Noise reduction by recycling dynamically coupled time series

M. Eugenia Mera and Manuel Morán

Dpto. Fundamentos del Análisis Económico I

Universidad Complutense, 28223 Madrid. Spain[*]

(Dated: July 28, 2011)

We say that several scalar time series are dynamically coupled if they record the values of measurements of the state variables of the same smooth dynamical system. We show that much of the information lost due to measurement noise in a target time series can be recovered with a noise reduction algorithm by crossing the time series with another time series with which it is dynamically coupled. The method is particularly useful for reduction of measurement noise in short length time series with high uncertainties.

PACS numbers: 05.45.Tp, 05.10.-a

In the study of real world dynamical phenomena it is common to record the values of a feature of the phenomenon as time varies (time series). Such features can be some of the state variables of the dynamics under study or some real function (observation) of such variables. Typically the measuring device used to record the time series has limited accuracy, and consequently the recorded signal differs from the clean signal by a measurement error. *Noise reduction* means any procedure which eliminates part of the noise, thereby yielding a cleaned signal with geometric, dynamical, and statistical properties as similar as possible to those of the clean signal. The level of noise reduction achieved can be seen as a measure of the reduction of risk, of the information gain, and of the increase in forecasting ability.

Powerful methods of noise reduction, such as the Kalman filter, obtain marginal gains in the information processed only through updating data points when new marginal information arrives. In this paper we show that sharing complementary information spread among different research teams or institutions makes possible a process exploiting and recycling existing information.

The results reported in this paper show that substantial reduction of additive measurement noise in a noisy time series can be achieved by crossing the given time series with other time series with which it is dynamically coupled. We say that several time series are *dynamically coupled* (see Ref.[1] for other approaches to coupled systems in the setting of noisy time series) if they are the time ordered values either of some of the state variables or of real functions of the state variables of the same smooth dynamical system. For instance, classical models in economics show that the pairs of

time series taken respectively from unemployment and inflation, or from interest rates and gross national product, are dynamically coupled. The method of reducing measurement noise by dynamical coupling is especially advantageous when the uncertainty of the time series under study is greater than that of the time series dynamically coupled to it. For instance, in a predator-prey pair, often the stock of one of the species can be observed with greater accuracy than can that of the other. In fact, typical real world dynamically coupled time series (think for instance of clinical, meteorological, or econometric measurements) are heteroscedastic[2,3] (have different levels of uncertainty). Some of the reasons for the heteroscedasticity are that the nature of the data recorded varies (e.g. pressure, temperature, rainfall, etc.), the devices used to record the data produce different sorts of measurement errors, those who record social data and the procedures they use to do so are different, and also because some observables are more difficult to register than others.

It should be remarked that in order to apply the methods proposed in this paper it is not necessary to know the underlying smooth dynamical system. Rather, there suffices the existence of some dynamical system through which the measurements are coupled.

The method of dynamical coupling is especially powerful for very short time series. For instance, in a laboratory time series generated by the logistic map corrupted by a measurement noise with a 50% amplitude and having a length of only 250 data points, we can obtain (see Subsection III B 2) noise reduction levels of up to 80%.

The good behavior of the dynamical coupling procedure for short time series makes it possible to extend the scope of application of noise reduction techniques to the social sciences, where many of the methods used in nonlinear time series analysis have not yet been applied because of the short lengths of the available time series. Also, as a consequence of the ideas in this paper, the potential importance of sharing information between different research teams or institutions becomes clearer.

I. INTRODUCTION

In noise reduction problems the mechanism generating the noisy time series is assumed to have the following general form:

$$\mathbf{y}_{i+1} = \mathbf{F}(\mathbf{y}_i, \xi_i), \tag{1}$$
$$\mathbf{X}_i = \mathbf{H}(\mathbf{y}_i, \mathbf{e}_i).$$

Here $\mathbf{y}_i \in \mathbb{R}^m$ is the unobservable state of a system at time i; the function $\mathbf{F}$ is the evolution law; $\mathbf{H}$ is the measurement function or *observable*; ξ_i, and $\mathbf{e}_i$ are unknown multivariate random variables; and $\mathbf{X}_i$, $i = 1, 2, ..., N$, is the available data set or noisy time series (for continuous time dynamics the model is analogous with $\dot{\mathbf{y}}(t) = \mathbf{F}(\mathbf{y}(t), \xi_t)$). In this scheme ξ_i is referred to as the *dynamical noise* and $\mathbf{e}_i$ as the *observational* or *measurement noise*. The noise is called *additive* if it enters in the corresponding equation in an additive way (there exist functions $\mathbf{f}$ and $\mathbf{h}$ such that $\mathbf{F}(\mathbf{y}_i, \xi_i) = \mathbf{f}(\mathbf{y}_i) + \xi_i$ and/or $\mathbf{H}(\mathbf{y}_i, \mathbf{e}_i) = \mathbf{h}(\mathbf{y}_i) + \mathbf{e}_i$). This paper treats additive measurement noise reduction with unknown $\mathbf{f}$ and $\mathbf{h}$.

The corruption of a deterministic signal by measurement noise may render impractical the estimation of invariant parameters of the underlying nonlinear dynamics such as fractal dimensions, entropies, or Lyapunov exponents (see Refs.[4–7]). Noise can also distort the behavior of forecasting algorithms[8], and tests for detecting nonlinearity, chaos[9], continuity[10], or smoothness[11,12]. For this reason it is important to design procedures robust with respect to measurement noise or to use a noise reduction algorithm prior to using such tests or to estimating these invariant parameters. It is important to point out that the use of a noise reduction algorithm does not imply that a given characteristic of the underlying dynamics may be better estimated using the cleaned time series than using the original noisy time series. The quality of the estimation will depend on the noise reduction algorithm, the level and type of noise, and on the characteristics it is desired to estimate. Consequently, posterior to the noise reduction procedure an analysis is needed in order to evaluate whether the cleaned time series facilitates improvements in the estimation of the parameters of interest.

There are many algorithms available for noise reduction[13–32], but almost all of them are focused on the reduction of noise in univariate time series[13–29]. However, in many experimental problems several observations (scalar time series) of the unknown state variables can be recorded. These observations typically contain measurement errors of different amplitudes. The efficiency of the univariate noise reduction algorithms is then limited by the fact that by treating each signal separately such procedures do not allow the exploitation of the information about the phenomenon contained in the remaining scalar time series.

The main contribution of this paper is to show the efficacy of the dynamical coupling procedure in noise reduction. The dynamical coupling procedure can potentially be used for other purposes than noise reduction, as for instance chaos detection. We remark that the dynamical coupling procedure can be applied to any noise reduction algorithm designed for multivariate time series (for instance those proposed in Refs.[30,32,33]). In this paper we use the algorithm proposed in Ref.[32] (described briefly in Section II), but we give also in Subsection III D an example showing the

utility of the dynamical coupling procedure when using the algorithm proposed in Ref.[33]. Some of the advantages and drawbacks of the different noise reduction algorithms are discussed below and in Subsection II B.

In order to show the usefulness of the dynamical coupling procedure we have studied the level of noise reduction achieved by the noise reduction algorithm with and without the dynamical coupling procedure for noisy time series obtained from logistic, Hénon, and Lorenz dynamics. Our algorithm[32] is an improved version of the one proposed in Ref.[31]. The main virtue of the first version of the algorithm[31] is that its noise reduction scheme takes into account the possible variations in the unknown levels of uncertainty that may exist in the coordinates of a multivariate time series. In the updated algorithm[32] the main improvement was the incorporation of an adaptative neighborhood sizing scheme, which allows reduction of the noise in time series corrupted by high amplitude measurement noise.

Our algorithm proceeds through orthogonal projections of the data points of the time series within certain neighborhoods onto local linear subspaces. These linear subspaces are those closest to the data points in the neighborhoods with respect to a specific distance: that induced by an estimation of the covariance matrix of the measurement errors (see the explanation of this metric in Section II B). This distance gives less weight to coordinates (scalar time series) recorded with lower precision. The noise reduction scheme is iterative. The output time series at a given iteration is the input time series in the next iteration of the algorithm. The effectiveness of the noise reduction scheme means that noise levels decrease as the iterations advance. Then the sizes of the neighborhoods must be reduced accordingly, thereby minimizing the errors due to nonlinearity. In our algorithm the sizes and shape of the neighborhoods[32] are also determined by the remaining level of uncertainty in each coordinate. The neighborhoods will be ellipsoids rather than simply spheres, with semi-axial lengths related to the uncertainties in the corresponding coordinates. We explain in Section II C how these unknown levels of uncertainty are estimated using only the noisy time series.

