Manifold learning approach for chaos in the dripping faucet

Hiromichi Suetani,[1,2,3] Karin Soejima,[1] Rei Matsuoka,[4] Ulrich Parlitz,[5,6] and Hiroki Hata[1]

[1]*Department of Physics and Astronomy, Kagoshima University, Kagoshima 890-0065, Japan*
[2]*Decoding and Controlling Brain Information, Precursory Research for Embryonic Science and Technology,*
Japan Science and Technology Agency, Kawaguchi 332-0012, Japan
[3]*Flucto-Order Functions Research Team, RIKEN-HYU Collaboration Research Center, RIKEN Advanced Science Institute,*
Wako 351-0198, Japan
[4]*Department of Energy Engineering and Science, Nagoya University, Nagoya 464-8603, Japan*
[5]*Biomedical Physics Group, Max Planck Institute for Dynamics and Self-Organization, Am Faßberg 17, 37077 Göttingen, Germany*
[6]*Institute for Nonlinear Dynamics, Georg-August-Universität Göttingen, Am Faßberg 17, 37077 Göttingen, Germany*

Dripping water from a faucet is a typical example exhibiting rich nonlinear phenomena. For such a system, the time stamps at which water drops separate from the faucet can be directly observed in real experiments, and the time series of intervals τ_n between drop separations becomes a subject of analysis. Even if the mass m_n of a drop at the onset of the nth separation, which is difficult to observe experimentally, exhibits perfectly deterministic dynamics, it may be difficult to obtain the same information about the underlying dynamics from the time series τ_n. This is because the return plot τ_{n-1} vs. τ_n may become a multivalued relation (i.e., it doesn't represent a function describing deterministic dynamics). In this paper, we propose a method to construct a nonlinear coordinate which provides a "surrogate" of the internal state m_n from the time series of τ_n. Here, a key of the proposed approach is to use ISOMAP, which is a well-known method of manifold learning. We first apply it to the time series of τ_n generated from the numerical simulation of a phenomenological mass-spring model for the dripping faucet system. It is shown that a clear one-dimensional map is obtained by the proposed approach, whose characteristic quantities such as the Lyapunov exponent, the topological entropy, and the time correlation function coincide with the original dripping faucet system. Furthermore, we also analyze data obtained from real dripping faucet experiments, which also provide promising results.

I. INTRODUCTION

Dripping of water from a faucet is ordinarily seen in our daily life. At first glance, such a motion of dripping looks very common. It provides, however, a variety of rich nonlinear dynamics including the period-doubling bifurcation to chaos, intermittency, crisis, hysteresis, etc. In particular, Shaw and his collaborators [1] first found that there is a clear transition from a periodic motion to low-dimensional chaos by investigating the time intervals τ_n between dripping separations from the faucet, both theoretically and experimentally.

The dripping water is fluid dynamics (i.e., ideally described as an infinite-dimensional dynamical system). But as far as the dynamics is confined within a low-dimensional attractor, it can be modeled by a class of phenomenological models called "mass-spring" systems [1]. Since the pioneering work of Shaw *et al.*, many versions of the mass-spring system for the dripping faucet have been proposed. Among them, Kiyono and Fuchikami [2] significantly improved the mass-spring system on the basis of both numerical simulations of fluid dynamics [3] and real experiments [4]. They showed that their model can systematically explain various aspects of the complex behaviors observed in the real dripping faucet experiments.

One of the most prominent aspects of the Kiyono-Fuchikami model is that the essential feature of chaos in the dripping faucet is exactly represented as a one-dimensional map. More precisely, the mass m_n at the moment of the nth separation of a drop from the faucet obeys a one-dimensional mapping dynamical system [i.e., there exists a deterministic scalar function $f(\cdot)$ such that $m_n = f(m_{n-1})$].

In general, however, not all state variables are observable in real experiments. In the case of the dripping faucet system, a high precision balance with resolution of about 0.1 mg for every 0.1 sec is needed to obtain the time series of the mass m_n of drops. Instead, time intervals τ_n between successive drop separations can be recorded more easily in real experiments. As investigated by Shaw *et al.*, depending on the degree of flux of water, the return plot τ_{n-1} vs τ_n also shows a clear functional relationship. At the same time, however, they have also shown that it often takes the form of a multivalued relation. Namely, there are two or more candidates of τ_n against a single value of τ_{n-1}, which prevents us from interpreting the dripping faucet as a simple one-dimensional mapping system. This multivaluedness problem often occurs in general chaotic dynamical systems such as the Kuramoto-Sivashinsky equation [5].

On the other hand, the existence of the one-dimensional map f associated with the mass m_n means that an embedding of the dripping-time interval τ_n into a sufficiently high, say d-dimensional Euclidean space $\mathbb{R}^d$ as $\mathbf{s}_n = (\tau_{n-d+1}, \dots, \tau_{n-1}, \tau_n)$, is lying on a one-dimensional manifold $\mathcal{S} \subset \mathbb{R}^d$. Then, a point $\mathbf{s}_n \in \mathcal{S}$ obeys a deterministic law as $\mathbf{s}_n = \mathbb{F}(\mathbf{s}_{n-1})$ where $\mathbb{F}(\cdot)$ is a d-dimensional vector valued function whereas the relationship between τ_{n-1} and τ_n is a multivalued one. Actually, in the case of the Kiyono and Fuchikami's mass-spring model, embedding τ_n into a three-dimensional space generally results in a filamentlike one-dimensional manifold without crossing. Therefore, if a new coordinate u is spanned along $\mathcal{T}$, which plays the role of a surrogate variable for the internal state m_n, then we obtain

a more simplified expression as $u_n = g(u_{n-1})$ where $g(\cdot)$ is a scalar function of u. Even if the original one-dimensional mapping system $m_n = f(m_{n-1})$ is not available, important dynamical features can be obtained from the mapping associated with the surrogate variable.

