Google matrix and Ulam networks of intermittency maps

L. Ermann and D. L. Shepelyansky

Laboratoire de Physique Théorique (IRSAMC), Université de Toulouse–UPS, F-31062 Toulouse, France
and LPT (IRSAMC), CNRS, F-31062 Toulouse, France
(Received 19 November 2009; published 30 March 2010)

We study the properties of the Google matrix of an Ulam network generated by intermittency maps. This network is created by the Ulam method which gives a matrix approximant for the Perron-Frobenius operator of dynamical map. The spectral properties of eigenvalues and eigenvectors of this matrix are analyzed. We show that the PageRank of the system is characterized by a power law decay with the exponent \( \beta \) dependent on map parameters and the Google damping factor \( \alpha \). Under certain conditions the PageRank is completely delocalized so that the Google search in such a situation becomes inefficient.

DOI: 10.1103/PhysRevE.81.036221

PACS number(s): 05.45.Ac, 89.20.Hh

I. INTRODUCTION

In the 1960s, Ulam proposed a method to construct a matrix approximant for a Perron-Frobenius operator of dynamical systems, which is now known as the Ulam method [1]. The Ulam conjecture was that, in the limit of small cell discretization of the phase space, this method converges and gives the correct description of the Perron-Frobenius operator of a system with continuous phase space. This conjecture was shown to be true for hyperbolic maps of the interval [2]. Various types of more generic maps of an interval were studied in [3-5]. Further mathematical results have been obtained in [6-10] with extensions and prove of convergence for hyperbolic maps in higher dimensions. The mathematical analysis of nonuniformly expanding maps is now in progress [11]. At the same time it is known that the Ulam method applied to Hamiltonian systems with integrable islands of motion destroys the invariant curves thus producing a strong modification of properties of the Perron-Frobenius operator of the system with continuous phase space (see e.g., [12]).

Recently it was shown that the Ulam method naturally generates a class of directed networks, named Ulam networks, which properties have certain similarities with the world wide web (WWW) networks [12]. Thus, the Google matrix constructed for the Ulam networks built for the Chirikov typical map has a number of interesting properties showing a power law decay of the PageRank vector.

The classification of network nodes by the PageRank algorithm (PRA) was proposed by Brin and Page in 1998 [13] and became the core of the Google search engine used everyday by majority of internet users. The PRA is based on the construction of the Google matrix which can be written as (see e.g., [14] for details):

\[
G = \alpha S + (1 - \alpha)E/N.
\]  

Here, the matrix \( S \) is constructed from the adjacency matrix \( A \) of directed network links between \( N \) nodes so that \( S_{ij} = A_{ij}/\Sigma_i A_{ij} \) and the elements of columns with only zero elements are replaced by \( 1/N \). The second term in r.h.s. of Eq. (1) describes a finite probability \( 1-\alpha \) for WWW surfer to jump at random to any node so that the matrix elements \( E_{ij} = 1 \). This term stabilizes the convergence of PRA introducing a gap between the maximal eigenvalue \( \lambda = 1 \) and other eigenvalues \( \lambda_j \). Usually the Google search uses the value \( \alpha =0.85 \) [14]. The factor \( \alpha \) is also called the Google damping factor. By the construction \( \Sigma_j E_{ij} = 1 \) so that the asymmetric matrix \( G \) belongs to the class of Perron-Frobenius operators. Such operators naturally appear in the ergodic theory [15] and dynamical systems with Hamiltonian or dissipative dynamics [16,17].

The right eigenvector at \( \lambda =1 \) is the PageRank vector with positive elements \( p_j \) and \( \Sigma_j p_j = 1 \), the components \( p_j \) of this vector are used for ordering and classification of nodes. The PageRank can be efficiently obtained by a multiplication of a random vector by \( G \) which is of low cost since in average there are only about ten nonzero elements in a typical line of \( G \) of WWW. This procedure converges rapidly to the PageRank. All WWW nodes can be ordered by decreasing \( p_j p_j + p_{\max} \) so that the PageRank plays a significant role in the ordering of websites and information retrieval. The classification of nodes in the decreasing order of \( p_j \) values is used to classify importance of network nodes as it is described in more detail in [14].