Although algorithms for univariate time series predominate in the noise reduction literature, there are algorithms designed for multivariate time series[30] or that may be used for both univariate and multivariate time series (see the algorithm proposed in Ref.[21], which is implemented in the program *ghkss* in the TISEAN library[33]). The algorithms proposed in Refs.[30,33] are also based on orthogonal projections onto linear subspaces. These linear subspaces are those closest to the data points in certain neighborhoods, but with respect to a metric that does not take into account the different levels of uncertainty in the individual coordinates of the time series. Since these algorithms are designed for multivariate time series they also allow exploitation of the information contained in multiple observations of a single dynamical system. However for the problem we address in this paper, our algorithm has some advantages:

(i) The metric we choose for determining the linear subspaces is optimal (in a maximum likelihood sense[34]) whereas

the metric used in the cited algorithms is optimal (in a maximum likelihood sense) only if the noise levels in all of the univariate time series are similar.

(ii) Our algorithm adapts the shape and size of the neighborhoods to the uncertainty remaining in each individual coordinate, whereas in many of the noise reduction algorithms the neighborhoods are built by setting in advance a minimum number of neighbors, which is fixed for all neighborhoods and in all the iterations of the algorithm. Obviously for high noise levels the number of neighbors in early iterations must be significantly higher than in later iterations, and in many of these algorithms it is not possible to control the number of neighbors in an automatic way or the computational cost of such control is too high (see Ref.[22]).

(iii) Our algorithm incorporates a stopping criterion (see Section II D) without computational cost, which does not require knowledge of the clean time series. In these algorithms the number of iterations is an input parameter, and only looking at the final results it is possible to determine what number of iterations is best.

The paper is organized as follows. In Section II we describe the main ideas of our algorithm and in Section III we present some results obtained by our algorithm for logistic, Hénon, and Lorenz time series corrupted by noise. In order to show the effectiveness of the dynamical coupling procedure we compare these results with those obtained by processing each time series separately.

II. THE ALGORITHM

We consider the model (1) with additive measurement noise, and an observable $\mathbf{h}$ taking values in $\mathbb{R}^d$. Then, there are d available univariate noisy time series $\{X_i^j, \ i = 1, ..., N\}$, $j = 1, ...d$ where

$$X_i^j = s_i^j + e_i^j; \quad i = 1, ..., N; \quad j = 1, ..., d.$$

For each j, $j = 1, ..., d$, we assume that $\{e_i^j; \ i = 1, ..., N\}$ are independent and identically distributed (i.i.d.) random variables with mean zero and Gaussian distribution, and $\{s_i^j, i = 1, ..., N\}$ is an unknown deterministic time series obtained from the observation of the first N points of an orbit of a smooth chaotic dynamical system,

$$s_i^j = h^j(\mathbf{y}_i) \text{ and } \mathbf{y}_{i+1} = \mathbf{f}(\mathbf{y}_i),$$

where $\mathbf{f} : M \subset \mathbb{R}^m \to M$ and $h^j : M \subset \mathbb{R}^m \to \mathbb{R}$, $j = 1, ..., d$ are unknown smooth functions. For the case of continuous time dynamics, the d noisy scalar time series are obtained in an analogous way. We assume first that $d \geq m$, and we will explain how to proceed in the case $d < m$ in Subsection II D.

The multivariate time series that the algorithm takes as input is $\{\mathbf{X}_i, \ i = 1, ..., N\} \subset \mathbb{R}^d$ where $\mathbf{X}_i := \mathbf{s}_i + \mathbf{e}_i$, $\mathbf{X}_i := (X_i^1, X_i^2 ..., X_i^d)$, $\mathbf{s}_i := (s_i^1, s_i^2 ..., s_i^d)$, and $\mathbf{e}_i := (e_i^1, e_i^2, ..., e_i^d)$. The aim of our noise reduction procedure is the separation of the deterministic component $\{\mathbf{s}_i, \ i = 1, ..., N\}$ from the observational noise $\{\mathbf{e}_i, i = 1, ..., N\}$, in the sense that the output time series of the algorithm, denoted $\{\widehat{\mathbf{s}}_i, \ i = 1, ..., N\}$, must recover the main geometric, dynamical, and statistical properties of $\{\mathbf{s}_i, \ i = 1, ..., N\}$.

We present now the main ideas underlying our algorithm. The details of the implementation can be found in Ref.[32].

A. Noise reduction through orthogonal projections onto linear subspaces

The three delay embedding $\{(\mathbf{s}_{i-1}, \mathbf{s}_i, \mathbf{s}_{i+1}), i = 2, ..., N - 1\}$ of the clean data belongs to an m-dimensional submanifold of $\mathbb{R}^{3d}$. Given any point of this three delay embedding, and a small neighborhood of such a data point, all the data points of the three delay embedding within such a neighborhood, are close to an m-dimensional linear subspace (the space tangent to the submanifold at the center of the neighborhood). As result of the noise, the observed data within the small neighborhoods are separated from the tangent spaces, so the noise can be partially removed by estimating the m-dimensional linear subspace closest to the data points in each neighborhood and projecting the data points within the neighborhood onto it. To our knowledge this idea was implemented for the first time in an algorithm for reducing the noise of a multivariate time series in Ref.[30].

Our algorithm requires a three delay embedding $\{\mathbf{Z}_i := (\mathbf{X}_{i-1}, \mathbf{X}_i, \mathbf{X}_{i+1}), i = 2, ..., N - 1\} \subset \mathbb{R}^{3d}$ of the observed time series $\{\mathbf{X}_i, i = 1, ..., N\}$, a neighborhood U_i for each point $\mathbf{Z}_i$, and the center of mass $\langle \mathbf{Z}_i \rangle$ of the points $\{\mathbf{Z}_k, \text{ for } \mathbf{Z}_k \in U_i\}$. The noise is partially removed by replacing $\mathbf{Z}_k$ by $\widehat{\mathbf{Z}}_{k(i)} := \langle \mathbf{Z}_i \rangle + P_{\mathbf{T}_i}(\mathbf{Z}_k - \langle \mathbf{Z}_i \rangle)$, where $\mathbf{T}_i$ is the m-dimensional linear subspace closest (with respect to a distance specified below) to the data points in U_i, and $P_{\mathbf{T}_i}(\mathbf{Z}_k - \langle \mathbf{Z}_i \rangle)$ is the orthogonal projection of the vector $\mathbf{Z}_k - \langle \mathbf{Z}_i \rangle$ onto $\mathbf{T}_i$. Thus $\widehat{\mathbf{Z}}_{k(i)}$ is the estimation of $(\mathbf{s}_{k-1}, \mathbf{s}_k, \mathbf{s}_{k+1}) \in \mathbb{R}^{3d}$ obtained from the neighborhood U_i, provided that $\mathbf{Z}_k \in U_i$. Since a given point $\mathbf{Z}_k$ can belong to many neighborhoods U_i, we have an estimation $\widehat{\mathbf{Z}}_{k(i)}$ for each such neighborhood. The final estimation $(\widehat{\mathbf{s}}_{i-1}, \widehat{\mathbf{s}}_i, \widehat{\mathbf{s}}_{i+1})$ of $(\mathbf{s}_{i-1}, \mathbf{s}_i, \mathbf{s}_{i+1}) \in \mathbb{R}^{3d}$ is the average of all of these estimations. This gives a $3d$-dimensional time series and therefore three different estimations for each $\mathbf{s}_i, i = 2, ..., N - 1$: those given respectively by the final, intermediate, and initial d coordinates of $(\widehat{\mathbf{s}}_{i-2}, \widehat{\mathbf{s}}_{i-1}, \widehat{\mathbf{s}}_i)$, $(\widehat{\mathbf{s}}_{i-1}, \widehat{\mathbf{s}}_i, \widehat{\mathbf{s}}_{i+1})$, and $(\widehat{\mathbf{s}}_i, \widehat{\mathbf{s}}_{i+1}, \widehat{\mathbf{s}}_{i+2})$. We take as the estimation of $\mathbf{s}_i$ a weighted average of these three estimates. The weights are the same for all the points and inversely proportional to the sample variances of the corrections $\epsilon_i := (\mathbf{X}_{i-1}, \mathbf{X}_i, \mathbf{X}_{i+1}) - (\widehat{\mathbf{s}}_{i-1}, \widehat{\mathbf{s}}_i, \widehat{\mathbf{s}}_{i+1})$, $i = 2, ..., N - 1$.

The metric we implement in our algorithm is that induced by $\widehat{\Sigma}$, where $\widehat{\Sigma}$ is an estimation of the covariance matrix

of the errors contained in the data points $\mathbf{Z}_j$, $j = 2, ..., N-1$. In Subsection II C we explain how this covariance matrix is estimated using only the information contained in the noisy time series. The metric induced by $\widehat{\Sigma}$ takes into account the different levels of uncertainty for the different coordinates of the multivariate time series. The m-dimensional linear subspace $\mathbf{T}_i$ is that subspace, in the set $\mathcal{L}_m$ of m-dimensional linear subspaces, closest to the data points $\{\mathbf{Z}_j - \langle \mathbf{Z}_i \rangle, \mathbf{Z}_j \in U_i\}$ with respect to the distance induced by $\widehat{\Sigma}$. Then $\mathbf{T}_i$ is the solution of the problem,

$$\min_{\mathbf{T} \in \mathcal{L}_m} \sum_{j:\ \mathbf{Z}_j \in U_i} \mathbf{w}_j^t\, \widehat{\Sigma}^{-1} \mathbf{w}_j$$

where $\mathbf{w}_j := \mathbf{Z}_j - \langle \mathbf{Z}_i \rangle - \mathbf{P_T}(\mathbf{Z}_j - \langle \mathbf{Z}_i \rangle)$ and $\mathbf{P_T Z} := \arg\min_{\mathbf{u} \in \mathbf{T}} (\mathbf{u} - \mathbf{Z})^t \widehat{\Sigma}^{-1}(\mathbf{u} - \mathbf{Z})$. This is an optimization problem with an explicit solution easily implemented numerically (see Ref.[32]).