To identify lower-dimensional representations of the dynamics, dimension reduction methods can be employed. Dimension reduction is an important task of data (pre)processing with applications in pattern and speech recognition, image processing, bioinformatics, psychology, etc. Linear subspaces containing or approximating the available data can be identified using principal component analysis (PCA), independent component analysis (ICA), and many useful methods in various fields [6]. However, when the data set of interest is located on or close to a (sub)manifold with significant curvature, the applicability of these linear methods is limited and nonlinear dimension reduction methods have to be employed.

Recently, in the field of statistical machine learning, methods of *manifold learning* have been developed for providing a low-dimensional representation when data is lying on a nonlinear low-dimensional manifold embedded in a high-dimensional Euclidean space. A number of methods have been proposed and in the present study we employ ISOMAP [7] for such a purpose. ISOMAP, which is an abbreviation of the term "isometric feature mapping," is a method of manifold learning where the geodesics between training samples are employed as the dissimilarity information in multidimensional scaling (MDS) [8].

In this paper, we demonstrate that ISOMAP is very useful to extract a surrogate state variable u and to construct a well-defined one-dimensional map $g(\cdot)$ for both numerical and real experimental data. It is shown that dynamical characteristics such as the Lyapunov exponent and the time correlation function can be computed from $g(\cdot)$.

The present paper is organized as follows. In Sec. II, we explain the dripping faucet system. In Sec. III, we first introduce the method of ISOMAP, then we apply it to data generated from the mass–spring model mentioned in the previous section. Finally, in Sec. IV, we give a summary and discuss possible directions of future research.

II. DRIPPING FAUCET SYSTEM: MODEL AND EXPERIMENT

A. Basic mechanism

Let us begin with a brief introduction of the basic mechanism how a water drop separates from a faucet. Figure 1(a) shows an apparatus for dripping faucet experiments that we prepared. Using such an experimental apparatus, we take a snapshot at just the moment when a water drop separates from a thin plastic pipe as shown in Fig. 1(b). Here, the shape of the drop is determined by the balance between the surface tension and the weight of water. When increasing the mass of a drop by injecting water, the following processes are repeated with time. (i) A "neck" which connects between the drop and the faucet is formed by the break of the balance between the tension and the mass of water. (ii) When the weight reaches a critical value, the neck is broken (i.e., a portion of the drop separates from the faucet). (iii) Just after its separation, the
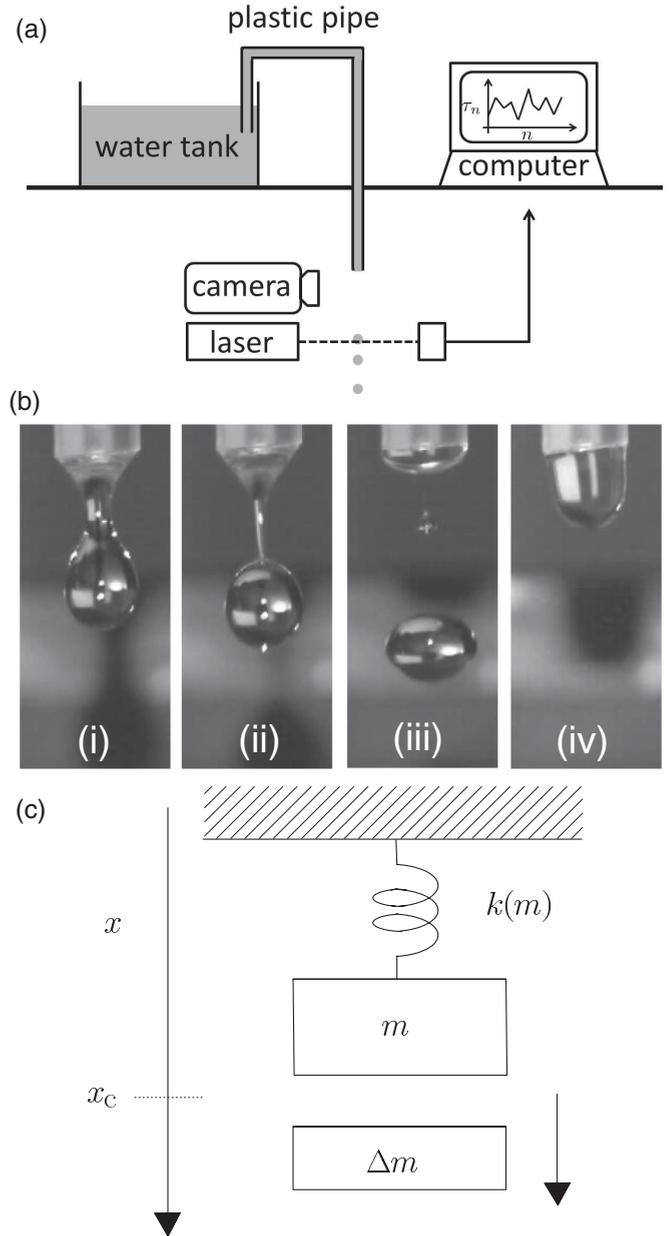
FIG. 1. (a) Schematic illustration of the apparatus for the dripping faucet experiment. (b) Snapshots of a high-speed movie showing the separation of a drop. (c) Illustration of the mass-spring model.

remainder of the drop rapidly shrinks by the surface tension to the upward direction. (iv) Finally, the drop grows again with oscillations.

B. Mass-spring model

Based on observations as mentioned in the previous subsection, the following equations of motion can be considered as a phenomenological model for the dripping faucet experiment [1]

$$m\ddot{x} + \dot{x}\dot{m} = -kx - \gamma\dot{x} + mg, \quad (1)$$

$$\dot{m} = Q \text{ (const.)}. \quad (2)$$

Such a model is called a mass-spring model and its schematic illustration is depicted in Fig. 1(c). Here, x is the vertical position of the forming drop to the downward direction, m is its mass, g is the gravitational acceleration, k is the stiffness of the spring, and γ is the damping parameter. The restoring force given by the surface tension is represented as a spring force in Eq. (1), and the mass of the forming drop linearly depends on time with the rate Q as described in Eq. (2), because a drop grows with time due to the influx of water from the faucet. It is also assumed that when the position of the drop reaches the critical point x_c, the drop loses its mass by Δm due to the separation of a portion of the drop and this portion falls to the ground. In spite of its simplicity, this model can explain many dynamical aspects of the real dripping faucet [1].

