction purchased by an optimal 1 bit of information about the perturbation, but s can be *interpreted* as the number of Hilbert-space dimensions explored randomly by the perturbed vectors. Thus when s increases exponentially, as our numerical work indicates for the Balazs–Voros–Saraceno map, it signals that the perturbed vectors are populating an exponentially increasing number of dimensions in a random way. Linear entropy increase and exponential fidelity decay do not provide information about this property of chaotic quantum dynamics.

The numerical hypersensitivity results in this paper are obtained for single-qubit perturbations of the five-qubit baker's maps. Investigating more general perturbations would involve dealing with more qubits and thus would require considerably greater computational resources. To reduce the required computational resources in future investigations of hypersensitivity, it would be highly desirable to have an analytical argument or sufficient numerical evidence to demonstrate convincingly that the hypersensitivity parameter s is, by itself, a reliable signature of hypersensitivity to perturbation. Were this established, numerical investigations of hypersensitivity could be reduced from computing the entire information–entropy trade-off to calculating the trade-off only for the case of 1 bit of acquired information.

Acknowledgments

This work was supported in part by US Office of Naval Research Grant No N00014-00-1-0578, US National Science Foundation Grant No CCF-0448658, and the European Union IST-FET project EDIQIP. AJS acknowledges support from CIAR, CSE, iCORE and MITACS.

Appendix A. information–entropy trade-off for random vectors

Consider $\mathcal{N}$ state vectors distributed randomly in a d-dimensional Hilbert space, where we assume that $\mathcal{N} \geqslant d$. Given an entropy $H \leqslant \log d$, we group the vectors into groups that on average have this entropy and then ask how much information is required to specify a group.

For a given H, we are interested in the grouping that minimizes the required information, as in equation (2.9). The relation between $I_{\min}$ and H is the *information–entropy trade-off*. In this appendix we formulate an approximate information–entropy trade-off for random vectors by grouping the vectors into spheres on projective Hilbert space whose radius is given by a Hilbert-space angle ϕ. This being an approximate trade-off relation, we denote the information by I instead of $I_{\min}$.

The sphere-grouping model is based on results, given in [14], for the volume and entropy of a Hilbert-space sphere. The model was formulated in [7] and refined in [11].

The number of spheres of radius ϕ that can be packed into projective Hilbert space is given by equation (A.18) of [14],

$$\mathcal{N}_d(\phi) = \frac{\mathcal{V}_d}{\mathcal{V}_d(\phi)} = (\sin^2\phi)^{-(d-1)}, \tag{A.1}$$

where $\mathcal{V}_d(\phi)$ is the volume of a sphere of radius ϕ and $\mathcal{V}_d$ is the total volume of projective Hilbert space in d dimensions. The entropy of a mixture of vectors distributed uniformly within a sphere of radius ϕ is given by equations (B.5)–(B.6) of [14],

$$H_d(\phi) = -\lambda_0 \log \lambda_0 - (1-\lambda_0)\log\left(\frac{1-\lambda_0}{d-1}\right) = H_2(\lambda_0) + (1-\lambda_0)\log(d-1), \tag{A.2}$$

where $H_2(\lambda_0)$ is the binary entropy corresponding to the largest eigenvalue

$$\lambda_0 = 1 - \frac{d-1}{d}\sin^2\phi. \tag{A.3}$$

If we group the $\mathcal{N}$ vectors into groups of radius ϕ, the number of vectors per group is

$$\mathcal{N}_V(\phi) = \frac{\mathcal{N}}{\mathcal{N}_d(\phi)} = \mathcal{N}(\sin^2\phi)^{d-1}, \tag{A.4}$$