Notice that in the homoscedastic case, $\widehat{\Sigma} = \sigma^2 \mathrm{I}$, where I is the identity matrix, the distance $d_\Sigma(\mathbf{u}, \mathbf{Z}) = \left((\mathbf{u} - \mathbf{Z})^t \widehat{\Sigma}^{-1}(\mathbf{u} - \mathbf{Z}) \right)^{1/2}$ gives the same solution as does the Euclidean distance $d_2(\mathbf{u}, \mathbf{Z}) = \|\mathbf{u} - \mathbf{Z}\|$. In the uncorrelated and heteroscedastic case, $\widehat{\Sigma}$ is a diagonal matrix with entries $\widehat{\sigma}_i^2$, $i = 1, ..., 3d$, and d_Σ is the weighted distance $d_\Sigma(\mathbf{u}, \mathbf{Z}) = \sqrt{w_1(u^1 - Z^1)^2 + ... + w_{3d}(u^{3d} - Z^{3d})^2}$ with weights $w_i = \frac{1}{\sigma_i^2}, i = 1, ..., 3d$. Observe that when σ_i^2 decreases, w_i increases correspondingly, with the final effect that the difference $(u^i - Z^i)$ in the solution $\mathbf{u} = (u^1, ..., u^{3d})$ which minimizes $d_\Sigma(\mathbf{u}, \mathbf{Z})$ is forced to be small.

B. The metric induced by $\widehat{\Sigma}$ and alternative metrics

Noise reduction via orthogonal projections of the data in neighborhoods onto linear subspaces is also the basis of other noise reduction algorithms designed for scalar[21–29] and multivariate[30] time series. The performance of the algorithms depends strongly on the metric considered, because the metric determines which linear subspace T_i is closest to the data points in the neighborhood U_i, what is the orthogonal projection $P_{\mathbf{T}_i}(\mathbf{Z}_j - \langle \mathbf{Z}_i \rangle)$ of a vector $\mathbf{Z}_j - \langle \mathbf{Z}_i \rangle$ onto the linear subspace $\mathbf{T}_i$, and, in our algorithm, also what are the shapes of the neighborhoods U_i. Furthermore it determines the statistical properties of the estimations[34]. For these reasons we think that the metric used should not be the same for scalar and multivariate time series. The algorithms for scalar time series work with delay vectors, and the level of uncertainty in all the coordinates of the delay vectors is the same. Therefore to consider a metric which weights all coordinates equally, as is done in Ref.[23,24,28,29], seems a reasonable choice. In fact it gives good results. It has been proved[21] that a metric which focuses the noise reduction on the most stable middle coordinates of the delay vectors (taking the weights corresponding to the first and last coordinates to be large relative to the other weights, which are set equal to one) gives excellent results for chaotic scalar time series. For this reason

this algorithm, implemented in the TISEAN package[33], is among those most frequently used for noise reduction of scalar time series.

Although for multivariate heteroscedastic noisy time series, it is known that the Euclidean metric produces biased estimators (see Refs.[14,35,36]), this is the metric most frequently used. However the estimations of the true values of the multivariate time series obtained taking as the metric that induced by $\widehat{\Sigma}$ have very good statistical properties: for linear models they are unbiased and consistent estimations of the true values and, for Gaussian errors, they are those of maximum likelihood (see Ref.[34]). This metric gives elliptical neighborhoods rather than simply spherical ones. The univariate time series with more accuracy have more weight in the determination of the linear subspaces T_i where the data points in neighborhoods are projected, and the greater the degree of uncertainty in a coordinate, the greater the correction that the algorithm will make to it. Obviously if the multivariate noisy time series is homoscedastic our algorithm can be applied. The estimation of the covariance matrix of the errors will be a multiple of the identity matrix, and our metric gives results analogous to those obtained with the Euclidean metric. However, in the homoscedastic case, the algorithm proposed in Ref.[30] might perform even better than ours because it uses a metric that takes into account that the d central coordinates of the points $\{\mathbf{Z}_i, i = 2, ..., N - 1\} \subset \mathbb{R}^{3d}$ are, for dynamical reasons, more stable than the remaining $2d$ coordinates.

C.   Estimation of the covariance matrix of the errors

Our algorithm requires at each iteration the estimation of the covariance matrix of the errors in the data points. This estimation is made using only the noisy time series.

Let $\{\widehat{\mathbf{s}}_i^k, i = 1, ..., N\} \subset \mathbb{R}^d$ be the output of the algorithm at iteration $k \geq 1$, which is also the multivariate input time series at iteration $k + 1$. Let $\{\mathbf{Z}_i^0 := (\mathbf{X}_{i-1}, \mathbf{X}_i, \mathbf{X}_{i+1}), i = 2, ..., N - 1\}$ be the three-embedding of the initial noisy time series, and let $\{\mathbf{Z}_i^k := (\widehat{\mathbf{s}}_{i-1}^k, \widehat{\mathbf{s}}_i^k, \widehat{\mathbf{s}}_{i+1}^k), i = 2, ..., N - 1\}$. An estimate $\widehat{\Sigma}^k$ of the covariance matrix of the errors in $\{\mathbf{Z}_i^k, i = 2, ..., N - 1\}$ is needed for obtaining the linear subspace $\mathbf{T}_i$ closest, with respect to the metric $\widehat{\Sigma}^k$, to the data points in the neighborhood U_i. The matrix $\widehat{\Sigma}^k$ is also needed in the construction of the neighborhoods $U_i := \{\mathbf{Z}_j^k : (\mathbf{Z}_j^k - \mathbf{Z}_i^k)^t (\widehat{\Sigma}^k)^{-1} (\mathbf{Z}_j^k - \mathbf{Z}_i^k) \leq r\}$ where r is taken as small as possible subject to the requirement that, with a given confidence level, the relevant information is contained in the neighborhoods (see Ref.[32]). Finally, the matrix $\widehat{\Sigma}^k$ also determines the orthogonal projection $\mathbf{P}_{\mathbf{T}_i}\mathbf{Z} := \arg \min_{\mathbf{u} \in \mathbf{T}_i} (\mathbf{u} - \mathbf{Z})^t (\widehat{\Sigma}^k)^{-1} (\mathbf{u} - \mathbf{Z})$ of a vector $\mathbf{Z}$ onto $\mathbf{T}_i$.

If we do not have any a priori information about the initial level of uncertainty in each coordinate of the time series

(i.e. about the covariance matrix Σ^0), then we run l iterations of our algorithm in order to obtain an initial estimation $\widehat{\Sigma}^0$ of Σ^0. In the first of these l iterations, we take $\widehat{\Sigma}_0^0 = \mathrm{I}$, where I is the identity matrix, as the estimation of the covariance matrix of the errors in $\{\widehat{\mathbf{Z}}_i^0, i = 2, ..., N - 1\}$. For the $(k + 1)$st iteration, $k > 1$, we take as the estimation $\widehat{\Sigma}^k$ of the covariance matrix of the errors in $\{\widehat{\mathbf{Z}}_i^k, i = 2, ..., N - 1\}$ the sample covariance matrix of the corrections

$$\widehat{\mathbf{e}}_i^k := \widehat{\mathbf{Z}}_i^k - \widehat{\mathbf{Z}}_i^{k-1}, \;\; i = 2, ..., N - 1,$$

made by the algorithm at the kth iteration. At the end of the lth iteration, the estimate $\widehat{\Sigma}_l^0$ of Σ^0 is the sample covariance matrix of the total corrections

$$\widehat{\epsilon}_j := \widehat{\mathbf{Z}}_i^l - \widehat{\mathbf{Z}}_i^0, \;\; i = 2, ..., N - 1,$$

made by the algorithm in these l iterations. The number l of iterations must be large enough to guarantee that $\widehat{\Sigma}_l^0$ stabilizes. We have observed that eight iterations suffice for guaranteeing stabilization, so $\widehat{\Sigma}_8^0$ is the estimate of Σ^0 we have used in all the experiments below.

For the $(k + 1)$st iteration, $k \geq 1$, we take the sample covariance matrix $\widehat{\Sigma}^k$ of the corrections made by the algorithm at the kth iteration as the estimation of the covariance matrix Σ^k of the true remaining errors in the time series $\{\mathbf{Z}_i^k, i = 2, ..., N - 1\}$.