Kiyono and Fuchikami improved the above phenomenological model [2] based on the knowledge of the numerical simulations of fluid dynamics [3] and real experiments of the dripping faucet [4]. They first modified the equation of motion in Eq. (1) as

$$m\ddot{x} + (\dot{x} - v_0)\dot{m} = -kx - \gamma\dot{x} + mg, \qquad (3)$$

where v_0 is the velocity of the influx of water. Note here that there is a relation between v_0 and Q in Eq. (2) as $Q = \pi a^2 v_0$ where a is the radius of the faucet. Then, based on their real experiments, they considered that the stiffness k in Eq. (3) also depends on the mass of the drop as

$$k(m) = \begin{cases} -11.4m + 52.5 & (m < m_c), \\ 0 & (m \geqslant m_c), \end{cases} \qquad (4)$$

where $m_c = 4.61$. Equation (4) means that when the mass m amounts to m_c, the value of the stiffness becomes zero, then the drop undergoes free-fall. In their experiments, the units of the length, time, and mass are chosen as $(\gamma/\rho g)^{1/2} = 0.27$ cm, $(\Gamma/\rho g^3)^{1/4} = 0.017$ sec and $\rho(\gamma/\rho g)^{3/2} = 0.020 g$, respectively, where Γ is the surface tension and ρ is the density. Using these units, parameters are set to $\gamma = 0.05$, $g = 1$, $x_c = 5.5$, $\Delta m = 0.8m - 0.3$, and $a = 0.916$, and the constants in Eq. (4) are also determined from their experiments. They also assumed that just after a portion of the drop separates, the position and velocity are reset to $x = x_0 = 2.0$, $\dot{x} = \dot{x}_0 = 0$.

Figure 2 shows a trajectory of the above mentioned model with $v_0 = 0.1130$ after some transient. In Fig. 2(a) we can see that the trajectory is tracing a chaotic attractor. In real experiments, however, it is in general impossible to observe all state variables of the system. In the case of the dripping faucet, time intervals $\tau_n (n = 1, 2, \ldots)$ between successive drop separations are observed in experiments. How to determine τ_n from the signal of the position $x(t)$ is depicted in Fig. 2(b). Here the variable m_n is the value of the mass at the moment of the nth drop separation.

Figure 3(a) shows the return plot m_{n-1} vs m_n for $v_0 = 0.1130$. We can see that there is a clear scalar function between m_{n-1} and m_n as $m_n = f(m_{n-1})$. In Fig. 3(b), however, the return plot τ_{n-1} vs τ_n is a multivalued relation (i.e., the right-hand side of the return plot shows the 1 to 2 values). From a different viewpoint, if we regard this return plot as the time-delay embedding of τ_n into the two-dimensional plane $\mathbb{R}^2$ as $s_n = (\tau_{n-1}, \tau_n)$ denoting the manifold on which the states s_n
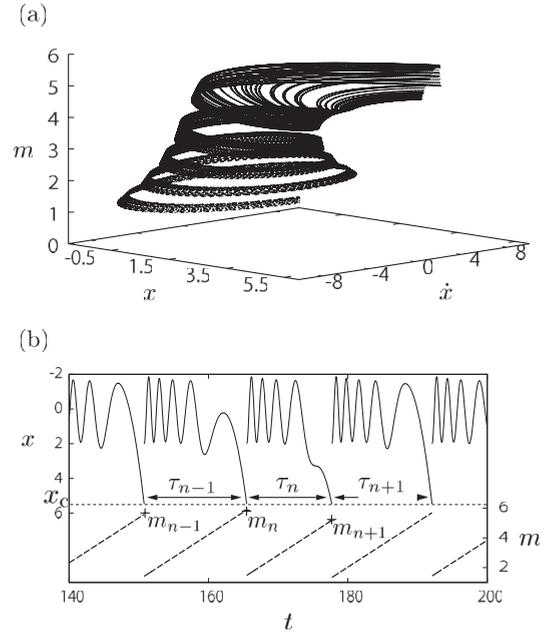
FIG. 2. (a) Chaotic trajectory of the Kiyono-Fuchikami model. (b) Time series of drop position x and the time interval τ between drop separations.

are lying as $\mathcal{S}$, then, we can see that there is a deterministic function between s_{n-1} and s_n as $s_n = \mathbb{F}(s_{n-1})$ in $\mathbb{R}^2$.