provided this number is not less than 1. There is a critical angle, ϕ_b, at which there is only one vector per group, i.e., $(\sin^2\phi_b)^{d-1} = 1/\mathcal{N}$. For $\phi \geqslant \phi_b$, there are $\mathcal{N}_d(\phi)$ groups, each containing approximately $\mathcal{N}_V(\phi)$ vectors, but for $\phi \leqslant \phi_b$, there are $\mathcal{N}$ groups, each containing just one vector. The information required to specify a group at resolution angle ϕ is thus $I(\phi) = \log\mathcal{N}$ for $\phi \leqslant \phi_b$ and $I(\phi) = \log\mathcal{N}_d(\phi)$ for $\phi \geqslant \phi_b$. There is another critical angle, ϕ_d, at which there are only two groups, i.e., $(\sin^2\phi_d)^{d-1} = 1/2$. For $\phi \geqslant \phi_d$, we cannot talk about grouping the vectors into spheres of equal radius, so we remove these angles ϕ from consideration. Thus we write the information to specify a group as

$$I(\phi) = \begin{cases} \log\mathcal{N}, & \phi \leqslant \phi_b, \\ \log\mathcal{N}_d(\phi) = -(d-1)\log(\sin^2\phi), & \phi_b \leqslant \phi \leqslant \phi_d. \end{cases} \tag{A.5}$$

For $\phi_b \leqslant \phi \leqslant \phi_d$, we have $\sin^2\phi = 2^{-I/(d-1)} = \mathrm{e}^{-I\ln 2/(d-1)}$, which shows that there are two important cases in terms of the number of vectors. If $\log\mathcal{N} \ll d$ ($\mathcal{N} \ll 2^d$), a situation we refer to as a *sparse* collection of random vectors, we have $I \leqslant \log\mathcal{N} \ll d$, giving

$$\sin^2\phi \approx 1 - \frac{I\ln 2}{d-1} \quad \Longrightarrow \quad \phi \approx \frac{\pi}{2} - \sqrt{\frac{I\ln 2}{d-1}} \tag{A.6}$$

over the entire range $\phi_b \leqslant \phi \leqslant \phi_d$. In particular, we have $\phi_b \approx \pi/2 - \sqrt{\ln\mathcal{N}/(d-1)}$. The number of groups increases so fast as ϕ retreats from $\pi/2$ that for a sparse collection, there is a group for each vector when the radius ϕ is still quite close to $\pi/2$. In contrast, if $\log\mathcal{N} \gg d$ ($\mathcal{N} \gg 2^d$), which we call a *dense* collection of vectors, then $\phi_b \approx \sin\phi_b = 2^{-\log\mathcal{N}/2(d-1)} \ll 1$, meaning that to get to one vector per group, the radius ϕ_b must be small.

When we turn to the entropy of the groups, it becomes clear that there is yet another critical angle, ϕ_c, the angle at which the number of vectors per group equals the Hilbert-space

dimension, i.e., $\mathcal{N}_V(\phi_c) = d$ or $I(\phi_c) = \log\mathcal{N} - \log d$. For $\phi \geqslant \phi_c$, there are sufficiently many vectors in each group to explore all the available Hilbert-space dimensions, so the entropy is close to the entropy $H_d(\phi)$ of a mixture of vectors distributed uniformly within a sphere of radius ϕ in d dimensions. In contrast, for $\phi_b \leqslant \phi \leqslant \phi_c$, the vectors in a group can explore roughly only $\mathcal{N}_V(\phi) = \mathcal{N}(\sin^2\phi)^{d-1} = 2^{-I}\mathcal{N}$ dimensions, thus giving an entropy close to $H_{\mathcal{N}_V(\phi)}(\phi)$. Finally, for $\phi \leqslant \phi_b$, there is only one vector per group, so $H = 0$.

Our main interest is the relation between H and I, so we eliminate the radius ϕ from the above expressions. The region $\phi \leqslant \phi_b$ gives $H = 0$ and $I = \log\mathcal{N}$. For $\phi_b \leqslant \phi \leqslant \phi_c$, i.e., $\log\mathcal{N} \geqslant I \geqslant \log\mathcal{N} - \log d$, we have

$$H = H_{\mathcal{N}_V(\phi)}(\phi) = H_2(\lambda) + (1-\lambda)\log(2^{-I}\mathcal{N} - 1), \tag{A.7}$$

where

$$\lambda = 1 - \frac{2^{-I}\mathcal{N} - 1}{2^{-I}\mathcal{N}} 2^{-I/(d-1)} = 1 - 2^{-I/(d-1)}\left(1 - \frac{2^I}{\mathcal{N}}\right). \tag{A.8}$$