D. Stopping criterion and embedding issues

We have explained before that any noise reduction algorithm proceeds iteratively. The stopping criterion we use is based on our adaptive neighborhood construction. As the iterations increase, the noise levels and the sizes of the neighborhoods U_i decrease. A clear indication that further iterations of the algorithm will not reduce the remaining noise is the stabilization of the mean number of points in the neighborhoods, and this is our stopping criterion.

Now we indicate how to treat the case $d < m$. That is, the case in which the number of scalar time series d is less than the dimension m of the space where the dynamics is defined. In this case we have to ensure that the three delay embedding of the d-dimensional clean time series reproduces the main geometric, dynamical, and statistical properties of an orbit of the original dynamics (see Ref.[37]). This is guaranteed if $d = m$, for generic observables h^j, and also if $d < m$, provided $3d > 2m$ holds. If $d < m$ but $3d \leq 2m$, then we take a number n of additional univariate time series obtained by taking delays of some of the d original noisy time series, with n satisfying $3(d + n) > 2m$ in order to guarantee that the three delay embedding $\{(\mathbf{s}_{i-1}, \mathbf{s}_i, \mathbf{s}_{i+1}), i = 2, ..., N - n\}$ of the unknown new clean time series $\{\mathbf{s}_i, i = 1, ..., N - n + 1\} \subset \mathbb{R}^{d+n}$ provides a good reconstruction[37] of the unknown dynamics.

III.   NUMERICAL RESULTS

In this section we show the advantages of dynamic coupling for time series generated by logistic, Hénon, and Lorenz dynamics corrupted by additive measurement Gaussian noises with zero mean and different amplitudes.

The unknown noise level in the jth observed time series is given by the noise-to-signal standard deviation ratio

$$NSR^j := \sqrt{\frac{\sum_{i=1}^{N} \left(e_i^j - \langle e^j \rangle\right)^2}{\sum_{i=1}^{N} \left(s_i^j - \langle s^j \rangle\right)^2}}, \ j = 1, ..., d, \tag{2}$$

where $\langle e^j \rangle := \frac{1}{N} \sum_{i=1}^{N} e_i^j$ and $\langle s^j \rangle := \frac{1}{N} \sum_{i=1}^{N} s_i^j$.

We quantify the noise removed by the algorithm in each of the scalar time series using the method standard in the noise reduction literature[7,17–21,23–32]: comparing the pointwise distances between the clean time series and the noisy time series before and after the noise reduction procedure. If the pointwise distance after the noise reduction

$$d_p(\widehat{\mathbf{s}}^j) := \left(\frac{1}{N} \sum_{i=1}^{N} \left(s_i^j - \widehat{s}_i^j\right)^2\right)^{1/2},$$

is less than before the noise reduction $d_p(\mathbf{X}^j)$, then the noise level in $\widehat{\mathbf{s}}^j$ is less than the noise level in the jth coordinate of the input noisy time series $\mathbf{X}^j$. The percentage of pointwise noise reduction in the jth coordinate of the time series is $R^j := 100 \left(1 - \frac{d_p(\widehat{\mathbf{s}}^j)}{d_p(\mathbf{X}^j)}\right)$. A high value of R^j guarantees that any measurement depending on the values of the clean time series $\mathbf{s}^j$ (for instance its mean, standard deviation etc.) estimated from $\widehat{\mathbf{s}}^j$ yields better results than those that would be obtained from $\mathbf{X}^j$.

The calculation of R^j requires knowledge of the clean time series. We use such knowledge only in quantifying the level of noise reduction. It is used neither in the noise reduction scheme nor in deciding when the algorithm must stop. In fact, the evaluation of the performance of the noise reduction algorithm is independent of and posterior to the noise reduction process. The noise reduction requires as information only the given noisy time series, whereas an evaluation of the performance of the algorithm based on the pointwise distance requires knowledge of the clean time series. Thus the noise reduction algorithm can work for data generated by an unknown process, although an evaluation of its performance based on the pointwise distance would not be possible in such a case. However there exist other measures of noise reduction that do not require knowledge of the clean time series (for instance $\widehat{R}_{dyn}$ defined in (4) in Subsection III F).

We have computed other noise reduction measures in all the experiments below. These alternative noise reduction measures support in a similar manner the effectiveness of the dynamical coupling procedure. In the interest of brevity,

we present below only the results corresponding to the noise reduction measure most common in the literature, that based on the pointwise distance. However, in one example, recorded in Table V in Section III F, we report the noise reduction values obtained using these other noise reduction measures. We also report there the drawbacks of the noise reduction measures based on the pointwise distance to the clean time series.

A. Description of the experiments

The equation of the logistic map is

$$y(k+1) = 4y(k)(1 - y(k)).$$

The Hénon map is given by the equations

$$y_1(k+1) = 1 - ay_1(k)^2 + y_2(k), \quad y_2(k+1) = by_1(k),$$

where the parameter values are taken to be $a = 1.4$ and $b = 0.3$.

The Lorenz dynamics is defined by

$$\dot{y}_1 = \sigma(y_2 - y_1), \ \dot{y}_2 = y_1(R - y_3) - y_2, \ \dot{y}_3 = y_1 y_2 - by_3,$$

where the parameter values are taken to be $\sigma = 16$, $R = 45.92$ and $b = 4$. These equations were integrated using a fourth order Runge-Kutta algorithm with an integration step of $\Delta t = 0.001$. The sampling time we consider is $30\Delta t = 0.03$.

Since one of our aims is to show the applicability of dynamic coupling in problems coming from the social sciences, where time series are usually short, we focus on showing the results of the algorithm for time series of short and medium lengths, although for completeness we also present some results for long time series.

For short time series the results of the algorithm depend strongly on the realization of the error term, and also on the clean time series considered. The level of noise reduction achieved by any noise reduction algorithm must be understood in terms of the distribution of the outputs of the algorithm. For this reason we repeat each experiment with a given set of conditions several times for different clean time series and realizations of the error term. We take a long clean time series of 50000 data points. This time series is observed using different observables h^j, $j = 1, ..., d$. These d time series of 50000 data points are corrupted using d independent and uncorrelated Gaussian noises of different variances. Finally we split each of these d time series into L time series each having $N = 50000/L$ data points. We take as the length N of the time series $N \in \{250, 500, 1000, 2500, 5000\}$, in order to study the behavior of

the algorithm in time series of very short, short, medium and large lengths. We quantify the level of noise reduction using the mean $\langle R^j \rangle$ and the standard deviation s_{R^j} of R^j, $j = 1, ..., d$, obtained using the L time series. Thus for $N = 250$, the sample means $\langle R^j \rangle$ and the sample standard deviations s_{R^j}, $j = 1, ..., d$, are calculated using 200 time series, whereas for $N = 5000$ these values are calculated using 10 time series.

B. Results for the logistic map with and without dynamical coupling

1. Medium amplitude and moderated heteroscedastic noises $NRS = (10\%, 5\%, 7.5\%)$

Table I presents for comparison results obtained by applying our algorithm using two and three dynamically coupled noisy time series, and Table II records the corresponding results obtained without using the dynamical coupling procedure. The aims of this experiment are to show the advantages of the dynamical coupling in reducing the noise of the target time series (the corresponding to the highest noise level), and to show some statistical properties of the distribution of the outputs of the algorithm when the dynamical coupling procedure is used.

The three noisy scalar time series $\mathbf{X}^1$, $\mathbf{X}^2$, and $\mathbf{X}^3$ correspond to the logistic map, the observables $h^1(y) = \sqrt{5y + 4}$, $h^2(y) = (y + 2)^2$, and $h^3(y) = 2y + 3$, respectively, and to the noise to signal ratios 10%, 5%, and 7.5%, respectively.

We record in the second column of Table I the sample mean $< R^1 >$ and the sample standard deviation s_{R^1} of the noise reduction levels obtained for the target time series $\mathbf{X}^1$, when it is processed using the coupled time series $\mathbf{X}^2$, corresponding to the observable h^2 and $NRS^2 = 5\%$. The third column of Table I contains the corresponding values of $< R^2 >$ and s_{R^2}. The fourth column of Table I displays the results obtained for $\mathbf{X}^1$ using as input one additional coupled time series $\mathbf{X}^3$, that corresponding to the observable h^3 and $NRS^3 = 7.5\%$. Again for completeness there appear in the last two columns of the table the noise reduction values $(< R^j >, s_{R^j})$, $j = 2, 3$, obtained for the time series $\mathbf{X}^2$ and $\mathbf{X}^3$, respectively. These results can be used to assess the gains, in terms of the improvements in the noise reduction achieved for the time series $\mathbf{X}^1$ and $\mathbf{X}^2$, caused by the additional information provided by the time series $\mathbf{X}^3$ corrupted with noise having the intermediate level $NSR^3 = 7.5\%$.