C. Bifurcation Structure

We also investigated how the statistical property of the mass-spring model [Eqs. (3) and (4)] depends on the water influx v_0, and the result is shown in Fig. 4(a). One can see repetitions of the period doubling bifurcation route to chaos, as well as periodic windows and their reverses for increasing v_0. We also made experiments to check whether the true dripping faucet system also exhibits this bifurcation structure. Figure 4(b) shows a time series of the time intervals of drop separations over a long time period. Here, in our experiments, the surface of water of the bath decreases very slowly because no water is supplied from outside, which plays a role of changing the water influx. Therefore, this figure represents a kind of "bifurcation" diagram. One can see that there is a significant qualitative similarity between the numerical simulations [Fig. 4(a)] and the real experiments [Fig. 4(b)].
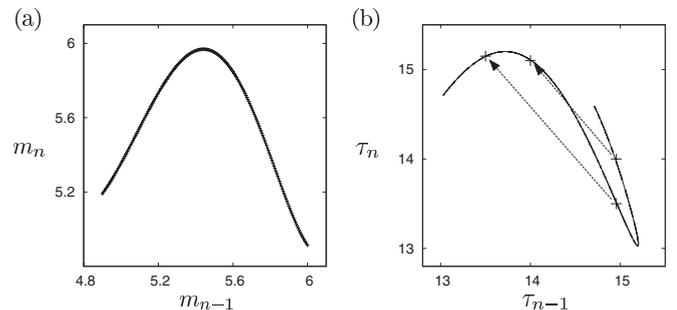
FIG. 3. (a) Return plot of m_{n-1} vs m_n. (b) Return plot of τ_{n-1} vs τ_n. Movements of points as the time delay embedding vector $s_n = (\tau_{n-1}, \tau_n)$ are also depicted.
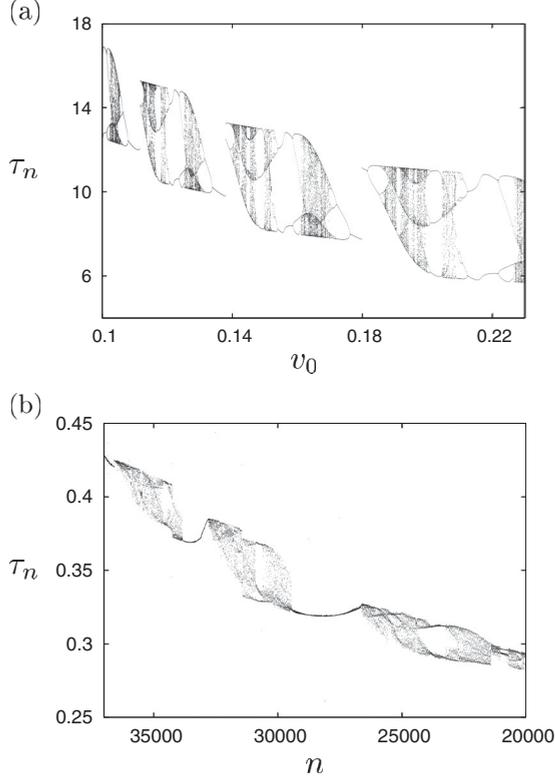
FIG. 4. (a) Bifurcation diagram of the Kiyono-Fuchikami model. Note that the unit of time is rescaled (e.g., $\tau = 14$ corresponds to 0.238 sec). (b) Time series of a real dripping faucet experiment.

FIG. 5. Neighboring graph $\mathcal{G}$. (a) Schematic illustration of the sample points ($\circ$) and the three-neighboring graph $\mathcal{G}$ (solid line). The heavy solid line is the shortest path between points i and j and the dashed line is the Euclidean distance. (b) 500 sample points ($+$) and four-neighboring graph $\mathcal{G}$ (line) for $v_0 = 0.1130$. (c) Graph $\mathcal{G}$ for two-band chaos ($v_0 = 0.1128$). The two clusters B_1, B_2 caused by two-band chaos are connected by the shortest path C.

III. EXTRACTING ONE-DIMENSIONAL MAPS OF INTERNAL STATE VARIABLES BY ISOMAP

As shown in the previous section, this dripping model is essentially described by a one-dimensional map $f(m)$. Its experimental observables τ_n also show a one-dimensional filament in the d-dimensional space ($\tau_{n-d+1}, \tau_{n-d+2}, \ldots, \tau_n$), but the relationship between τ_{n-1} vs τ_n is not always given by a one-dimensional map directly. In this paper, we discuss the case $d = 2$ as shown in Fig. 3. This suggests that the dripping time interval τ isn't appropriate for the simple description of the dynamics. If we can construct a new coordinate u along the filament, the dynamics must be described by a one-dimensional map $u_n = g(u_{n-1})$ and easily analyzed using the theory for one-dimensional maps. In this section, we try to construct a new coordinate u by applying ISOMAP to the time series τ_n and get the one-dimensional map $g(u)$. In addition, we test whether we can recover the statistical properties of the dripping faucet system from the one-dimensional map $g(u)$.

A. ISOMAP

ISOMAP is one of several widely used low-dimensional embedding methods, which is an extension of classical MDS (multidimensional scaling) [7]. MDS seeks a low-dimensional representation of the sample points. This is achieved by plotting data points in a low-dimensional space preserving the "dissimilarity" (generalized distance) between sample points (in the original higher-dimensional space) as much as possible.

For the dripping time series τ_n, when we use the Euclidian distances (pairwise distances between sample points) in the (τ_{n-1}, τ_n) plane as the dissimilarity, we shall not obtain a one-dimensional embedding, because the sample points are located on a curved filament. In ISOMAP, the geodesic distance on the low-dimensional structure instead of the Euclidean distance is used and then we can embed the sample points into a one-dimensional space and get the new coordinate u. A concrete procedure is described as follows.

First, we compute the geodesic distance d_{ij} between ith and jth sample points, which is approximated by the shortest path from one to the other on the neighboring graph $\mathcal{G}$. This graph $\mathcal{G}$ is constructed by locally connecting among sample points (we employ the Euclidean distance to construct k-nearest neighbors). This procedure with $k = 3$ is illustrated in Fig. 5(a). The point i is connected to i_1, i_2, i_3 which are k-nearest neighbors of the point i and all sample points are connected in the same manner. The distance d_{ij} between points i and j is not defined by the Euclidean distance (dashed line), but by the shortest path distance (heavy solid line). As a result, d_{ij} is an approximation of the geodesic distance on the manifold. Figure 5(b) shows the graph $\mathcal{G}$ of the dripping faucet (3) and (4) for $v_0 = 0.1130, N = 500$, and $k = 4$. And d_{ij} is obviously a good approximation of the geodesic distance along the filament.

The bifurcation diagram for this dripping faucet is shown in Fig. 4. When the attractor has n bands or exhibits strong intermittency, the neighboring graph $\mathcal{G}$ may be separated into a

few clusters and then we cannot estimate the geodesic distances d_{ij}. In this case we adopt the shortest connection among the clusters to construct the global graph $\mathcal{G}$. An example for two-band chaos ($v_0 = 0.1128$) is shown in Fig. 5(c).

