Random Walk Equivalence to the Compressible Baker Map and the Kaplan-Yorke Approximation to Its Information Dimension

William Graham Hoover
Corresponding Author email : hooverwilliam@yahoo.com
Ruby Valley Research Institute
Highway Contract 60, Box 601, Ruby Valley, Nevada 89833, USA ;

and Carol Griswold Hoover
Ruby Valley Research Institute
Highway Contract 60, Box 601, Ruby Valley, Nevada 89833, USA ;

(Dated: September 23, 2019)

Abstract

Hoover and Posch studied a single deterministic compressible time-reversible version of Eberhard Hopf's "Baker Map". This particular version was selected to serve as a "Toy Model" of nonequilibrium molecular dynamics. Iterating the map generates an ergodic fractal distribution with measure everywhere in a square domain. Two versions of the map are isomorphic to a third model, a one-dimensional random walk confined within the unit interval $(0 < y < 1)$ with leftward steps twice as likely as rightward. This correspondence between two-dimensional time-reversible deterministic (q, p) and (x, y) maps and the stochastic bounded one-dimensional y-direction walk is used here to characterize discrepancies between three different formulations of the fractal information dimension D_I of the compressible Baker Map.

We compare the Kaplan-Yorke conjecture, believed valid for these models, $D_I \stackrel{?}{=} D_{KY}$, to brute-force computations of either a few, or as many as a trillion iterations of the one- and two-dimensional Maps. The (q, p) Baker Mapping contains square roots. These irrational numbers provide sufficient detail, or "noise", to avoid the periodic orbits that would otherwise short-circuit the convergence of the Baker Map to the steady state solutions found here. We find that the Kaplan-Yorke Lyapunov dimension, D_{KY}, accepted as correct for simple maps like ours based on the mapping of smooth density functions, differs from the simplest of several information dimension estimates. The latter estimate is based on repeated mappings of an initial point, as measured on a series of meshes ranging from $(1/3)$ in width down to $(1/3)^{19}$. Other families of meshes support the conventional wisdom that the small-mesh limiting information dimension is given *exactly* by the 40-year-old Kaplan-Yorke "conjecture".