Table II records the results obtained without the dynamical coupling procedure. Comparison of these results with those given in Table I makes apparent the advantages of dynamical coupling over processing each time series separately. Such a comparison is made below.

The results reported in the second column of Table I show high levels of noise reduction for the target time series $\mathbf{X}^1$ even for very short time series (only 250 data points). The mean noise reduction together with the standard

deviation value also indicate that surprisingly high levels of noise reduction are attained for all the 200 time series corresponding to the length $N = 250$. We can see that the level of noise reduction increases with the number N of points in the time series and that the standard deviation decreases with N. Although we do not pretend to show that the dynamical coupling procedure is advantageous in reducing the noise of the most accurate time series $\mathbf{X}^2$, the results reported in the third column also show high mean noise reduction for this time series, especially for large values of N.

The righthand columns of the table display the results obtained using the additional information provided for the time series $\mathbf{X}^3$ having an intermediate noise level $NSR^3 = 7.5\%$. We can see that the mean noise reduction values for the target time series $\mathbf{X}^1$ are higher, for all the lengths N, than those obtained without using the additional coupled time series $\mathbf{X}^3$. The same behavior can also be observed for the most accurate time series $\mathbf{X}^2$. The explanation is that although NSR^3 is higher than NSR^2, it is less than NSR^1. Thus it is more advantageous to process the most accurate time series $\mathbf{X}^2$ jointly with $\mathbf{X}^1$ and $\mathbf{X}^3$ than it is to process it only with the noisiest time series $\mathbf{X}^1$.

To facilitate visualization of the behavior pattern, the results contained in Table I corresponding to the target time series $\mathbf{X}^1$ (the second and fourth columns) are plotted in Figure 1(a).

In Table II we present the results obtained using our algorithm without the dynamical coupling procedure. In this case the algorithm only uses the information provided for one of the scalar time series. The input time series for the algorithm are a scalar time series (denoted as $dm = 1$) or a dm-dimensional delay embedding of such a scalar time series with $dm = 2, 3$. The results for the dm-embeddings for $dm = 1, 2, 3$ of the target time series $\mathbf{X}^1$ are on the left of the table (the second, third, and fourth columns, respectively), and of the most accurate time series $\mathbf{X}^2$ are on the right of the table (the fifth, sixth, and seventh columns, respectively).

To facilitate visualization of the behavior pattern, the results contained in Table II corresponding to $\mathbf{X}^1$ (the second, third, and fourth columns) are plotted in Figure 1(b).

If we compare the results of Table II corresponding to the target time series $\mathbf{X}^1$ (the second, third, and fourth columns, which are plotted in Figure 1(b)) with those given in Table I (the second and fourth columns, which are plotted in Figure 1(a)) we can see that the dynamical coupling procedure significantly improves the levels of noise reduction (see also Figure 1(c)). For instance, the best result obtained without the dynamical coupling is $< R^1 >= 74.4198$ for $dm = 3$ and $N = 5000$, and values of $< R^1 >$ higher than 70% require $N = 2500$ data points. However, levels of noise reduction higher than 70% are attained with the dynamical coupling procedure (with one or two dynamically coupled time series) for only 250 data points, and for long time series the values of $< R^1 >$ are

higher than 85%.

Even for the time series $\mathbf{X}^2$ corresponding to the initial noise levels of 5%, (compare the results in the fifth, sixth and seventh columns of Table II with those contained in the third and fifth columns of Table I), the levels of noise reduction for short or medium length time series are higher when it is processed with the other one or two time series dynamically coupled with it, even when these time series present a higher level of uncertainty. If the difference between the initial noise levels were higher, a similar result would not be possible. There do not appear in Table II the results for $\mathbf{X}^2$, and $dm = 3$ and short time series because our algorithm is not able to reduce the noise efficiently for such low initial noise level and time series so short.

The results reported in Table II also show that the noise reduction of scalar time series requires a significantly higher number of points than is needed when we have several dynamically coupled time series, as it indicates the fact that the values of $< R^j >$, $j = 1, 2$ for $dm = 3$ are smaller than those for $dm = 2$ if $n < 5000$. This problem is more serious when working with dynamics defined in spaces of high dimensions, where in order to guarantee a good embedding it will be necessary to take an embedding dimension dm large enough (see Ref.[37]).

2. Highly heteroscedastic noises $NRS = (50\%, 5\%)$

The advantage of dynamic coupling is greater when the initial noise levels are very different. This can be seen in Table III, which gives the results obtained for two dynamically coupled time series corresponding to the logistic map, the observables $h^1(y) = \sqrt{5y + 4}$ and $h^2(y) = (y + 2)^2$, and the noise levels $NSR^1 = 50\%$ and $NSR^2 = 5\%$ respectively. The table presents on its righthand side the results obtained without the dynamical coupling procedure for the target time series $\mathbf{X}^1$ and on its left the results using the dynamical coupling procedure (for completeness the table shows also the results for the dynamically coupled time series $\mathbf{X}^2$). The results contained in the second and fifth columns of Table III are plotted in Figure 2 in order to aid visualization of the patterns of behavior with and without dynamical coupling (for $dm = 2$) for the target time series $\mathbf{X}^1$.

The results reported in this table and Figure 2 show that, using our algorithm, the advantage of the dynamical coupling procedure is clear. For $N = 250$, the values $(< R^1 >, s^1) = (81.4044, 3.3442)$ obtained with dynamical coupling are, respectively, approximately double and half the best result $(< R^1 >, s^1) = (42.5436, 7.6653)$, obtained without dynamical coupling. This result gives, at least for short length time series, clear evidence of the difficulty, if using our algorithm, of improving the noise reduction levels achieved for the target time series $\mathbf{X}^1$ without using the information provided for some more accurate dynamically coupled time series. Notice in Table III, for the dynamical

coupling case, the high levels of noise reduction achieved in the target time series $\mathbf{X}^1$ (up to 93% for long time series and up to 80% for only 250 data points), the moderate values of the standard deviations of such noise reduction levels, and how the standard deviations decrease quickly as the number of points N increases. However, using only the information provided by the target time series $\mathbf{X}^1$ we can not reach a noise reduction level of 50% for any of the cases considered.

Although we focus on the advantages of the dynamical coupling procedure for the noisiest time series $\mathbf{X}^1$, in this example its effects for the most accurate time series are also remarkable. The algorithm significantly reduces the noise for both short and long time series (up to 50% noise reduction for $N = 250$ and up to 64% for $N = 2500$), in a case where the level of initial noise is low (only 5%) and the coupled time series has 10 times its level of uncertainty. In spite of such values, since the difference between the initial noise to signal ratios of the two time series is so high, in order to reduce the noise of the scalar time series measured with more accuracy it is better to process such time series separately.

The output time series obtained with and without the dynamical coupling procedure for one of the numerical examples reported in Table III (one of the $L = 10$ time series of length $N = 5000$ corresponding to the target time series $\mathbf{X}^1$) can be compared from a geometric point of view using Figure 3. It can be seen that the clean time series is closer to the output obtained using the dynamical coupling procedure than it is to the output obtained by processing the target time series alone.

C.   Statistical properties of the outputs of the algorithm

In Figure 4 we have plotted the two dimensional delay embeddings of each of the coordinates of the outputs obtained with the dynamical coupling procedure, corresponding to two time series $\mathbf{X}^1$ and $\mathbf{X}^2$ from observations of the Hénon map using the linear observables $h^1(y_1, y_2) = 2y_1 + y_2$ and $h^2(y_1, y_2) = y_1 + 2y_2$, respectively, and corrupted by noises with very different noise to signal standard deviation ratios: 50% and 5%, respectively. The aim is to visualize the statistical properties of the outputs obtained with the dynamical coupling procedure in each of these two scalar time series.

We have plotted the two dimensional delay embedding of the outputs corresponding to the target time series $\mathbf{X}^1$ (Figures 4(a) and 4(b)) and to the dynamically coupled time series $\mathbf{X}^2$ (Figures 4(c) and 4(d)), together with the two delay embeddings of the corresponding clean and noisy time series. In order to analyze the dependence of the outputs on the number of points of the time series, we present in (a) and (c) the results for short time series ($L = 200$ time

series of only $N = 250$ data points) and for long time series ($L = 10$ time series of $N = 5000$ data points) in (b) and (d). The outputs for both time series seem to be unbiased estimators of the corresponding clean time series, and the outputs tend to concentrate quickly around the clean time series when the lengths of the time series increase. These greater geometric accuracies are also supported for the values of $< R^j >$, s_{R^j}, $j = 1, 2$ reported in the caption of the figure. The results for the time series $\mathbf{X}^1$ corrupted by 50% noise (target time series) are amazing. Even for short time series, the outputs seem to recover some of the geometric structure of the clean time series, and for long time series the mean noise reduction values are very high and the standard deviations small (see these numerical values in the caption of the figure). The low values of the mean noise reduction and the high values of the standard deviation obtained for the time series $\mathbf{X}^2$ clearly confirm that the dynamical coupling is not advantageous in reducing the noise of this more accurate time series, especially for $N = 250$, and therefore it is better to process the time series $\mathbf{X}^2$ alone. However it is remarkable that the mean noise reduction value $< R^2 >$ increases and the standard deviation s_{R^2} decreases quickly when the length N of the time series increases, and also that the dynamical coupling procedure does not seem to distort the signal heavily, even when the dynamically coupled time series $\mathbf{X}^1$ has 10 times its level of uncertainty.