Finally, for $\phi_c \leqslant \phi \leqslant \phi_d$, i.e., $\log\mathcal{N} - \log d \geqslant I \geqslant 1$, we have $H = H_d(\phi)$, with

$$\lambda_0 = 1 - \frac{d-1}{d} 2^{-I/(d-1)}. \tag{A.9}$$

Summarizing, we have

$$H = \begin{cases} H_{\mathcal{N}_V(\phi)}(\phi) = H_2(\lambda) + (1-\lambda)\log(2^{-I}\mathcal{N} - 1), & \log\mathcal{N} \geqslant I \geqslant \log\mathcal{N} - \log d, \\ H_d(\phi) = H_2(\lambda_0) + (1-\lambda_0)\log(d-1), & \log\mathcal{N} - \log d \geqslant I \geqslant 1, \end{cases} \tag{A.10}$$

with λ and λ_0 given by equations (A.8) and (A.9). Equation (A.10) is the approximate trade-off relation we are seeking.

The important part of the trade-off relation is the part that is independent of the number of random vectors, i.e., for $1 \leqslant I \leqslant \log\mathcal{N} - \log d$. Note that to investigate this region, we need $\mathcal{N} \gg d$, but we do *not* need $\mathcal{N}$ so large that the random vectors sample generic vectors, which would require at least $\mathcal{N} \sim 2^d$ vectors, i.e., a dense collection. We emphasize that *we do not need a dense collection of vectors to investigate the important part of the trade-off relation*.

Before going further, it is useful to put the trade-off relation (A.10) in other forms, which can be easily specialized to the case of a sparse collection of vectors. For the second case, which is the case of interest, we can write

$$\begin{aligned} H_d(\phi) = \log d - \frac{1}{d}([d(1 - 2^{-I/(d-1)}) + 2^{-I/(d-1)}] \\ \times \log[d(1 - 2^{-I/(d-1)}) + 2^{-I/(d-1)}] - I2^{-I/(d-1)}). \end{aligned} \tag{A.11}$$

For a sparse collection of vectors, for which $I \leqslant \log\mathcal{N} - \log d \leqslant \log\mathcal{N} \ll d$, or anytime we have $I \ll d$, we can approximate this by

$$H_d(\phi) = \log d - \frac{1}{d}((1 + I\ln 2)\log(1 + I\ln 2) - I). \tag{A.12}$$

We can manipulate the first case in equation (A.10) in a similar way:

$$H_{\mathcal{N}_V(\phi)}(\phi) = \log\mathcal{N} - I - \lambda\log\left(\frac{\mathcal{N}\lambda}{2^I}\right) - (1-\lambda)\log\left(\frac{1-\lambda}{1 - 2^I/\mathcal{N}}\right). \tag{A.13}$$

The factor $2^I/\mathcal{N}$ increases from $1/d$ at $I = \log\mathcal{N} - \log d$ to 1 at $I = \log\mathcal{N}$. For a sparse collection, we can approximate λ by

$$\lambda = \frac{2^I}{\mathcal{N}} + \frac{I\ln 2}{d-1}\left(1 - \frac{2^I}{\mathcal{N}}\right). \tag{A.14}$$

The second term is always small. When the first term dominates, the second two terms in equation (A.13) are small. When the first term is as small or smaller than the second, the second two terms in equation (A.13) are again small. Thus for a sparse collection, it is always a good approximation to use $H_{\mathcal{N}_V(\phi)}(\phi) = \log \mathcal{N} - I$.

The conclusion of these considerations is that for sparse collections, the trade-off relation (A.10) is well approximated by

$$H = \begin{cases} \log \mathcal{N} - I, & \log \mathcal{N} \geqslant I \geqslant \log \mathcal{N} - \log d, \\ \log d - \dfrac{1}{d}((1 + I \ln 2) \log(1 + I \ln 2) - I), & \log \mathcal{N} - \log d \geqslant I \geqslant 1. \end{cases} \tag{A.15}$$

This is the form of the trade-off relation that we use in section 2.1.2. When d is large, it is quite a good approximation for sparse collections of random vectors, certainly more than adequate given the approximate character of the entire sphere-grouping model. These approximate expressions are poorest at the knee between the two behaviours, which is also where the approximate treatment of the grouping is at its worst.