The second step is MDS whose purpose is to place a set of new points $\boldsymbol{u}_i, (i = 1, 2, \ldots, N)$ in a low-dimensional space so that the dissimilarities d_{ij} in the original state are well approximated by $|\boldsymbol{u}_i - \boldsymbol{u}_j|$ [i.e., we find points $\boldsymbol{u}_i$ that minimize $\sum_{i,j}(d_{ij} - |\boldsymbol{u}_i - \boldsymbol{u}_j|)^2$]. In MDS, for the sample points $\boldsymbol{s}_i, (i = 1, 2, \ldots, N)$ in the m-dimensional Euclidean space, we define the square distance matrix D as $D_{ij} = |\boldsymbol{s}_i - \boldsymbol{s}_j|^2$, and introduce

$$Z = -\tfrac{1}{2} J D J = (JS)(SJ)^{\mathrm{T}}, \tag{5}$$

where $S_{k\ell}$ is the ℓth component of $\boldsymbol{s}_k$ and $J = (J_{k\ell}) = (\delta_{k\ell} - 1/N)$ is called the centering matrix whose effect for S is $(JS)_{k\ell} = S_{k\ell} - \sum_k S_{k\ell}/N$. Next we decompose Z into its eigenvalues and eigenvectors as $Z\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i$ with $\lambda_i \geqslant \lambda_{i+1}$. Then we get $Z = P\Lambda P^{\mathrm{T}}$ where $\Lambda_{ij} = \lambda_i \delta_{ij}$ and P_{ij} is the jth component of $\boldsymbol{p}_i$. Therefore, we obtain the matrix

$$U = P\Lambda^{1/2}, \quad (\Lambda^{1/2})_{ij} = \sqrt{\lambda_i}\,\delta_{ij}, \tag{6}$$

which corresponds to the matrix JS and the new point $\boldsymbol{u}_i$ which is the ith row vector of U. Clearly the new points $\boldsymbol{u}_i$ are reconstructions of the original points $\boldsymbol{s}_i$ and recover the distance $|\boldsymbol{s}_i - \boldsymbol{s}_j|$. If $\lambda_n \gg \lambda_{n+1}$, we can approximate U by its projection into the subspace spanned by the eigenvectors $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$.

As we can find the points $\boldsymbol{u}_i$ only from the distances, we start with the geodesic distances d_{ij} instead of $|\boldsymbol{s}_i - \boldsymbol{s}_j|$ and get the new low-dimensional vector $\boldsymbol{u}_i$ corresponding to the ith sample point on the filament. Finally, we obtain n-dimensional representations $\boldsymbol{u}_i = (u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(n)})$ whose distances preserve the geodesic distances in the original space as much as possible. This procedure is essentially the same as the principal component analysis for the data matrix U [7,9].

B. Results for the mass-spring model

Figure 6(a) shows configurations of sample points for the dripping faucet data (τ_{n-1}, τ_n) onto the $(u^{(1)}, u^{(2)})$ plane obtained from ISOMAP. Here, the spread of points in the direction of the $u^{(2)}$ component is much smaller than that of the $u^{(1)}$ component (about 0.6%). This means that the points on the filament are almost explained only by the first component $u^{(1)}$, which implies that ISOMAP succeeds to unfold the attractor to a straight line and $u^{(1)}$ is considered as a new coordinate along the filament [Fig. 6(b)]. Hereafter, $u_n^{(1)}$ is abbreviated as u_n as we mainly use the first component. The successful unfolding can be also confirmed from the one-to-one relation between the mass m_n and the new coordinate u_n in Fig. 6(c).

As any point on the filament is deterministically mapped to another point, the time evolution of u is described by a one-dimensional map $u_n = g(u_{n-1})$, which is shown in Fig. 7. In addition, the points whose spread in the $u^{(2)}$ direction in Fig. 6(a) is relatively large are located around the folding point (critical point) of the one-dimensional map in Fig. 7(a).
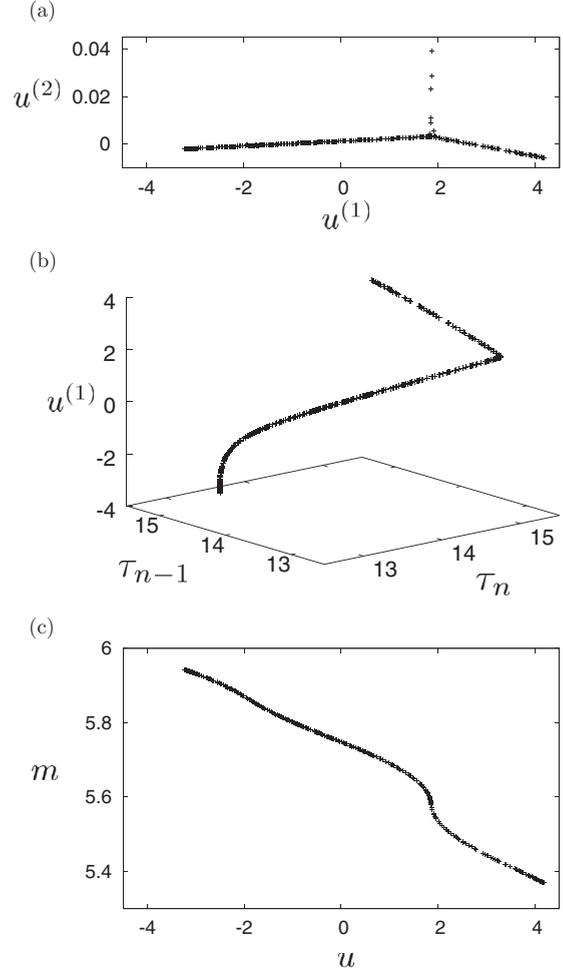
FIG. 6. Results of the application of ISOMAP for Fig. 5(b). (a) Configurations of sample points onto the $(u^{(1)}, u^{(2)})$ plane. (b) Plot of sample points in the $(\tau_{n-1}, \tau_n, u_n^{(1)})$ space. (c) Relation between the new coordinate u and the droplet mass m.