D. Higher dimensional dynamics

We present in Table IV and Figure 5 some results for the Lorenz dynamics with linear observables. The aim is to show how the dynamical coupling procedure behaves for a three dimensional continuous dynamical system.

The target time series is a noisy scalar time series corresponding to Lorenz dynamics, the linear observable $h^1(y_1, y_2, y_3) = \frac{-y_1 + 6y_2 + 3y_3}{4}$, and a noise level of $NRS^1 = 50\%$. The dynamically coupled time series correspond to the observables $h^2(y_1, y_2, y_3) = \frac{-y_1 + 2y_2 + 7y_3}{4}$ and $h^3(y_1, y_2, y_3) = \frac{-y_1 + 6y_2 + 3y_3}{4}$, with respective initial noise levels $SNR^2 = 5\%$ and $SNR^3 = 10\%$. The results (see Table IV) with the dynamical coupling procedure are significantly better than those obtained by processing the nosiest time series separately, and the advantage is maintained also for long time series.

In Figure 5 we show the three delay embedding of one of the $L = 200$ scalar noisy time series $\mathbf{X}^1$ of only $N = 250$ data points, together with the corresponding clean time series $\mathbf{s}^1$, and the time series $\widehat{\mathbf{s}}^1$ that the algorithm gives as output using the two additional dynamical coupled time series. This figure illustrates that even for a short time series with such a large noise level, the dynamical coupling procedure allows recovery of some of the geometric and dynamical properties of the clean time series. We can see in Table IV that for a length of only 250 data points the

mean noise reduction level obtained is 58.9386%.

1. Dynamical coupling with other noise reduction algorithms

In order to give an indication of the effectiveness of the dynamical coupling scheme when using other noise reduction algorithms, we have run the algorithm implemented in the TISEAN[33] package with the name *ghkss* on one of the examples of Table IV, corresponding to the length $N = 5000$. This algorithm can be applied to both scalar and multivariate time series, using in both cases the Euclidean metric. Without using the dynamical coupling procedure, and taking a $dm = 7$ delay embedding of the target time series as the input time series, there results a noise reduction of $R = 54.3435$. If the target time series is processed together with the two dynamically coupled time series, then the resulting noise reduction is $R = 74.8087$, which is significantly higher than the noise reduction obtained by processing the target time series separately.

E. Estimation of the covariance matrix of the errors

For all the examples, we have estimated the covariance matrix of the unknown error terms, at any iteration of the algorithm, as we have explained in Subsection II C. In this subsection we illustrate how the algorithm behaves in the estimation of the initial covariance matrix of the errors, and in the estimation

$$\widehat{NSR}^j := \sqrt{\frac{\sum_{i=1}^{N} \left(\widehat{\epsilon}_i^j - \langle \widehat{\epsilon}^j \rangle \right)^2}{\sum_{i=1}^{N} \left(\widehat{s}_i^j - \langle \widehat{s}^j \rangle \right)^2}}, \ j = 1, ..., d, \tag{3}$$

of the initial noise to signal ratios NSR^j, $j = 1, ..., d$, where $\widehat{s}_i^j$ denotes the estimation of the clean data s_i^j at the eighth iteration and $\widehat{\epsilon}_i^j := X_i^j - \widehat{s}_i^j$ (compare $\widehat{NSR}^j$, $j = 1, ..., d$ with the true values $NSR^j, j = 1, ..., d$ defined in (2)). The three noisy dynamically coupled scalar time series $\mathbf{X}^1$, $\mathbf{X}^2$, and $\mathbf{X}^3$ of this experiment correspond to the logistic map, the observables $h^1(y) = \sqrt{5y + 4}$, $h^2(y) = (y + 2)^2$, and $h^3(y) = 2y + 3$, respectively; to the noise to signal ratios 10%, 5%, and 7.5%, respectively; and to a length $N = 5000$. The

true covariance matrix of the errors is $\Sigma^0 = \begin{pmatrix} 1.2531 \times 10^{-3} & 1.0103 \times 10^{-5} & -2.8674 \times 10^{-5} \\ 1.0103 \times 10^{-5} & 7.6455 \times 10^{-3} & -5.8485 \times 10^{-5} \\ -2.8674 \times 10^{-5} & -5.8485 \times 10^{-5} & 2.8445 \times 10^{-3} \end{pmatrix}$ and the es-

timation obtained from the sample covariance matrix of the corrections made by the algorithm in eight itera-

tions is $\widehat{\Sigma}^0 = \begin{pmatrix} 1.2616 \times 10^{-3} & -9.7567 \times 10^{-5} & -2.1806 \times 10^{-5} \\ -9.7567 \times 10^{-5} & 6.2035 \times 10^{-3} & -3.9168 \times 10^{-4} \\ -2.1806 \times 10^{-5} & -3.9168 \times 10^{-4} & 2.7860 \times 10^{-3} \end{pmatrix}$. The true noise to signal ratios are $NSR =$

$(10.0059\%, 4.9148\%, 7.5070\%)$ and the estimations are $\widehat{NSR} = (10.0520\%, 4.7969\%, 7.4322\%)$. Considering that $\widehat{\Sigma}^0$

and $\widehat{NSR}^j$, $j = 1, 2, 3$, are obtained using only eight iterations of the algorithm, they are reasonable estimations for

the unknown Σ^0 and NSR^j, $j = 1, 2, 3$.

F. Effectiveness of dynamical coupling with other measures of noise reduction

In all the examples above we have quantified the noise reduction achieved by the algorithm using a measure

based on the pointwise distance to the clean time series. The other standard way of measuring noise reduction is

based on the dynamical properties, in particular the forecasting power, of the time series before and after the noise

reduction[7,13,14,17–21,24,30–32]. If we denote by $\mathbf{f}^*$ the shift dynamics for the clean time series (i.e. $\mathbf{f}^*(\mathbf{s}_i) := \mathbf{s}_{i+1}$), then

a measure of how well the cleaned time series represents the dynamics is

$$e_{dyn}(\widehat{\mathbf{s}}) := \left(\frac{1}{N-1} \sum_{i=1}^{N-1} (\widehat{\mathbf{s}}_{i+1} - \mathbf{f}^*(\widehat{\mathbf{s}}_i))^2 \right)^{1/2},$$

and a measurement of the level of noise reduction achieved by a noise reduction algorithm is obtained by comparing

$e_{dyn}(\widehat{\mathbf{s}})$ with $e_{dyn}(\mathbf{X})$, taking for instance $R_{dyn} := 100(1 - \frac{e_{dyn}(\widehat{\mathbf{s}})}{e_{dyn}(\mathbf{X})})$ (we can define an analogous measure of the noise

reduction in each of the d components of the cleaned time series). If $\mathbf{f}^*$ is unknown then an analogous measure of

noise reduction $\widehat{e}_{dyn}$ can be obtained by replacing the unknown shift dynamics $\mathbf{f}^*$ by its local linear estimation $\widehat{\mathbf{f}}_i^*$ at

$\mathbf{s}_i$ in each of the terms of the sum defining e_{dyn}. Then

$$\widehat{R}_{dyn} := 100(1 - \frac{\widehat{e}_{dyn}(\widehat{\mathbf{s}})}{\widehat{e}_{dyn}(\mathbf{X})}) \tag{4}$$

is a measure of the noise reduction that requires knowledge of neither the clean time series nor the dynamics.

The two standard measures of noise reduction (the based on the pointwise distance d_p and the dynamical distance

e_{dyn}), might not guarantee the efficacy of the cleaned time series for the estimation of every geometric, dynamical,

or statistical feature of the underlying dynamics. In Ref.[38], relating to applications of noise reduction to clinical

measurements, it is reported that some features of clinical relevance contribute little to the variance of the signal,

so a small pointwise distance to the clean time series does not guarantee small distortion of the clinically important

features. The opposite situation is also possible: in Ref.[4,7] the authors have pointed out that a noise reduction

algorithm can shift the data slightly in a systematic way, or that the cleaned time series could be considered as a cleaner data set for a slightly different value of the parameters in the generating model, resulting in either case in large values of d_p and e_{dyn}. One of the alternatives they propose is to consider the short-term predictability of the data using any nonlinear predictor (this is also the spirit of $\widehat{e}_{dyn}$). Even a third possible situation is reported in Ref.[31]: algorithms using respectively the pointwise and the dynamical distance resulted in very different evaluations of the noise reduction achieved. For this reason we proposed in that paper two other measures of noise reduction that use a Hausdorff-like distance d_h instead of the pointwise distance. The d_h-distance to the clean time series of the cleaned time series is defined by

$$d_h(\widehat{\mathbf{s}}) := \max_{i=1,\ldots,N} \min_{j=1,\ldots,N} \|\mathbf{s}_j - \widehat{\mathbf{s}}_i\|,$$

and $R_h := 100(1 - \frac{d_h(\widehat{\mathbf{s}})}{d_h(\mathbf{X})})$ is the d_h-based measure of the percentage of noise reduction. By the definition of d_h, poor performance of a noise reduction scheme at even a single point can perturb badly the global noise reduction measure R_h. Conversely, a high level of noise reduction in an R_h sense gives strong evidence of the convergence of the support of the empirical measure associated to the cleaned time series to that of the underlying invariant measure. This might be useful, for instance, in the case of clinical data where outliers can be very relevant.