Keywords: Random Walk, Fractal, Baker Map, Information Dimension

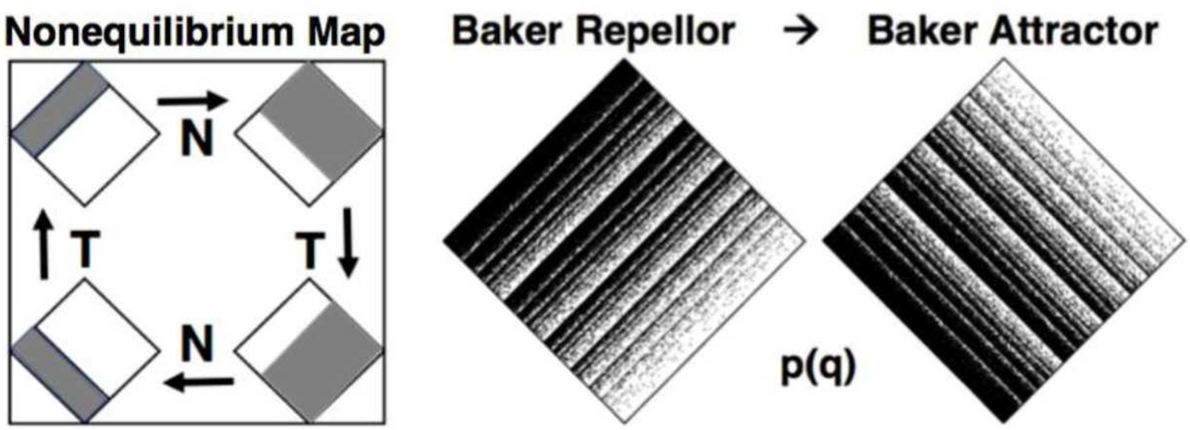

FIG. 1: The nonequilibrium Baker Map "**N**" maps the white southeast region with two-thirds the area, at the far upper left, to the adjacent white southwest one-third the area region at the top right of the leftmost figure. Moving right, we see the flow *from* the repellor, at the center here, *to* the attractor at the right can only be time-reversed by storing and reversing a forward trajectory : $\mathbf{N^{-1}} \equiv \mathbf{TNT}$. The flows here are achieved by repeated mappings of a single point. The momentum-like vertical variable p has its sign changed by time reversal, $\mathbf{T}(\pm q, \pm p) = (\pm q, \mp p)$. The coordinate-like horizontal variable q is unchanged by reversal. The repellor has zero probability, with two-thirds of its (vanishing) measure located in its northwest third. The mirror-image attractor has unit probability, with two-thirds of it in the southwest third of the attractor image shown at the far right. The steady-state information dimensions, based on meshes of width $(1/3)^n$, of the fractals at the center and right are 1.741, as is discussed in the text. The Kaplan-Yorke Lyapunov dimension is only 1.7337 for this map. This discrepancy is a puzzle. For another puzzle see **Figure 2**.

I. THREE MODELS OF COMPRESSIBLE TIME-REVERSIBLE BAKER MAPS

Nonequilibrium computer simulations make use of deterministic time-reversible boundary regions so as to simulate shear flows and heat flows. The boundary regions are maintained at specified velocities and/or kinetic temperatures with feedback forces of the type pioneered by Shuichi Nosé in 1983[1,2] and clarified by Bill, Harald Posch, and Franz Vesely in 1984-1985[3,4]. By 1987 it became apparent that these time-reversible nonequilibrium systems generate fractal attractors in phase space, "strange attractors", where neighboring systems typically depart from one another exponentially fast, within a negligibly small, but ergodic and fractal, fraction of the occupied phase space[5–7]. A well-attended week-long Budapest workshop was influential in spreading these ideas world-wide[8] in 1997.

A. Information Dimension from Bin Probabilities

Grebogi, Ott, and Yorke[9] were followed by Dellago and Hoover[10] in showing that the length of the stable numerical periodic orbits in a coarse-grained digital phase space follows a simple stochastic model. Imagining that trajectories jump randomly from state to state reproduces the observed periodic orbit lengths for a variety of deterministic models. The orbit length varies as the square root of the number of coarse-grained states, $(1/\delta)^{D_c/2}$, where D_C is the correlation dimension and δ is the grain size of the states. The correlation dimension describes the power-law variation of separated attractor points in the vicinity of another attractor point from a relatively distant time. Typical basins of attraction of these sparse computational fractals have the dimensionality of the full phase space[9,10]. Typically these fractals, though sparse with zero volume, are also ergodic, with some nonvanishing measure throughout their domains. For that reason we use the information dimension of their measure as a way of distinguishing ergodic fractals with some measure everywhere. The information dimension is the small-mesh limit, $\delta \rightarrow 0$, of

$$D_I(\delta) = \sum_{bins} P \ln P / \ln \delta \text{ with } \sum_{bins} P \equiv 1 \ .$$

Because the probability P of a D-dimensional volume element is proportional to δ^D, $P \propto \delta^D$.

The specific map we consider here is the compressible, time-reversible, deterministic Baker Map of **Figure 1**. It was developed by Posch and Hoover, and analyzed further by Kumičák[11,12]. It serves as a Toy Model for the generic steady states generated by nonequilibrium molecular dynamics. To mimic molecular dynamics we sought a map which was not only deterministic but also time-reversible. This is the (q, p) Baker Map.

We describe three different versions of these Baker Map motion equations, and comment on their time reversibility, next. The three versions are [1] the basic time-reversible (q, p) map, [2] a more conventional but equivalent Cartesian (x, y) map, which lacks reversibility, and [3] a one-dimensional stochastic random walk on the interval $(0 < y < 1)$ which captures the chaotic motion of its two-dimensional deterministic relatives.

B. Time Reversibility and the (q, p) Baker Map

A map **M** is said to be time-reversible when it can be reversed by a three-step process: [1] changing the signs of the momentum-like variables, [2] propagating all the variables one ("backward") iteration, and then changing the signs of the momenta once more, so that the inverse of the map $\mathbf{M}$ is given by $\mathbf{M^{-1}} = \mathbf{TMT}$. In ordinary Hamiltonian mechanics the $\mathbf{T}$ mapping simply maps $(\pm q, \pm p) \rightarrow (\pm q, \mp p)$. Bill's conversations with Bill Vance during Vance's graduate work at the University of California's Davis campus led us to a nonequilibrium *rotated* version of the Baker Map $\mathbf{B}$ which we call $\mathbf{N}$, for "Nonequilibrium". This Map's domain is the diamond-shaped region, centered on $(q,p) = (0,0)$ and shown at the left of **Figure 1**. Now imagine that the map $\mathbf{N}$ is applied to a representative input point (q,p). This operation produces the next point (q',p').