Due to a spectacular success of the Google search the studies of PageRank properties became very active research filed in the computer science community. A number of interesting results in this field can be found in [18-20]. An overview of the field is available in [22]. It is established that for large WWW subsets \( p_j \) is satisfactory described by a scale-free algebraic decay with \( p_j \sim 1/j^\beta \) where \( j \) is the PageRank ordering index and \( \beta \sim 0.9 \) [14,23].

In this work we analyze the properties of Google matrix constructed from Ulam networks generated by one-dimensional (1D) intermittency maps. Such maps were introduced in [24] and studied in dynamical systems with intermittency properties (see e.g., [25-28]). A number of mathematical results on the measure distribution and slow mixing in such maps can be found in [29,30] (see also references therein). The mathematical properties of convergence of the Ulam method in such intermittency maps are discussed in a recent work [11]. The analysis of such 1D maps is simpler compared to the two-dimensional (2D) map considered in [12]: for example the PageRank at \( \alpha = 1 \) is described by the invariant measure of the map which can be find analytically as a function of map parameters. Following the approach discussed in [12,31] we study not only the PageRank but also the spectrum and the eigenstates of the
Google matrix generated by the intermittency maps. Indeed, the right eigenvectors \( \psi_i \) and eigenvalues \( \lambda_i \) of the Google matrix \( \mathbf{G} \psi_i = \lambda_i \psi_i \) are generally complex and their properties should be studied in detail to understand the behavior of the PageRank. We show that under certain conditions the properties of the PageRank can be drastically changed by parameter variation.

The results are presented in a following way: in Sec. II, we describe the class of intermittency maps and the distribution of links in the corresponding Ulam network, the spectral properties of the Google matrix and PageRank are considered in Secs. III and IV, the discussion of the results is presented in Sec. IV.

II. INTERMITTENCY MAPS

The intermittency maps of the interval considered in this paper are described by the two map functions depending on parameters and defined for the first model as:

\[
 f_1(x) = \begin{cases} 
 x + (2x)^z_1/2, & \text{for } 0 \leq x < 1/2 \\
 (2x - 1 - (1-x)^z_2 + 1/2^z_2)/(1 + 1/2^z_2), & \text{for } 1/2 \leq x \leq 1 
\end{cases}
\]  

(2)

and for the second model as

\[
 f_2(x) = \begin{cases} 
 x + (2x)^z_1/2, & \text{for } 0 \leq x < 1/2 \\
 a \sin[\pi(x - 1/2)], & \text{for } 1/2 \leq x \leq 1 
\end{cases}
\]  

(3)

The parameters \( z_1, z_2, a \) are positive numbers. The dynamics is given by the map \( \bar{x}=f_1(x) \) and \( \bar{x}=f_2(x) \). The map functions \( f_{1,2}(x) \) are shown in Fig. 1.

According to the usual theory of intermittency maps and ergodicity theory [24–26,29,30,32,33] in the case of chaotic dynamics the steady state invariant distribution \( g(x) \) of the map is proportional to a time \( t(x) \) spent by a trajectory at point \( x \) which is proportional to \( t \sim 1/x^{1-z_1} \) so that one has a power-law distribution at small values of \( x \):

\[
 g(x) \sim 1/x^{z_1-1}.
\]  

(4)

For \( f_1 \) map the dynamics is fully chaotic while for \( f_2 \) map a fixed point attractor appears for \( a > 0.945 \) when \( f_2(x)=x \).