The second noise reduction measure $R_{<h>}$ we proposed in Ref.[31] avoids the drawback pointed out for R_h by replacing d_h by $d_{<h>}$ where

$$d_{<h>}(\widehat{\mathbf{s}}) := \frac{1}{N} \sum_{i=1}^{N} \min_{j=1,\ldots,N} \|\mathbf{s}_j - \widehat{\mathbf{s}}_i\|.$$

A joint analysis of the values R_h and $R_{<h>}$ might be used to detect bad behavior of the noise reduction algorithm localized at a small number of points. These measures of noise reduction R_h and $R_{<h>}$ require the use of some technique of fast neighbor search such as the box-assisted method[39] in order to make the algorithm more efficient. We also use such a technique in the construction of the neighborhoods.

Finally, the most important drawback of noise reduction measures based on the pointwise distance or on e_{dyn} is that they require the knowledge of either the clean time series or the shift dynamics $\mathbf{f}^*$, which are unknown in most applications (for instance in time series coming from clinical data such as electrocardiograms or electroencephalo-grams). Noise reduction measures which do not require knowledge of the time series are proposed in Ref.[7] (see also the references therein) and in Refs.[29,40]. Furthermore, there is an extensive literature[41] about different techniques for estimating the noise level in time series (through correlation dimension estimations, entropies, recurrence plots, false neighborhoods, etc.), and many of these techniques can be used for quantifying the noise reduction by comparing the

estimated noise levels of the time series before and after the noise reduction.

Table V reports the levels of noise reduction obtained with our algorithm using measures of noise reduction other than the one based on the pointwise distance ($< R_{dyn} >$, $< \widehat{R}_{dyn} >$, $< R_h >$ and $< R_{<h>} >$). We have designed an experiment where the advantage of the dynamical coupling procedure is subtle, in order to check which of these measures can detect this subtle advantage. The target time series has a medium noise level of $NSR^1 = 10\%$, and corresponds to a Hénon map, the nonlinear observable $h^1(y_1, y_2) = \ln(4 + y_1)$, and a length $N = 5000$. In the first case this target time series is processed with only one dynamically coupled noisy time series corresponding to the nonlinear observable $h^2(y_1, y_2) = (2y_1 + 5)^2$ and a noise level $NSR^2 = 5\%$. In the second case we add a second dynamically coupled time series with a noise level $NSR^3 = 7.5\%$ and corresponding to the linear observable $h^3(y_1, y_2) = 2y_1 + 3$. The mean values of noise reduction obtained with $L = 10$ time series of length $N = 5000$ that appear in Table V show, for all these noise reduction measures, higher values of noise reduction when using the additional information provided by the third of the dynamically coupled time series. Although they are not reported in this paper, this agrees with the results we have obtained in all the previous examples, so we infer that all these measures are able to detect the effectiveness of the dynamical coupling in noise reduction.

IV. CONCLUSIONS

This paper shows the advantages of sharing and recycling all the information available from different research teams or institutions. The advantage is clear for the owners of the most uncertain information. The owners of the information with smaller uncertainty may also benefit from the sharing of information in the case of short length time series or when their levels of uncertainty are only slightly smaller.

Moreover, the method we propose allows a firm or institution to asses the value of new information in terms of the shadow price of the noise reduction level. This shadow price is the maximum monetary quantity that a buyer would agree to pay for the extra reduction in uncertainty achieved with the new information. For instance, as can be seen in Tables I and II, for short length noisy logistic time series with $NSR = 10\%$, the mean noise reduction level with only one observable is 48.76%. If we use one additional noisy time series dynamically coupled with the first with $NSR = 5\%$, then the mean noise reduction achieved is 72.16%, and if we use two dynamically coupled time series, being the noise to signal ratios of 5% and 7.5% respectively, then the mean noise reduction achieved is 73.1%. The shadow price of the information provided by these new time series is the quantity the buyer is willing to pay for this extra noise reduction. We can also see in the mentioned tables that for long time series the gains in noise reduction

with new information decrease drastically, so the shadow price will be lower for long time series.

In order to asses such shadow price, the final level of noise reduction must be quantified. In this paper the noise reduction levels are calculated using the information provided for the clean time series, which in real situations is unknown. However, there exist several methods that allows to estimate such noise reduction levels without using the information provided for the clean time series (see Ref.[4,41]).

As a consequence of the results of this paper the algorithms for noise reduction for multivariate time series acquire a relevance similar to those designed for scalar time series, which are pervasive in the noise reduction literature. The dynamical coupling procedure can be implemented using any noise reduction algorithm for multivariate time series (see Ref.[30,33] for instance), and yields better results than those that would be obtained where each time series is processed separately. However, if there can be obtained a dynamically coupled time series with a noise level significantly lower than that of the target time series, our algorithm is preferable because it takes into account the noise levels in each of the time series used in the noise reduction scheme.

The dynamical coupling procedure is especially useful for short time series, and surprisingly this good behavior does not decay seriously for higher dimensional dynamics (see for instance the results for the Lorenz dynamics). Since these are common features in social sciences - high dimensional dynamics, short time series, and an abundance of complementary information - we hope that the dynamical coupling method will soon have applications in these areas.

* Electronic address: mera@ccee.ucm.es,mmoranca@ccee.ucm.es

[1] T. I. Netoff, T. L. Carroll, L. M. Pecora, and S. J. Schiff. "Detecting coupling in presence of noise and nonlinearity", in "*Handbook of Time Series Analysis.*" Edited by B. Scheler, M. Winterhalder, and J. Timmer. (Wiley-VCH, 2006).

[2] G. A. Milliken, and D. E. Johnson. "*Analysis of Messy Data. Volume III: Analysis of Covariance*" (Chapman & Hall/CRC, 2002).

[3] W. H. Greene. "*Econometric Analysis.*"(Prentice Hall, 2007).

[4] H. Kantz and T. Schreiber. "*Nonlinear time series analysis.*" (Cambridge University Press, 2004).

[5] J. C. Sprott. "*Chaos and Time-Series Analysis.*" (Oxford University Press, 2004).

[6] C. Diks, "*Nonlinear time series analysis.*" (World Scientific, 1999).

[7] E. J. Kostelich and T. Schreiber. "Noise reduction in chaotic time-series data: A survey of common methods." Physical Review. E **48**, 1752 (1993).

[8] M. Casdagli. "Chaos and Deterministic versus Stochastic Non-linear Modelling." Journal of the Royal Statistical Society. Series B **54**, 2, 303-328 (1991).

[9] W. A. Barnett, A. R. Gallant, M. J. Hinich, J. A. Jungeilges, D. T. Kaplan, and M. J. Jensen. "A single-blind controlled competition among test of nonlinearity and chaos." Journal of Econometrics **82**, 1, 157-192 (1997).

[10] L. M. Pecora, L. Moniz, J. Nichols and T. L. Carroll. "A unified approach to attractor reconstruction." Chaos **17**, 013110 (2007).

[11] L. M. Pecora, T. L. Carroll, and J. F. Heagy. "Statistics for mathematical properties of maps between time series embeddings." Physical Review E **52**, 4, 3420-3439 (1995).

[12] M.E. Mera and Morán. "Conververgence of the Eckmann and Ruelle algorithm for the estimation of Lyapunov exponents." Ergodic Theory and Dynamical Systems, **20**, 531-546 (2000).

[13] E. J. Kostelich and J. A. Yorke. "Noise reduction in dynamical systems." Physical Review A **38**, 3, 1649-1652 (1988).

[14] E. J. Kostelich and J. A. Yorke. "Noise reduction: Finding the simplest dynamical system consistent with the data." Physica D **41**, 183 (1990).

[15] S. M. Hammel. "A noise reduction method for chaotic systems." Physics Letters A **148**, 421-428 (1990).

[16] J. D. Farmer and J. Sidorowich. "Optimal shadowing and noise reduction." Physica D **47**, 373-392 (1991).

[17] T. Schreiber and P. Grassberger. "A simple noise-reduction method for real data." Physics Letters A **160**, 5, 411-418 (1991).

[18] T. Schreiber. "Extremely simple nonlinear noise-reduction method." Physical Review E **47**, 4, 2401-2404 (1993).