The results in Fig. 7 show the expected relationship between u_{n-1} and u_n. We approximate this one-dimensional map $g(u)$ by a locally quadratic function

$$g(u) = \sum_{j=0}^{2} a_j(u) u^j, \tag{7}$$
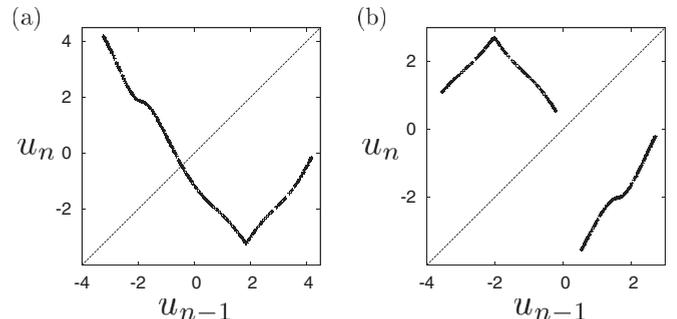
FIG. 7. (a), (b) Return plots of the new variable $u^{(1)}$ generated by ISOMAP for Figs. 5(b) and 5(c), respectively.

where $a_j(u)$ is determined by the local least squares method, i.e., $a_j(u)$ satisfies

$$\min_{\{a_0(u),a_1(u),a_2(u)\}} \sum_n (g(u_{n-1}) - u_n)^2 \exp\left\{-\left(\frac{u_{n-1} - u}{\sigma}\right)^2\right\}$$

and we use $\sigma = 0.05$.

C. Statistical properties

From the one-dimensional map $g(u)$, we can get the natural invariant density $\rho(u)$, which is the base for the discussion of the statistical properties. Here, $\rho(u)$ satisfies the equation

$$\rho(u) = H[\rho(u)] = \int \rho(v)\delta[u - g(v)]dv$$
$$= \sum_{u_n:u=g(u_n)} \frac{\rho(u_n)}{|g'(u_n)|}, \quad g'(u) = \frac{dg}{du} \quad (8)$$

where H is called Frobenius-Perron operator and $\rho(u)$ is generally expected to be the empirical distribution $\lim_{N\to\infty} \sum_{n=1}^{N} \delta(u - u_n)/N$ for a chaotic orbit.

To solve approximately Eq. (8), we divide the domain of $g(u)$ into the intervals $I_i, (i = 1, 2, \dots)$ whose edges are inverse mapping points $g^{-n}(u_*), (n = 0, 1, 2, \dots)$, where u_* is the critical point [the minimum in Fig. 7(a) and the maximum in Fig. 7(b)] of the function $g(u)$, and expand $\rho(u)$ as

$$\rho(u) = \sum_i \alpha_i e_i(u), \quad e_i(u) = \begin{cases} 1, & (u \in I_i) \\ 0, & (\text{others}). \end{cases} \quad (9)$$

We substitute Eq. (9) in Eq. (8) and get

$$\alpha_i = \sum_j H_{ij}\alpha_j, \quad H_{ij} = \frac{\beta_{ij}}{|g'(u_j^{(c)})|},$$
$$\beta_{ij} = \begin{cases} 1, & \text{if } I_j \subset g(I_i) \\ 0, & \text{if } I_j \not\subset g(I_i) \end{cases} \quad (10)$$

where $u_j^{(c)}$ is the center of I_j. The solution α_i is given by the eigenvector corresponding to the eigenvalue 1 of the matrix H. The above-mentioned method is a kind of the Galerkin approximation which is often used in the study of the one-dimensional map [10].

The internal state variable u is not a natural physical quantity for the dripping faucet system. However we can derive any physical quantity A from u, because the quantity A on the filament is determined by u [i.e., $A = A(u)$ and its long-time average is calculated by $\langle A \rangle = \int A(u)\rho(u)du$]. Actually, as the results of ISOMAP provide the relationship between u_n and (τ_{n-1}, τ_n) as shown in Fig. 8(a), the important observable variable τ of the dripping faucet system is determined by

$$\tau = \phi(u), \quad (11)$$

where the function $\phi(u)$ is approximated in the same way as $g(u)$ [see Eq. (7)]. First we get the distribution function

$$P(\tau) = \frac{\rho(u)}{|\phi'(u)|}, \quad (12)$$

which is one of the most basic properties of the dripping faucet. The result in Fig. 8(b) shows $P(\tau)$ with many peaks which are
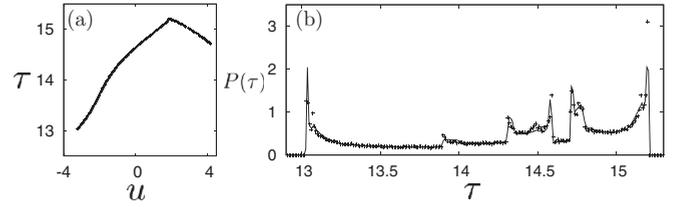
FIG. 8. (a) Relation between the new variable u and the dripping interval τ. (b) Distribution function of dripping intervals $P(\tau)$ which is derived from Eq. (12). The result from the direct simulation (10^6 droplets) are also plotted by the symbol $+$.

generated by the folding processes of $g(u)$. These properties are consistent with the result from the direct simulation of Eqs. (3) and (4).

Next, we calculate the Lyapunov exponent [11] and the topological entropy [12], which characterize the stability and the variety or complexity of chaotic orbits, respectively. The Lyapunov exponent of $g(u)$ is defined by $\Lambda = \lim_{N\to\infty}(1/N) \ln |du_N / du_0| = \lim_{N\to\infty}(1/N) \sum_{n=0}^{N-1} \ln |du_{n+1}/du_n|$ and given by

$$\Lambda = \langle \ln |g'(u)| \rangle. \quad (13)$$

The topological entropy is calculated by [11]

$$H = \frac{1}{N} \ln \left\langle \left| \frac{dg^N(u)}{du} \right| \right\rangle, \quad N \gg 1. \quad (14)$$

As both quantities are invariant under the transformation from u to m, the results from $g(u)$ should coincide with the Lyapunov exponent and topological entropy of $f(m)$ [if the map $g(u)$ is an appropriate description of the dripping faucet dynamics given by Eqs. (3) and (4)]. They are cited in Table I. The results are in good agreement, but the standard deviations of both quantities calculated from $g(u)$ is larger than the results from $f(m)$. This may be caused by enhancing the fluctuation of the return plot by the shortcut in ISOMAP. We can calculate the time series τ_n from the one-dimensional map (7) and Eq. (11), but its long-time behavior is very different from the direct simulation of Eqs. (3) and (4), because the dynamics is chaotic.