The Ulam networks generated by the intermittency maps Eqs. (2) and (3) are constructed in a way similar to one described in [1,12]: the whole interval \( 0 \leq x \leq 1 \) is divided on \( N \) equal cells and \( N_t \) trajectories (randomly distributed inside cell) are iterated on one map iteration from cell \( j \) to obtain matrix elements for transitions to cell \( i \): \( S_{ij}=N_t(j)/N_t \) where \( N_t(j) \) is a number of trajectories arrived from cell \( j \) to cell \( i \).

The image of the density of Google matrix elements is shown in Fig. 2 for the first model. The structure of the matrix repeats the form of the map function \( f_1(x) \). We used from \( 10^4 \) to \( 10^6 \) cell trajectories \( N_c \), the obtained results are not sensitive to \( N_c \) variation in this interval.

The differential distribution of number of nodes \( N_t(\kappa) \) with ingoing or outgoing links \( \kappa \) is shown in Fig. 3. The first model shows a sharp drop of ingoing links and a power law decay of outgoing links. For the second model the situation is inverted. These properties can be understood from the following arguments. For the first model, the number of outgoing links is \( \kappa = d\bar{x}/dx = df_1(x)/dx \), the derivative is diverging.

FIG. 1. (Color online) Two types of intermittency map of the interval given by the map functions \( f_1(x) \) and \( f_2(x) \), the functions are identical for \( 0 \leq x < 1/2 \) but have different branches at \( 1/2 \leq x \leq 1 \) with \( f_1(x) \) (red/gray) and \( f_2(x) \) (blue/black); map functions are shown at \( z_1=2 \) and at different values of parameters \( z_2 \) and \( a \); the straight light gray line shows \( f(x)=x \).

FIG. 2. (Color online) Google matrix at \( \alpha=1 \) generated by the intermittency map \( f_1(x) \) at \( z_1=2, z_2=0.2, N=50, \) and \( N_t=10^6 \) [amplitude of matrix elements is changing from zero (black/blue) to 1 (red/gray)].
For the data of Fig. 3, x̄a is a general case where we have 1/κ0.4 out-going (black points) or ingoing (red/gray squares) links, respectively. Insets show data for small κ values in linear scale. The straight line shows the theoretical slope for outgoing links \( N_\text{out}(\kappa) \propto 1/\kappa^{0.4} \) first model, top panel) and for ingoing links \( N_\text{in}(\kappa) \propto 1/\kappa^3 \) second model, bottom panel).

For our case with \( z_1 = 2, \ a = 0.9 \), the condition \( N_\text{in}\sim x \sim 1/\kappa^{1/2} \) and

\[
N_\text{in} \sim dN_\text{in}/d\kappa \sim 1/\kappa^{\mu}, \ \mu = (2 - z_2)/(1 - z_2). \tag{5}
\]

For the data of Fig. 3 (top panel) at \( z_2 = 0.2 \) and \( z_2 = 0.9 \), the estimate \( \mu = 9/4 \) in a good agreement with the numerical data. For the second model the differential distribution of nodes \( dN(x)/dx \) is always finite and we have a sharp drop for outgoing links distribution. The number of ingoing links is \( \kappa = dx/d\kappa \sim 1/\kappa^{1/2} \) and since we have \( x \sim 1 - (1-x)^{2\nu} \) near \( x = 1 \) (in our case \( x = 1 \) but we consider here a general case), Hence, the number of nodes with \( \kappa \) links is \( N_\text{in}\sim \kappa \sim 1/\kappa^{2\nu(2\nu-1)} \) and

\[
N_\text{in} \sim dN_\text{in}/d\kappa \sim 1/\kappa^{\mu}, \ \mu = (4\nu - 1)/(2\nu - 1). \tag{6}
\]

For our case with \( \nu = 1 \) we have \( \mu = 3 \). This value is in a good agreement with the data of Fig. 3. For the first model \( dx/d\kappa \) is always finite and we have a sharp drop of ingoing links distribution.

This analysis allows understanding the origin of power law distributions of links in the Ulam networks generated by 1d maps.