Appendix B. Entropy of equal partitions of projective Hilbert space

Let $|e_j\rangle$, $j = 1, \ldots, d$, be an orthonormal basis for a d-dimensional Hilbert space, and let

$$\hat{P}_+ = \sum_{j=1}^{n} |e_j\rangle\langle e_j| \tag{B.1}$$

be the projector onto the subspace S_+ spanned by the first n vectors,

$$\hat{P}_- = \sum_{j=n+1}^{n+m} |e_j\rangle\langle e_j| \tag{B.2}$$

be the projector onto the subspace S_- spanned by the next m vectors, and $\hat{P}_0 = \hat{\mathbb{1}} - \hat{P}_+ - \hat{P}_-$ be the projector onto the subspace S_0 spanned by the remaining $d - n - m$ vectors. An arbitrary normalized vector can be expanded uniquely as

$$|\psi\rangle = \cos\xi(\cos\theta|\chi\rangle + \sin\theta|\eta\rangle) + \sin\xi|\phi\rangle, \tag{B.3}$$

where $|\chi\rangle \in S_+$, $|\eta\rangle \in S_-$, and $|\phi\rangle \in S_0$ are normalized vectors. The angle ξ is the Hilbert-space angle between $|\psi\rangle$ and the span of S_+ and S_-, and θ is the Hilbert-space angle between the projection of $|\psi\rangle$ into the span of S_+ and S_-, i.e., $(\hat{P}_+ + \hat{P}_-)|\psi\rangle$, and the subspace S_+.

We are interested in the density operator formed from all pure states whose projection into the span of S_+ and S_- is closer to S_+ than an angle Θ,

$$\hat{\rho} = \mathscr{N} \int_{\theta \leqslant \Theta} \mathrm{d}\mathcal{S}_{2d-1} |\psi\rangle\langle\psi|, \tag{B.4}$$

where $\mathscr{N}$ is a normalization factor. Here and throughout $\mathrm{d}\mathcal{S}_j$ denotes the standard integration measure on the j-sphere, and $\mathcal{S}_j = \int \mathrm{d}\mathcal{S}_j$ is the volume of the j-sphere. This region of states is the analogue of the intersection in three real dimensions of a wedge of opening angle 2Θ with the unit sphere. It is clear that $\hat{\rho}$ is invariant under unitary transformations that are block-diagonal in the three subspaces, which implies that $\hat{\rho}$ has the form

$$\hat{\rho} = \lambda_+ \hat{P}_+ + \lambda_- \hat{P}_- + \lambda_0 \hat{P}_0. \tag{B.5}$$

Our job is to determine the three eigenvalues, $\lambda_\pm$ and λ_0, which satisfy

$$n\lambda_+ + m\lambda_- + (d - n - m)\lambda_0 = 1. \tag{B.6}$$

It turns out that $\lambda_0 = 1/d$, as we show below, so we have

$$\lambda_- = \frac{1}{d}\left(1 + \frac{n}{m}(1 - d\lambda_+)\right). \tag{B.7}$$

A small change in $|\psi\rangle$ can be written as

$$\begin{aligned}|\mathrm{d}\psi\rangle = {} & \mathrm{d}\xi(-\sin\xi(\cos\theta|\chi\rangle + \sin\theta|\eta\rangle) + \cos\xi|\phi\rangle) + \sin\xi|\mathrm{d}\phi\rangle \\ & + \cos\xi(\mathrm{d}\theta(-\sin\theta|\chi\rangle + \cos\theta|\eta\rangle) + \cos\theta|\mathrm{d}\chi\rangle + \sin\theta|\mathrm{d}\eta\rangle).\end{aligned} \tag{B.8}$$