TABLE I. Lyapunov exponents and topological entropies calculated from $f(m)$ and $g(u)$. Here, we approximate the one-dimensional map $f(m)$ and get the invariant density $\rho(m)$ in the same way as Eqs. (7)–(10) for the return plot of m_n (Fig. 3). And we use $N = 64$ in Eq. (14). The cited values are the averages and the standard deviations for 20 different sets of time series.

v_0	$f(m)$ or $g(u)$	Lyapunov exp.	topologcal entropy
0.1128	$f(m)$	0.254 ± 0.008	0.280 ± 0.002
	$g(u)$	0.249 ± 0.011	0.274 ± 0.009
0.1129	$f(m)$	0.320 ± 0.008	0.334 ± 0.001
	$g(u)$	0.316 ± 0.021	0.331 ± 0.019
0.1130	$f(m)$	0.360 ± 0.005	0.378 ± 0.003
	$g(u)$	0.349 ± 0.019	0.375 ± 0.012
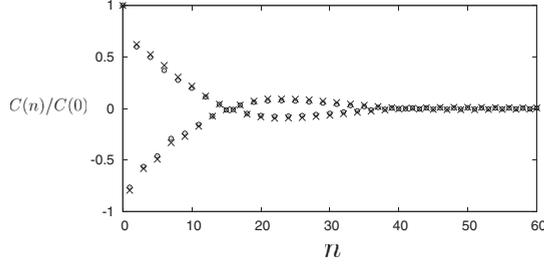
FIG. 9. Time correlation function of τ_n. The results from $g(u)$ (○) for a return plot (500 droplets) and the direct simulation (10^6 droplets) from Eqs. (3) and (4) (×) are plotted. The size of fluctuation for the different return maps is about 30% around $n = 25$. The two curves show the decay for a period-two and a long-period (about 23) oscillation.

We also calculate the time correlation function of τ_n

$$C(n) = \lim_{N \to \infty} \frac{1}{N} \sum_{m=0}^{N-1} (\tau_m - \bar{\tau})(\tau_{m+n} - \bar{\tau})$$

$$= \langle \phi(u)\phi(g^n(u)) \rangle - \langle \phi(u)^2 \rangle, \qquad (15)$$

where $\bar{\tau} = \lim_{N \to \infty} (1/N) \sum_{m=0}^{N-1} \tau_m$. The result in Fig. 9 shows the exponential decay for a period-two and long-period (about 23) oscillation. Their properties are also shown in the result, which is calculated directly from τ_n by the simulation of Eqs. (3) and (4). The good coincidence shows that by using the internal variable u, we can discuss statistical properties of the original observable τ.

D. Application to real experimental data

Last we show preliminary results of the application of the dimension reduction method to our real experimental data. A short time series whose water flux is almost stationary and its first-return plot τ_{n-1} vs τ_n are shown in Figs. 10(a) and 10(b), respectively. This return plot shows a multivalued function in the wide region and cannot describe the dripping faucet dynamics. The application of our method leads to the internal variable u and the one-dimensional map $u_n = g(u_{n-1})$, which is shown in Fig. 10(c). The absolute value of its slope is nearly one in almost all regions (i.e., the instability of the orbits is weak), which is related to the fact that the time series contains one or two periodic-like motions.

It has been pointed out that ISOMAP is topologically unstable for small noise [13]. In actuality, the neighboring graph $\mathcal{G}$ around the folding point of the filament is affected by experimental noise or high-dimensional dynamics and has some shortcuts, which are out of the filament. Therefore, the first-return map of u has a multivalued structure around the critical point (maximum point) of $g(u)$. However, this effect is small and localized so we can say our method is a promising method for not only the numerical study but also experimental data.

Our experiment is a first step and currently we cannot keep it stationary to measure long time series. We work on a revised setup and extended experimental results are a subject for future research.
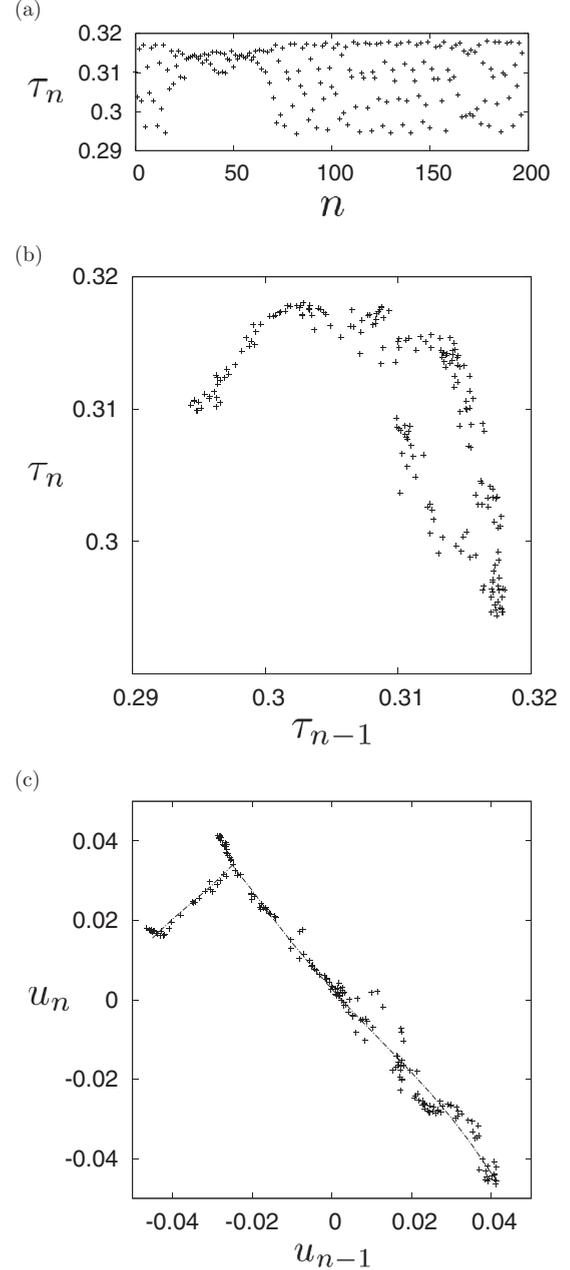
FIG. 10. Results for our experimental data. (a) Dripping time series τ_n, which contains one or two periodic motion around $n = 50, 170$. (b) First-return plot of τ. It shows a multivalued function in the right region. (c) First-return plot of inner state variable u, which suggests that there is a one-dimensional map $g(u)$. It may have a wavelike pattern in the right region.