Our nonequilibrium Map, $\mathbf{N}(q,p) \rightarrow (q',p')$ has the following form : For twofold expansion, $q < p - \sqrt{2/9}$:

$$q' = (11q/6) - (7p/6) + \sqrt{49/18} \ ; \ p' = (11p/6) - (7q/6) - \sqrt{25/18} \ .$$

For twofold contraction, $q > p - \sqrt{2/9}$:

$$q' = (11q/12) - (7p/12) - \sqrt{49/72} \ ; \ p' = (11p/12) - (7q/12) - \sqrt{1/72} \ .$$

Figure 1 shows the resulting concentration of probability into bands parallel to the bottom left and upper right edges of the diamond.

In conventional Cartesian coordinates, with $0 < x, y < 1$, a topologically equivalent map $\mathbf{B}$ is

$$x > (1/3) \rightarrow x' = (3x - 1)/2 \ ; \ y' = (y/3) \ ;$$

$$x < (1/3) \rightarrow x' = 3x \ ; \ y' = (1 + 2y)/3 \ .$$

Although the algebra is more cumbersome we have chosen to use the rotated (q,p) version of this map, centered on the origin and confined to a diamond-shaped region of sidelength 2, as shown in **Figure 1**. We regard the horizontal q variable as a coordinate and the vertical p variable as a momentum. The Figure illustrates the time-reversibility of the (q,p) map. This similarity to nonequilibrium molecular dynamics, along with the square roots generating the 45^o rotation, are the twin advantages of the nonequilibrium diamond-shaped map $\mathbf{N}$. The square roots eliminate many of the artificial periodic orbits due to finite computer precision. Beginning at the center point of the rational-number square map, $(x,y) = (0.5, 0.5)$, leads to a periodic orbit of just 3096 single-precision iterations. Starting instead at the equivalent central point of the irrational-number diamond map, $(q,p) = (0.0, 0.0)$, leads to a single-precision periodic orbit of 1,124,069 iterations. With double-precision arithmetic the orbits are much longer, 146,321,810 for the (x, y) map. 10^{12} such double-precision (q, p) iterations from the equivalent initial condition gave no repeated points.

A single one-dimensional Baker-Map mapping of a uniform distribution of many points (millions or billions) on the interval $(0 < y < 1)$ puts 2/3 of the measure into the lefthand interval of width 1/3. The remaining 1/3 of the measure is mapped into the remaining interval of length 2/3. **Figure 2** illlustrates the iterated operation of the Baker Map for 1, 2, 3, and 4 iterations applied to an initially uniform distribution. For simplicity here we have projected the result of the mapping onto the unit interval in y rather than the 2×2 diamond or unit square. The singly-mapped measure corresponds to measures of (2/3) and (1/6) and (1/6) in three equal-width intervals, and so to an approximate one-step information dimension after a single iteration of many uniformly-dense points :

$$D_I(1) = [\ (2/3)\ln(2/3) + (1/6)\ln(1/6) + (1/6)\ln(1/6)\ ]/\ln(1/3) = 0.78969\ .$$

Although it is a surprise to find that the same information dimension results for 2 or 3 or 4 iterations, that result is fully consistent with the scale-model nature of the distribution, as shown in **Figure 2**. Evidently the mapping of a uniform distribution thus gives results quite different to those from the iteration of a single point (or a delta function).

If we do follow the alternative many-iteration history of a typical point, iterated 20 billion times we find probabilities for the three intervals corresponding to a somewhat reduced information dimension along with a substantial difference in the populations of the middle and rightmost intervals. Here is the accurate three-bin approximation using 20 billion iterations:

$$D_I \simeq [\ (2/3)\ln(2/3) + 0.29516\ln(0.29516) + 0.03818\ln(0.03818)\ ]/\ln(1/3) = 0.6874\ .$$

The vanishing bin-size limit, (with many iterations per identically-sized bin) of this procedure, entropy, divided by the logarithm of the mesh size evidently approaches $D_I = 0.741$ if bins of size $(1/3)^n$ are used[13].