This gives a line element on normalized vectors,

$$\mathrm{d}s^2 = \langle\mathrm{d}\psi|\mathrm{d}\psi\rangle = \mathrm{d}\xi^2 + \sin^2\xi\langle\mathrm{d}\phi|\mathrm{d}\phi\rangle + \cos^2\xi(\mathrm{d}\theta^2 + \cos^2\theta\langle\mathrm{d}\chi|\mathrm{d}\chi\rangle + \sin^2\theta\langle\mathrm{d}\eta|\mathrm{d}\eta\rangle), \tag{B.9}$$

and a corresponding volume element on the $(2d-1)$-sphere of normalized vectors,

$$\begin{aligned}\mathrm{d}\mathcal{S}_{2d-1} = {} & \sin^{2(d-n-m)-1}\xi\cos^{2(n+m)-1}\xi\,\mathrm{d}\xi \\ & \times\cos^{2n-1}\theta\sin^{2m-1}\theta\,\mathrm{d}\theta\,\mathrm{d}\mathcal{S}_{2(d-n-m)-1}\,\mathrm{d}\mathcal{S}_{2n-1}\,\mathrm{d}\mathcal{S}_{2m-1}.\end{aligned} \tag{B.10}$$

Normalizing the density operator gives

$$\begin{aligned}1 = \mathrm{tr}(\hat{\rho}) &= \mathscr{N}\int_{\theta\leqslant\Theta}\mathrm{d}\mathcal{S}_{2d-1} \\ &= \mathscr{N}\mathcal{S}_{2(d-n-m)-1}\mathcal{S}_{2n-1}\mathcal{S}_{2m-1}\int_0^{\pi/2}\mathrm{d}\xi\,\sin^{2(d-n-m)-1}\xi\cos^{2(n+m)-1}\xi \\ &\quad\times\int_0^{\Theta}\mathrm{d}\theta\cos^{2n-1}\theta\sin^{2m-1}\theta.\end{aligned} \tag{B.11}$$

We first verify that $\lambda_0 = 1/d$. Letting $|e_0\rangle$ be any normalized vector in S_0, we have

$$\begin{aligned}\lambda_0 = \langle e_0|\hat{\rho}|e_0\rangle &= \mathscr{N}\int_{\theta\leqslant\Theta}\mathrm{d}\mathcal{S}_{2d-1}|\langle e_0|\psi\rangle|^2 \\ &= \mathscr{N}\mathcal{S}_{2n-1}\mathcal{S}_{2m-1}\int_0^{\pi/2}\mathrm{d}\xi\,\sin^{2(d-n-m)+1}\xi\cos^{2(n+m)-1}\xi \\ &\quad\times\int_0^{\Theta}\mathrm{d}\theta\cos^{2n-1}\theta\sin^{2m-1}\theta\int\mathrm{d}\mathcal{S}_{2(d-n-m)-1}|\langle e_0|\phi\rangle|^2.\end{aligned} \tag{B.12}$$

Using

$$\int\mathrm{d}\mathcal{S}_{2(d-n-m)-1}|\langle e_0|\phi\rangle|^2 = \frac{\mathcal{S}_{2(d-n-m)-1}}{d-n-m} \tag{B.13}$$

and the expression for the normalization constant from equation (B.11) and changing integration variable to $u = \sin^2\xi$, we get

$$\begin{aligned}\lambda_0 &= \frac{1}{d-n-m}\frac{\int_0^1\mathrm{d}u\,u^{d-n-m}(1-u)^{n+m-1}}{\int_0^1\mathrm{d}u\,u^{d-n-m-1}(1-u)^{n+m-1}} \\ &= \frac{1}{d-n-m}\frac{\Gamma(d-n-m+1)\Gamma(n+m)/\Gamma(d+1)}{\Gamma(d-n-m)\Gamma(n+m)/\Gamma(d)} \\ &= \frac{1}{d}.\end{aligned} \tag{B.14}$$