IV. SUMMARY AND DISCUSSION

In this paper, employing the dripping faucet system as an illustrative example, we studied the problem of constructing a surrogate variable for the internal state of a chaotic dynamical system from time series using manifold learning analysis. Especially, when the time-delay embedding of the observed time series forms a one-dimensional curved structure, we succeeded in obtaining one-dimensional deterministic maps associated with the surrogate variable. The statistical properties of the

original chaotic system were successfully reproduced by its surrogate system.

In real-world applications, not all original state variables of the system can be directly observed. Instead, only some of the original state variables or their transformations are observed in experiments. So, it is often seen that the manifolds obtained from the time-delay embedding may be very complex even if their dimensionality is low, which leads to multivaluedness in the return plots [5]. For example, the time series of interspike intervals is mainly observed in neural systems and its return plot often exhibits a multivalued relation [14]. Besides neural systems, there are a number of such examples (e.g., laser systems [15,16], passive biped walkers [17], and social activity models [18]). Extracting the deterministic relationship from the observations of these models may also be done using dimension reduction methods.

Besides, the return plot of u_{n-1} vs u_n (Fig. 7) is less smooth compare to that of m_{n-1} vs m_n [Fig. 3(a)]. This is because methods of manifold learning are generally unsupervised ones, only using the information of τ_n. If the assumption that the data is generated from a dynamical system with a simple mapping form (say the logistic parabola) can be incorporated additionally to the manifold learning as some additional constraint term then, we can obtain a more refined return plot, which may be more interpretable to us.

In this paper, we have been only concerned with the case in which the internal state behind the observed time series obeys a one-dimensional dynamical system. This is of course

an ideal case. As formal methodologies (application of ISOMAP or other manifold learning methods) are not restricted to one-dimensional manifolds, the presented approach can in principle be extended to higher-dimensional cases. It should be noted, however, that ways of acquiring training samples to obtain a lower-dimensional representation become more important. For example, let us consider the situation where the Hénon attractor $\mathcal{A}$ is lying on a two-dimensional nonlinear manifold $\mathcal{M}$ embedded in, say, the three-dimensional Euclidean space $\mathbb{R}^3$. In order to obtain a lower-dimensional representation of $\mathcal{M}$, not only the data on $\mathcal{A}$, but also the data associated with transient dynamics are needed because $\mathcal{A}$ is too thin to recover the whole two-dimensional structure of $\mathcal{M}$. In addition, the nonuniformity in the natural measure on $\mathcal{A}$ and its transient area affect the performance of manifold learning.

ACKNOWLEDGMENTS

[1] R. Shaw, *The Dripping Faucet as a Model of Chaotic System* (Aerial Press, Santa Cruz, 1984).

[2] K. Kiyono and N. Fuchikami, J. Phys. Soc. Jpn. **68**, 3259 (1999).

[3] N. Fuchikami, S. Ishioka, and K. Kiyono, J. Phys. Soc. Jpn. **68**, 1185 (1999).

[4] T. Katsuyama and K. Nagata, J. Phys. Soc. Jpn. **68**, 396 (1999).

[5] F. Christiansen, P. Cvitanović, and V. Putkaradze, Nonlinearity **10**, 55 (1997).

[6] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer-Verlag, New York, 2001).

[7] J. B. Tenenbaum, V. de Silva, and J. C. Langford, Science **290**, 2319 (2000).

[8] T. F. Cox and M. A. A. Cox, *Multidimensional Scaling* (Chapman and Hall, London, 2000).

[9] C. Williams, Machine Learning **46**, 11 (2002).

[10] T. Kohda and K. Murao, IEICE **73**, 793 (1990); Erik Bollt, Paweł Góra, Andrzej Ostruszka, and Karol Życzkowski, SIAM J. Appl. Dyn. Syst. **7**, 341 (2008).

[11] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, 2002).

[12] B.-L. Hao and W.-M. Zheng, *Applied Symbolic Dynamics and Chaos* (World Scientific Publishing Company, Singapore, 1998).

[13] M. Balasubramanian and E. L. Schwartz, Science **295**, 7 (2002).

[14] U. Feudel, A. Neiman, X. Pei, W. Wojtenek, H. Braun, M. Huber, and F. Moss, Chaos **10**, 231 (2000).

[15] U. Hübner, C.-O. Weiss, N. B. Abraham, and D. Tang, in *Time-Series Prediction: Forecasting the Future and Understanding the Past*, edited by A. S. Weigend and N. A. Gershenfeld, Vol. 73 (Westview Press, Boulder, 1993).

[16] J. Used and J. C. Martín, Phys. Rev. E **79**, 046213 (2009); **82**, 016218 (2010).

[17] A. Goswami, B. Thuilot, and B. Espiau, Int. J. Robotics Res. **17**, 1282 (1998).

[18] G. Feichtinger, L. L. Ghezzi, and C. Piccardi, Int. J. Bifurcation Chaos **5**, 255 (1995).