FIG. 3. (Color online) Distribution of links between nodes of the Ulam network of \( N=10^6 \) size for the first model with \( z_2 = 2, \ z_2 = 0.2 \) (top panel) and the second model with \( z_1 = 2, \ a = 0.9 \) (bottom panel). Here, \( N_\kappa(\kappa) \) gives the number of nodes which have \( \kappa \) outgoing (black points) or ingoing (red/gray squares) links, respectively. Insets show data for small \( \kappa \) values in linear scale. The straight line shows the theoretical slope for outgoing links \( N_\kappa(\kappa) \propto 1/\kappa^{0.4} \) first model, top panel) and for ingoing links \( N_\kappa(\kappa) \propto 1/\kappa^3 \) second model, bottom panel).

FIG. 4. (Color online) Distribution of eigenvalues \( \lambda \) in the complex plain for the Google matrix at \( a = 1 \) for the first \( (z_1 = 2, \ z_2 = 0.2, \) left panel) and second \( (z_1 = 2, \ a = 0.9, \) right panel) models at \( N = 12000 \). Color (grayscale) of small squares is determined by the value of PAR \( \xi \) associated with the corresponding eigenvector \( \psi \) as shown in the palette (the values of \( \xi \) are averaged over the states inside of the square size).

FIG. 5. (Color online) Dependence of density of states \( W(\gamma) \) on \( \gamma \) shown for different values of \( N \) for the first \( (z_2 = 2, \ z_2 = 0.2, \) top panel) and second \( (z_1 = 2, \ a = 0.9, \) bottom panel) models; \( G \) matrix is taken at \( a = 1 \).
both maps $f_1(x)$ and $f_2(x)$. This means that it appears due to the left branch of the map corresponding to intermittent motion near $x=0$. The density distributions $W(\gamma)=dN_i/d\gamma$ in the decay rate defined as $\gamma=-2\ln|\lambda|$ are shown in Fig. 5 (here $dN_i$ is a number of states in the interval $d\gamma$). It is clear that in the limit of large matrix size $N$ we have a convergence to a limiting distribution, which has a characteristic peak at $\gamma=2$.

Examples of few eigenstates $\psi(i)$ with values of $\gamma_m=-2\ln|\lambda_m|$ equal and close to zero are shown in Fig. 6 (the index $1\leq i\leq N$ gives the cell position $x_i=(i-1)/N$, index $m$ orders $\gamma_m$ from zero to maximum $\gamma$). The first state $\psi_1(i)$ with $\lambda_1=1$ is the steady state distribution generated by the map $f_1(x)$ (the states for the map $f_2(x)$ have similar structure and we do not show them here). We have $\psi_1(i)\approx 1/i^\beta$ with $\beta=1$ for $z_1=2$ is agreement with the theoretical expression (4) (the numerical fit gives $\beta=0.97$). The state $\psi_1(i)$ is monotonic in $i$ so that it coincides with the PageRank $p_j$ up to a constant factor. Eigenstates with next values of $\gamma$ are characterized by the same decay at large $i$ with additional minima at certain values of $i$ similar to few nodes of eigenstates in quantum mechanics.

The structure of eigenstates is changed when the value of $\gamma$ is increased. Typical states are shown in Fig. 7. The states on the first circle of $|\lambda|$ have peaked structure at certain $i$ with a plateau at large $i$. For $\gamma$ values at the maximum of $W(\gamma)$ (see Fig. 5) the eigenstates are delocalized over the whole interval of $1\leq i\leq N$.

An effective number of sites contributing to an eigenstate can be characterized by the PAR $\xi$. For the PageRank the value of $\xi$ is independent of the matrix size $N$ as it is clearly shown in Fig. 8. This is due to the power law decay of the PageRank $p_j\sim 1/j$ which corresponds to an algebraic localization. The dependence of $\xi$ on $\gamma$ is shown in Fig. 9. For small $\gamma$ it can be fitted by a power-law growth $\xi\sim \gamma^{1.2}$. The origin of the exponent of this growth requires further analysis.