Similarly, to find λ_+, we let $|e_+\rangle$ be any normalized vector in S_+ and write

$$\begin{aligned}\lambda_+ = \langle e_+|\hat{\rho}|e_+\rangle &= \mathscr{N}\int_{\theta\leqslant\Theta}\mathrm{d}\mathcal{S}_{2d-1}|\langle e_+|\psi\rangle|^2 \\ &= \mathscr{N}\mathcal{S}_{2(d-n-m)-1}\mathcal{S}_{2m-1}\int_0^{\pi/2}\mathrm{d}\xi\,\sin^{2(d-n-m)-1}\xi\cos^{2(n+m)+1}\xi \\ &\quad\times\int_0^{\Theta}\mathrm{d}\theta\cos^{2n+1}\theta\sin^{2m-1}\theta\int\mathrm{d}\mathcal{S}_{2n-1}|\langle e_+|\chi\rangle|^2.\end{aligned} \tag{B.15}$$

Using

$$\int \mathrm{d}\mathcal{S}_{2n-1}|\langle e_+|\chi\rangle|^2 = \frac{\mathcal{S}_{2n-1}}{n} \tag{B.16}$$

and the expression for the normalization constant and changing integration variables to $u = \sin^2\xi$ and $v = \sin^2\theta$, we get

$$\begin{aligned}\lambda_+ &= \frac{1}{n}\frac{\int_0^1 \mathrm{d}u\, u^{d-n-m-1}(1-u)^{n+m}}{\int_0^1 \mathrm{d}u\, u^{d-n-m-1}(1-u)^{n+m-1}}\frac{\int_0^{\sin^2\Theta}\mathrm{d}v\, v^{m-1}(1-v)^n}{\int_0^{\sin^2\Theta}\mathrm{d}v\, v^{m-1}(1-v)^{n-1}}\\ &= \frac{1}{n}\frac{\Gamma(d-n-m)\Gamma(n+m+1)/\Gamma(d+1)}{\Gamma(d-n-m)\Gamma(n+m)/\Gamma(d)}\frac{\int_0^{\sin^2\Theta}\mathrm{d}v\, v^{m-1}(1-v)^n}{\int_0^{\sin^2\Theta}\mathrm{d}v\, v^{m-1}(1-v)^{n-1}}\\ &= \frac{n+m}{nd}\frac{\int_0^{\sin^2\Theta}\mathrm{d}v\, v^{m-1}(1-v)^n}{\int_0^{\sin^2\Theta}\mathrm{d}v\, v^{m-1}(1-v)^{n-1}}.\end{aligned} \tag{B.17}$$

We now specialize to the case of interest, $n = m$ and $\Theta = \pi/4$, so that $\hat{\rho}$ is constructed from pure states occupying one of two halves of Hilbert space:

$$\lambda_+ = \frac{2}{d}\frac{\int_0^{1/2}\mathrm{d}v\, v^{n-1}(1-v)^n}{\int_0^{1/2}\mathrm{d}v\, v^{n-1}(1-v)^{n-1}}. \tag{B.18}$$

The integrals can be evaluated as

$$\begin{aligned}\int_0^{1/2}\mathrm{d}v\, v^{n-1}(1-v)^n &= \frac{n!(n-1)!}{2(2n)!}\left(1+\frac{\Gamma(n+1/2)}{\sqrt{\pi}n!}\right),\\ \int_0^{1/2}\mathrm{d}v\, v^{n-1}(1-v)^{n-1} &= \frac{[(n-1)!]^2}{2(2n-1)!}.\end{aligned} \tag{B.19}$$

Plugging these results into equations (B.18) and (B.7), we get

$$\lambda_\pm = \frac{1}{d}\left(1 \pm \frac{\Gamma(n+1/2)}{\sqrt{\pi}n!}\right) = \frac{1}{d}\left(1 \pm \frac{(2n)!}{2^{2n}(n!)^2}\right). \tag{B.20}$$

When $n = 1$, we get $\lambda_+ = 3/2d$ and $\lambda_- = 1/2d$, and when $n = 2$, $\lambda_+ = 11/8d$ and $\lambda_- = 5/8d$. For large n (and d), we can use Stirling's formula to write

$$\lambda_\pm \approx \frac{1}{d}\left(1 \pm \frac{1}{\sqrt{\pi n}}\right). \tag{B.21}$$