The data of **Figure 2** show that iterating the map twice gives nine occupation probabilities for nine equal-sized bins,

$$\{\ (4/9),\ (1/9),\ (1/9),\ (1/9),\ (1/9),\ (1/36),\ (1/36),\ (1/36),\ (1/36)\ \}\ ,$$

for the nine strips of width (1/9) parallel to the x axis, and so to D_I :

$$D_I(2) = [\ (4/9)\ln(4/9) + (4/9)\ln(1/9) + (1/9)\ln(1/36)\ ]/\ln(1/9) = 0.78969\ [\text{ Again ! }]\ .$$

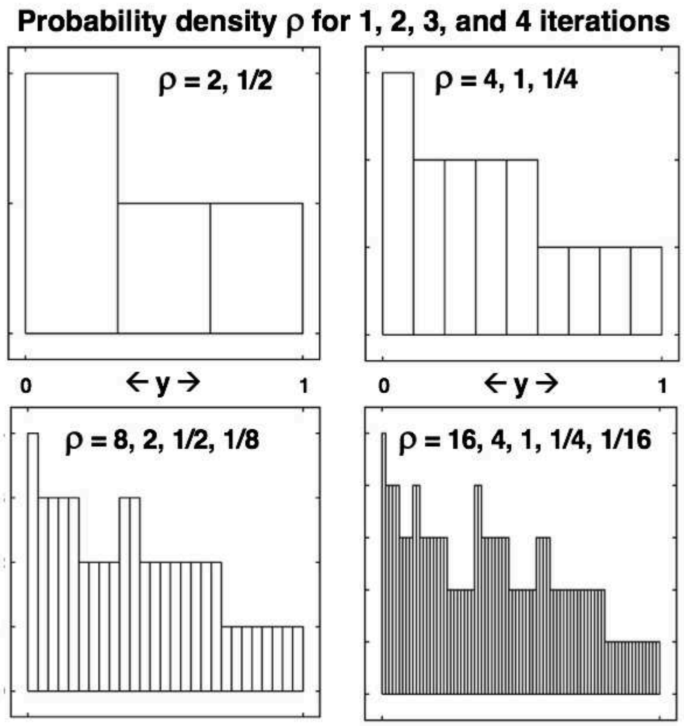

FIG. 2: Histograms of the (base-4 logarithm of) probability density $\rho(y)$ for 1, 2, 3, and 4 iterations of the y component of the Baker Map **B**. Notice that the rightmost (2/3) of each mapping is a perfect scale model of the leftmost (1/3). The information dimension of all these iterates is the same, $\sum P\ln(P)/\ln(\delta) = 0.78969$! This puzzling result differs from both the Kaplan-Yorke value of 0.7337 and the extrapolated 0.741 based on mesh sizes of $(1/3)^n$ and illustrated in **Figure 3**.

The same information dimension results with 3 or 4 or 5 or ... iterations of the map and 3^3 or 3^4 or 3^5 or ... equal-sized bins. Though suggestive, this result does not at all indicate that the information dimension of the pointwise *stationary solution* $D_I(\infty)$ *is* 0.78969. In fact for n bins of width $(1/3)^n$ the information dimension appears to approach 0.741 for large n, as we shall soon see.