Finally we note that we also determined the dependence of number of states $N_j$ with values of $\gamma>5$ on the matrix size $N$. Our data (not shown) are well described by the dependence $N_j\sim N$ so that in contract to the results presented in [12] there are no singes of the fractal Weyl law. We attribute this to the fact that in contrast to the dissipative map with a global contraction studied in [12] in the intermittency maps all dynamics takes place on the whole one-dimensional interval with inhomogeneous distribution of measure but without fractality.
GOOGLE MATRIX AND ULAM NETWORKS OF...

FIG. 10. (Color online) Dependence of gap $\Delta_{12} = 1 - |\lambda_2|$ between the first eigenstate with $\lambda_1 = 1$ and next one with maximum $|\lambda_2|$ on $N$ for the first $f_1(x)$ ($z_1=2$, $z_2=0.2$) and second $f_2(x)$ ($z_1=2$, $a=0.9$) models. The straight dashed line shows the dependence $\Delta_{12} \approx 1/N$.

IV. PROPERTIES OF THE PAGERANK

The spectral gap between $\lambda_1 = 1$ equilibrium state and the next state with maximum $|\lambda_2|$ has very small gap $\Delta_{12} = 1 - |\lambda_2|$ which goes to zero with the increase of $N$ like $\Delta_{12} \approx 3/N$ (see Fig. 10). This happens due to the dynamical properties of the maps Eqs. (2) and (3) where the time spent at small $x - 1/N$ is of the order $t_\epsilon \sim 1/x^{1-1}$ (see, e.g., [24,32,33]), so that the corresponding $\Delta_{12} \sim 1/t_\epsilon \sim 1/N^{z_1-1}$ that gives the exponent 1 for $z_1=2$.

Due to such decrease of $\Delta_{12}$ with $N$ the PRA has bad convergence at $\alpha = 1$ for large values of $N$. Up to $N \sim 14000$ we use the direct diagonalization of $G$ matrix which gives an algebraic decay $p_j \sim 1/j^\beta$ with $\beta = 1$ (see Fig. 6). For larger value of $N$ we used the continuous map obtaining $p_j$ from an equilibrium distribution over the cells of size $1/N^2$ after a larger number of map iterations $t_\epsilon \sim 10^5$ and large number of trajectories $N_{tr} \sim 10^5$. This distribution converges to a limiting one at large values of $t_\epsilon$ (see Fig. 11). Both methods give the same result for $N < 2 \cdot 10^4$. The numerical data for the exponent $\beta$ are in good agreement with the theoretical dependence Eq. (4) $\beta = z_1 - 1$ as it is shown in Fig. 12 (we attribute small deviations from the theoretical values to finite size effects of $N$).

FIG. 11. (Color online) Dependence of the PageRank $p_j$ on $j$ for the first model at different values of $a$ ($z_1=2$, $z_2=0.2$, top panel) and different values of $z_1$ ($z_2=0.2$, $a=1$, bottom panel). The data are obtained from the continuous map (see text) for $a=1$ and the PageRank algorithm at $\alpha < 1$, the number of nodes is $N=10^5$.

FIG. 12. (Color online) Dependence of the PageRank exponent $\beta(p_j \sim 1/j^\beta)$ on $z_1$ for the first model at $z_2=0.2$ and $a=1$, $N=10^5$. The straight dotted line shows the fit $\beta = 1.042 z_1 - 1.107$.

For $\alpha < 1$ the PRA, described in the Introduction, is stable and converges rapidly to the PageRank. It gives the same results as the exact diagonalization for $N < 2 \cdot 10^4$. The dependence of PageRank on $\alpha$ is shown in Fig. 11 (top panel). A small decrease down to $\alpha=0.999$ modifies $p_j$ at $j < 100$ making $p_j$ very flat in this region. For $\alpha = 0.875$ the PageRank becomes completely delocalized over the whole system size $N$.