The von Neumann entropy of $\hat{\rho}$ can be put in the form

$$\begin{aligned}H &= -n\lambda_+\log\lambda_+ - n\lambda_-\log\lambda_- - (d-2n)\lambda_0\log\lambda_0\\ &= \log d - \frac{2n}{d}(1 - H_2(d\lambda_+/2)).\end{aligned} \tag{B.22}$$

For fixed d, this is a decreasing function of n. For large n (and d), we can use equation (B.21) to write $H_2(d\lambda_+/2) \approx 1 - 1/2\pi n\ln 2$ and

$$H \approx \log d - \frac{1}{\pi d\ln 2} = \log d - \frac{0.46}{d}. \tag{B.23}$$

References

[1] Zurek W H and Paz J P 1994 *Phys. Rev. Lett.* **72** 2508
[2] Zurek W H and Paz J P 1995 *Physica* D **83** 300
[3] Peres A 1991 *Quantum Chaos* ed H A Cerdeira, R Ramaswamy, M C Gutzwiller and G Casati (Singapore: World Scientific) pp 73–102
[4] Peres A 1993 *Quantum Theory: Concepts and Methods* (Dordrecht: Kluwer)
[5] Caves C M 1993 *Physical Origins of Time Asymmetry* ed J J Halliwell, J Pérez-Mercader and W H Zurek (Cambridge: Cambridge University Press) p 47
[6] Schack R and Caves C M 1996 *Phys. Rev.* E **53** 3387
[7] Schack R and Caves C M 1996 *Phys. Rev.* E **53** 3257
[8] Caves C M and Schack R 1997 *Complexity* **3** 46
[9] Lichtenberg A J and Lieberman M A 1983 *Regular and Stochastic Motion* (New York: Springer) pp 268–75
[10] Schack R and Caves C M 2000 *Appl. Algebra Eng. Commun. Comput. (AAECC)* **10** 305
[11] Soklakov A N and Schack R 2000 *J. Mod. Opt.* **47** 2265
[12] Kraus K 1983 *States, Effects, and Operations. Fundamental Notions of Quantum Theory* (*Lecture Notes in Physics* vol 190) (Berlin: Springer)
[13] Schack R 1997 *Int. J. Theor. Phys.* **36** 209
[14] Schack R, D'Ariano G M and Caves C M 1994 *Phys. Rev.* E **50** 972
[15] Emerson J, Weinstein Y S, Saraceno M, Lloyd S and Cory D G 2003 *Science* **302** 2098
[16] Alekseev V M and Yakobson M V 1981 *Phys. Rep.* **75** 287
[17] Emerson J, Weinstein Y S, Lloyd S and Cory D G 2002 *Phys. Rev. Lett.* **89** 284102
[18] Weinstein Y W and Hellberg C S 2005 *Phys. Rev.* E **71** 035203(R)
[19] Gorin T, Prosen T, Seligman T H and Znidaric M 2006 *Preprint* quant-ph/0607050
[20] Balazs N L and Voros A 1989 *Ann. Phys.* **190** 1
[21] Saraceno M 1990 *Ann. Phys.* **199** 37
[22] Ermann L and Saraceno M 2006 *Preprint* nlin.CD/0606021
[23] Tracy M M and Scott A J 2002 *J. Phys. A: Math. Gen.* **35** 8341
[24] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[25] Soklakov A N and Schack R 2000 *Phys. Rev.* E **61** 5108
[26] Haake F 1991 *Quantum Signatures of Chaos* (Berlin: Springer)
[27] Anantharaman N and Nonnenmacher S 2005 *Preprint* math-ph/0512052
[28] Rudnick Z and Sarnak P 1994 *Commun. Math. Phys.* **161** 195
[29] Scott A J and Caves C M 2003 *J. Phys. A: Math. Gen.* **36** 9553
[30] Abreu R F and Vallejos R O 2006 *Preprint* quant-ph/0603261
[31] Ermann L, Paz J P and Saraceno M 2006 *Phys. Rev.* A **73** 012302
[32] Eiben A E and Smith J E 2003 *Introduction to Evolutionary Computing* (Berlin: Springer)