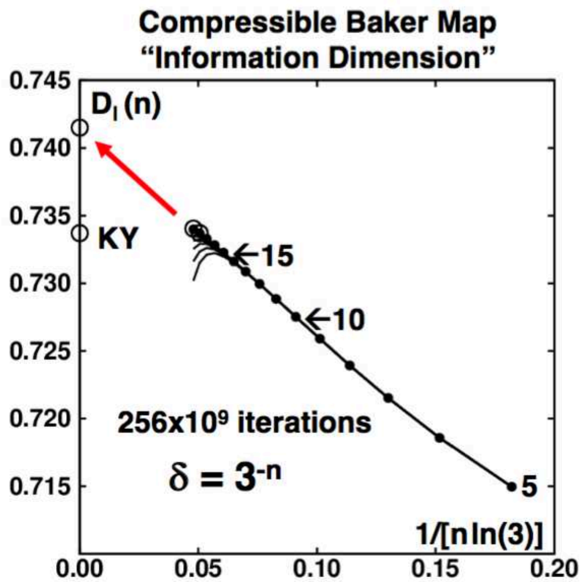

FIG. 3: Stationary estimates of D_I for the random-walk model of the Baker Map with results for $3^{5,10,15}$ equal bins emphasized. The lines join points computed with 1, 2, 4, 8 . . . 256 billion iterations of the random-walk version of the Baker Map. The two values shown at the zero bin-size limit ($\delta \to 0$) are the Kaplan-Yorke dimension, 0.7337, and a plausible extrapolation of two trillion-iteration computations with as many as 3^{19} bins. Note the qualitative difference of the mesh dependence (the slope is uniformly negative here) compared to those shown in **Figures 4 and 5**. The two open circles at $n = 18$ and 19 correspond to $2^{10} \times 10^9$ iterations. The directions of the individual random-walk steps were based on the FORTRAN subroutine `Random_Number(r)`.

II. A STATIONARY INFORMATION DIMENSION $D_I(\infty)$

It is quite practical to iterate the two-dimensional (q, p) map "**N**" as many as a trillion times, but for less than a still-practical $3^{19} = 1,162,261,467$ bins the information dimension converges relatively rapidly. Fewer iterations are needed. For a million iterations, 10^6 rather than 10^{12}, and $3^5 = 243$ bins the information dimension is 0.71447. For ten million 0.71488. For 100 million 0.71498 and for one billion 0.71501, which we would guess is within 0.00001 of the limiting value. The reason for the dimension different to 0.78969 stems from the latter's unrealistic initial condition, uniform coverage of the square domain. A billion iterations of a single point are sufficient to obtain the correct dimensionality within 0.0001 or 0.00001, as would be expected from the Central Limit Theorem.

Let us refine the mesh to $3^{10} = 59,049$ bins. The estimates of the information dimension

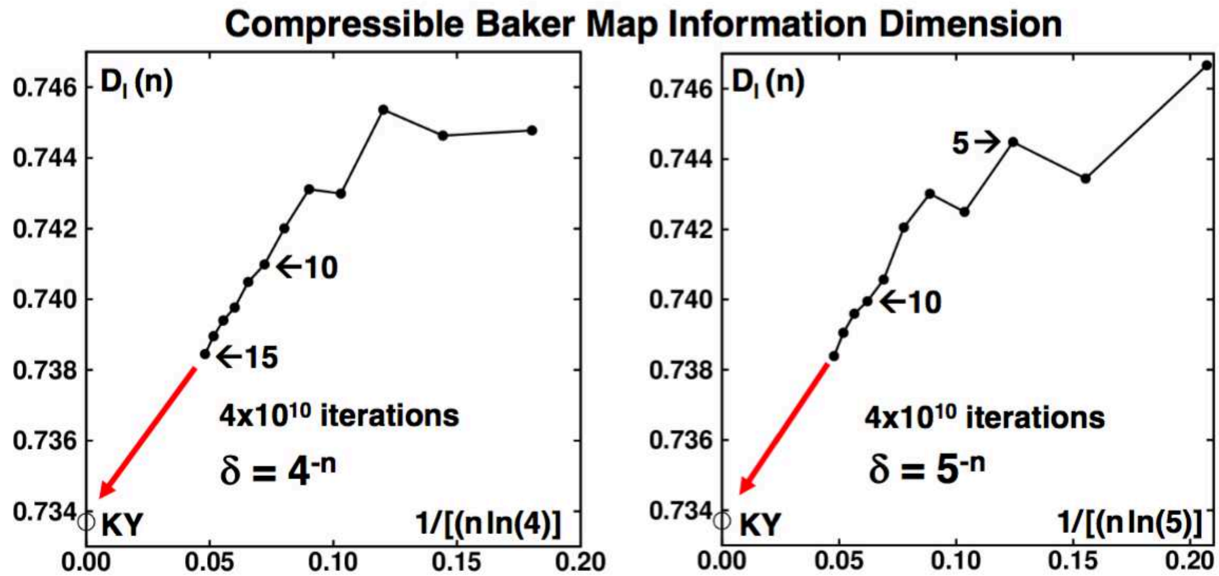

FIG. 4: Stationary estimates for the Baker Map Information Dimension using up to 4^{15} and 5^{13} bins of equal width. The rightmost data, based on forty billion iterations of the random walk mapping suggest agreement with the Kaplan-Yorke dimension of the (q, p) and (x, y) Baker Maps, $D_{KY} = 0.7337$.