For the second model the PageRank depends strongly on the value of $\alpha$. For $\alpha < 0.945$ when the dynamics is chaotic and the steady-state distribution is given by Eq. (4) the properties of the PageRank are similar to those of the first model described above, e.g., we have $\beta = 1$ being independent of $a$ for $\alpha = 1$, $z_1 = 2$ (see Fig. 13, bottom panel). However, for $\alpha > 0.945$ the map has a fixed point attractor and the PageRank becomes localized practically on one site at $\alpha = 1$. In this regime with fixed point attractor, the PageRank is very sensitive to $\alpha$ variation: at $\alpha < 1$ we have $p_j \sim 1/j^\beta$ with the fit values $\beta = 0.79$ at $\alpha = 0.98$, $\beta = 0.60$ at $\alpha = 0.9$, $\beta = 0.42$ at $\alpha = 0.8$ and $\beta = 0.32$ at $\alpha = 0.7$.

The delocalization of the PageRank from the fixed point attractor state is also clearly seen in the variation of PAR $\xi(a, \alpha)$ shown in Fig. 14. This shows that even if at $\alpha = 1$ the...
PageRank is dominated only by one node a decrease of $\alpha$ allows obtaining weighted contribution of other nodes.

We also note that in the phase of fixed point attractor the spectrum of eigenvalues $\lambda$ has globally a structure rather similar to one at $\alpha=0.9<0.945$ (see Fig. 4, right panel). However, the PAR values of all eigenstates at $\alpha>0.945$ become rather close to unity showing that almost all eigenstates are strongly localized in this phase. For example, for $\alpha=0.96$, we have almost all $\xi$ in the range from 1 to 4 for $N=10^4$, it is interesting that about 53% of the states have $|\lambda|<e^{-10}$ (for $\alpha=0.9$ this circle in $\lambda$ contains 23% of states, see Fig. 4 right panel).

V. DISCUSSION

The present studies allowed establishing a number of interesting properties of the Google matrix constructed for the Ulam network generated by intermittency maps. A general property of such networks is the existence of states with eigenvalues $|\lambda|$ being very close to unity. The PageRank of such networks at $\alpha=1$ is characterized by a power law decay with an exponent determined by the parameters of the map. It is interesting to note that usually for WWW it is observed that the decay of the PageRank follows the decay law of ingoing links distribution $N_L^\alpha(x)$ (see, e.g., [21]). In our case the decay of PageRank is independent of $N_L^\alpha(x)$ decay as it is clearly shown by Eqs. (5) and (6) and the data of Figs. 3, 11, and 13. In fact a map with singularities of both maps $f_1(x)$ and $f_2(x)$ (e.g., $f_3(x)$ which behaves like $x+x^{11}$ at small $x$, like $(1/2-x)^{11}$ at $x<1/2$ close to 1/2 and like $(1-x)^{6}$ near $x=1$) will have the asymptotic decay of links distribution given by Eqs. (5) and (6) but the decay of the PageRank will be given by $\beta=z_1-1$, hence, being independent of the decay of links distribution.

Our results also show that while at $\alpha$ close to unity the decay of the PageRank has the exponent $\beta=1$ but at smaller values $\alpha\approx 0.9$ the PageRank becomes completely delocalized (see Fig. 11). In this delocalized phase the PAR $\xi$ grows with the system size approximately as $\xi\approx N$. The delocalization of the PageRank can also take place at $\alpha=1$ due to variation of the parameters of the map (e.g., for $z_1\rightarrow 1$). It is rather clear that the delocalization of the PageRank makes the Google search inefficient.

We hope that the properties of Ulam networks generated by simple maps will be useful for future studies of real directed networks including WWW. Indeed, the whole world will go blind if one day the Google search will become inefficient. The investigations of the Ulam networks can help to understand the properties of directed networks in a better way that can help to prevent such a dangerous situation.

ACKNOWLEDGMENT

We thank A. S. Pikovsky for a useful discussion of his results presented in [27].