for the four choices of iteration now give 0.72525, 0.72725, 0.72753, 0.72757. A geometric series suggests 0.72758 as a good estimate for 3^{10} bins. A mesh with $3^{15} = 14,348,907$ bins provides the four estimates 0.70014, 0.72378, 0.73059, and 0.73205, suggesting adding 10 and 100 billion and a trillion iterations to our statistics, giving 0.73223, 0.73226_4, and 0.73226_7, confirming $D_I = 0.73227$ as accurate for 3^{15} bins. With the knowledge that the information dimension depends linearly on $1/\ln(\delta)$ the data using $3^{5,6,\ldots,19}$ give an estimate $D_I = 0.741$. Let us compare this to the Kaplan-Yorke estimate from the Lyapunov exponents of the map. For a preview see **Figure 3** which shows the discrepancy between the data and the estimate to be about one percent.

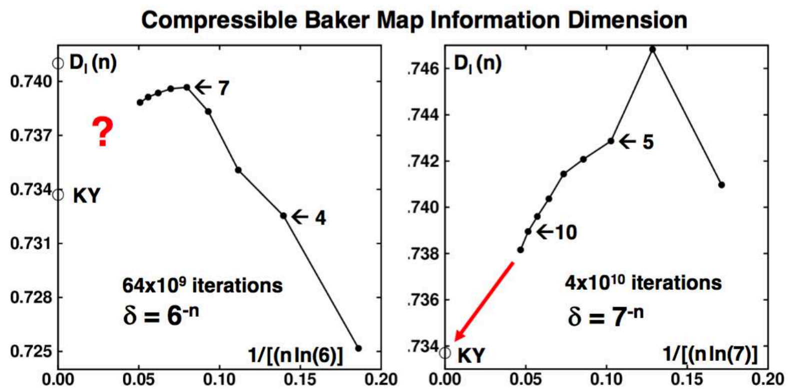

FIG. 5: Stationary estimates for the Baker Map Information Dimension using up to 6^{11} and 7^{11} bins of equal width. These data, like those in **Figure 4**, are based on forty billion iterations of the random walk mapping. Both the Kaplan-Yorke dimension 0.7337 and the estimate 0.741 from Reference 13, based on meshes with up to 3^{19} equal bins, are shown as open circles at the left border of the $\delta = 6^{-n}$ plot.

III. KAPLAN-YORKE OR "LYAPUNOV DIMENSION"

Kaplan and Yorke suggested that a linear interpolation formula between the last positive sum of exponents, starting with the largest, λ_1, and the first negative sum, $0 > \sum_{i \geq 1} \lambda_i$ would be a useful prediction for the information dimension[14,15] . In fact they cite many a case, including theoretical work carried out by L. S. Young, for which their conjectured estimate is exactly correct. It should be noted that the "heuristic proof" on page 162 of Reference 15 does not apply to ergodic attractors.

The white portion of the compressible Baker Map of **Figure 1** causes (2/3) of the measure to stretch by a factor (3/2) while the black portion causes (1/3) of the measure to stretch by a factor of 3. As a result

$$\lambda_1 = (2/3)\ln(3/2) + (1/3)\ln(3) = (1/3)\ln(27/4) = 0.63651 \ .$$

Likewise (2/3) of the measure shrinks by a factor 3 while (1/3) by a factor (2/3) so that

$$\lambda_2 = (2/3)\ln(1/3) + (1/3)\ln(2/3) = (1/3)\ln(2/27) = -0.86756 \ .$$

The linear interpolation gives $0.63651/0.86756 = 0.73368$. This Lyapunov dimension *is* the Kaplan-Yorke dimension.

IV. CONCLUSIONS AND DISCUSSION

Relatively simple numerical work, on the order of a few dozen lines of FORTRAN, and a few hours of laptop time, are enough to characterize the variety of results for D_I based on [1] iterating distributions or [2] generating representative series of points. These two different views of fractal structure are analogs of the Liouville and trajectory descriptions of particle mechanics. We think the singular anisotropy of fractals favors the pointwise approach. We found that pointwise analysis with the mesh series $(1/3)^n$ appears to contradict the Kaplan-Yorke dimension while the alternative series $(1/4)^n, (1/5)^n, (1/7)^n$ support it. The series $(1/6)^n$ is inconclusive.

Though the one-dimensional bounded random walk provides a fractal distribution $\{y\}$ indistinguishable from that for the compressible Baker Map, the walk analog lacks the Baker-Map Lyapunov exponents on which the Kaplan-Yorke dimension relies :

$$\lambda_1 = (1/3)\ln(27/4)\ ;\ \lambda_2 = (1/3)\ln(2/27) \rightarrow D_{KY} = 0.73368\ .$$

The variety of results obtained here for a specific map underline the value of studying particular, as opposed to general, models.

V. ACKNOWLEDGEMENT

Carl Dettmann, Thomas Gilbert, Harald Posch, Ed Ott, and James Yorke kindly provided advice and references helpful to the completion of this work. Carl's, Thomas', and James' emails eventually catalyzed our realization that mapping points is qualitatively different to mapping smooth densities. Surprisingly, it took us 20 years to come to this understanding. Perseverance and the stimulation of one's colleagues pays off. We thank them all. This project suggests a subject area for the upcoming 2020 Ian Snook Prize.

---

[1] S. Nosé, "A Molecular Dynamics Method for Simulations in the Canonical Ensemble", Molecular Physics **52**, 255-268 (1984).

[2] S. Nosé, "A Unified Formulation of the Constant Temperature Molecular Dynamics Methods", Journal of Chemical Physics **81**, 511-519 (1984).

[3] W. G. Hoover, "Canonical Dynamics: Equilibrium Phase-Space Distributions", Physical Review A **31**, 1695-1697 (1985).

[4] H. A. Posch, W. G. Hoover, and F. J. Vesely, "Canonical Dynamics of the Nosé Oscillator: Stability, Order, and Chaos", Physical Review A **33**, 4253-4265 (1986).

[5] B. Moran, W. G. Hoover, and S. Bestiale, "Diffusion in a Periodic Lorentz Gas", Journal of Statistical Physics **48**, 709-726 (1987).

[6] B. L. Holian, W. G. Hoover, and H. A. Posch, "Resolution of Loschmidt's Paradox: The Origin of Irreversible Behavior in Reversible Atomistic Dynamics", Physical Review Letters **59**, 10-13 (1987).

[7] W. G. Hoover, H. A. Posch, B. L. Holian, M. J. Gillan, M. Mareschal, and C. Massobrio, "Dissipative Irreversibility from Nosé's Reversible Mechanics", Molecular Simulation, **1**, 79-86 (1987).

[8] Proceedings of "Chaos and Irreversibility" at Eötvös University 31 August- 6 September, organized by T. Tél, P. Gaspard, and G. Nicolis, Chaos **8**, 309-461 (1998).

[9] C. Grebogi, E. Ott, and J. A. Yorke, "Roundoff-Induced Periodicity and the Correlation Dimension of Chaotic Attractors", Physical Review A **38**, 3688-3692 (1988).

[10] C. Dellago and Wm. G. Hoover, "Finite-Precision Stationary States At and Away from Equilibrium", Physical Review E **62**, 6275-6281 (2000). See the references to previous 1998 work of Grebogi, Lanford, Ott, and Yorke therein.

[11] Wm. G. Hoover and H. A. Posch, "Chaos and Irreversibility in Simple Model Systems", Chaos **8**, 366-373 (1998).

[12] J. Kumiĉák, "Irreversibility in a Simple Reversible Model", Physical Review E **71**, 016115 (2005).

[13] W. G. Hoover and C. G. Hoover, "Aspects of Dynamical Simulations, Emphasizing Nosé and Nosé-Hoover Dynamics and the Compressible Baker Map", Computational Methods in Science and Technology (in press, 2019).

[14] J. L. Kaplan and J. A. Yorke, "Chaotic Behavior of Multidimensional Difference Equations", pages 204-227 in *Functional Differential Equations and the Approximation of Fixed Points*, edited by H. O. Peitgen and H. O. Walther (Springer, Berlin, 1979).

[15] J. D. Farmer, E. Ott, and J. A. Yorke, "The Dimension of Chaotic Attractors", Physics **7** D, 153-180 (1983).