[19] M. Davies. "Noise reduction schemes for chaotic time series." Physica D **79**, 174-192 (1994).

[20] D. M. Walker and A. I. Mees. "Noise reduction of chaotic systems by Kalman filtering and by shadowing." International Journal of Bifurcations and Chaos **7**, 3, 769-779 (1997).

[21] P. Grassberger, R. Hegger, H. Kantz, C. Schaffrath and T. Schreiber. "On noise reduction methods for chaotic data." Chaos **3**, 2, 127-141 (1993).

[22] A. Kern, W.-H. Steeb and R. Stoop. "Projective noise cleaning with dynamic neighborhood selection." International Journal of Modern Physics C **11**, 1, 125-146 (2000).

[23] L. Matassini, H. Kantz, J. Holyst, and R. Hegger. "Optimizing of recurrence plots for noise reduction." Physical Review E **65**, 021102 (2002).

[24] A. Leontitsis, T. Bountis and J. Pagge. "An adaptive way for improving noise reduction using local geometric projection." Chaos **14**, 1, 106-110 (2004).

[25] K. Urbanowicz and J.A. Holyst. "Noise reduction in chaotic time series by a local projection with nonlinear constraints." Acta Physica Polonica B **35**, 9, 2175-2197 (2004).

[26] M.T. Johnson and R.J. Povinelli. "Generalized phase space projection for nonlinear noise reduction." Physica D **201**, 306-317 (2005).

[27] X. Luo, J. Zhang and M. Small. "Optimal phase-space projection for noise reduction." Physical Review E **72**, 046710 (2005).

[28] T. Sauer. "A noise reduction method for signals from nonlinear systems." Physica D, **58**, 1-4, 193-201 (1992).

[29] R. Cawley and G-H. Hsu. "Local-geometric-projection method for noise reduction in chaotic maps and flows." Physical Review A **46**, 6, 3057-3082 (1992).

[30] R. Hegger and T. Schreiber. "A noise reduction method for multivariate time series." Physics Letters A **170**, 305-310 (1992).

[31] M. E. Mera and M. Morán. "Geometric noise reduction for multivariate time series." Chaos, **16**, 013116 (2006).

[32] M. E. Mera and M. Morán. "Reduction of noise of large amplitude through adaptive neighborhoods." Physical Review E **80**, 016207 (2009).

[33] R. Hegger, H. Kantz, and T. Schreiber. "Practical implementation of nonlinear time series methods: The TISEAN package." Chaos **9**, 2, 413-435 (1999).

[34] W. A. Fuller, "*Measurement error models*". Wiley Series in Probability and Statistics. (John Wiley & Sons, 2006).

[35] J. Bröcker and U. Parlitz. "Efficient noncausal noise reduction for deterministic time series." Chaos **11**, 2, 319-326 (2001).

[36] L. Jaeger and H. Kantz. "Unbiased reconstruction of the dynamics underlying a noisy chaotic time series." Chaos **6**, 449 (1996).

[37] F. Takens "*Dynamical Systems and turbulence.*" Edited by D. A. Rang and L. -S. Young, Lecture Notes in Mathematics **898**, 366-381. (Springer, Berlin, 1981).

[38] T. Schreiber and D.T. Kaplan. "Nonlinear noise reduction for electrocardiograms." Chaos **6**, 1, 87-92 (1996).

[39] P. Grassberger, T. Schreiber and C. Schaffrath. "Nonlinear time sequence analysis." International Journal of Bifurcation and Chaos **1**, 3, 521-547 (1991).

[40] W. Orzeszko. "The new method of measuring the effects of noise reduction in chaotic data." Chaos, Solitons and Fractals **38**, 1355-1368 (2008).

[41] M. E. Mera and M. Morán. "Estimating measurement noise level through a noise reduction algorithm." in preparation.

TABLE I: *Noise reduction effectiveness using two and three dynamically coupled time series.*

Sample mean noise reduction level $< R^j >$, and sample standard deviation s_{R^j} of R^j, $j = 1, 2, 3$ (in parentheses) obtained using the dynamical coupling procedure with noisy logistic time series. On the left of the table there are the results using two noisy dynamically coupled time series, and on the right using three noisy dynamically coupled time series. The length of the time series is N, and L is the number of time series used to calculate $< R^j >$ and s_{R^j}, $j = 1, 2, 3$. The dynamically coupled noisy time series correspond to the logistic map, the observables $h^1(y) = \sqrt{5y + 4}$, $h^2(y) = (y + 2)^2$, and $h^3(y) = 2y + 3$ respectively, and Gaussian additive measurement noises with NSRs of 10%, 5%, and 7.5%, respectively.

(N, L)	Two observables $NSR = (10\%, 5\%)$		Three observables $NSR = (10\%, 5\%, 7.5\%)$		
	$< R^1 >$ (s_{R^1})	$< R^2 >$ (s_{R^2})	$< R^1 >$ (s_{R^1})	$< R^2 >$ (s_{R^2})	$< R^3 >$ (s_{R^3})
(250, 200)	72.1558 (3.1759)	57.2543 (4.3878)	73.0973 (3.2823)	59.2482 (4.6489)	69.2822 (3.2649)
(500, 100)	78.2957 (2.2286)	64.0965 (2.9998)	80.2954 (2.2635)	69.0632 (3.1964)	77.4221 (2.2967)
(1000, 50)	81.4837 (1.7280)	67.3679 (2.7840)	83.4224 (1.8337)	72.6099 (2.4393)	80.4309 (1.7521)
(2500, 20)	84.2583 (1.3913)	70.9704 (2.0683)	85.5285 (1.2452)	75.0332 (2.0689)	82.3555 (1.3295)
(5000, 10)	85.2670 (0.8255)	72.0466 (1.1132)	86.4148 (0.9055)	75.3407 (1.5972)	83.0978 (0.9629)

TABLE II: *Noise reduction effectiveness without using the dynamical coupling procedure.*

Sample mean noise reduction levels $< R^j >$ and sample standard deviations s_{R^j}, $j = 1, 2$ (in parentheses) obtained without the dynamical coupling procedure for the first two noisy scalar time series $\mathbf{X}^1$ and $\mathbf{X}^2$ of Table I. The results for $\mathbf{X}^1$ are on the left of the table, and the results for $\mathbf{X}^2$ are on the right of the table. The input time series are dm-dimensional delay embeddings of each of these two noisy scalar time series. The noisy time series $\mathbf{X}^1$ and $\mathbf{X}^2$ correspond to the logistic map, the observables $h^1(y) = \sqrt{5y + 4}$, and $h^2(y) = (y+2)^2$, and Gaussian additive measurement noises with NSRs of 10% and 5%, respectively.

(N, L)	$< R^1 >$ for h^1 and $NSR^1 = 10\%$ (s_{R^1})			$< R^2 >$ for h^2 and $NSR^2 = 5\%$ (s_{R^2})		
	$dm = 1$	$dm = 2$	$dm = 3$	$dm = 1$	$dm = 2$	$dm = 3$
(250, 200)	48.7575	39.2307	26.7612	49.6479	46.2020	
	(4.8831)	(7.7380)	(7.3026)	(4.2419)	(6.6799)	
(500, 100)	54.6265	53.1635	35.2890	58.4035	52.7725	
	(3.3760)	(4.5989)	(5.2448)	(3.2405)	(4.8983)	
(1000, 50)	55.6982	68.0615	50.9355	62.0200	66.7358	46.9185
	(2.8036)	(2.6997)	(4.6053)	(2.4349)	(4.4920)	(8.0153)
(2500, 20)	56.3709	73.9164	71.4198	66.4834	75.2390	63.4777
	(1.2945)	(1.2800)	(2.1565)	(2.2776)	(1.4355)	(4.0926)
(5000, 10)	58.3522	73.9624	74.4198	66.4813	76.5039	75.4329
	(1.5633)	(0.9374)	(1.2380)	(1.1258)	(1.1027)	(1.6512)

TABLE III: *Efficacy of the dynamical coupling for logistic time series corrupted by highly heteroscedastic noises.*

Sample mean noise reduction levels $< R^j >$ and sample standard deviations s_{R^j} of $R^j, j = 1, 2$ (in parentheses) obtained with L noisy logistic time series $\mathbf{X}^1$ of length N corresponding to the observable $h^1(y) = \sqrt{5y + 4}$ and a noise level of $NSR^1 = 50\%$, with and without the dynamical coupling procedure. The left side of the table shows the results obtained using one dynamically coupled time series $\mathbf{X}^2$ corresponding to the observable $h^2(y) = (y + 2)^2$ and a noise level of $NSR^2 = 5\%$. The right side of the table shows the results for the target time series $\mathbf{X}^1$ without using the dynamically coupled time series, and taking as the input time series a dm-dimensional delay embedding of $\mathbf{X}^1$ with $dm = 1, 2, 3$.

	Dynamical coupling Two observables (h^1, h^2) $NSR = (50\%, 5\%)$		Without dynamical coupling $< R^1 >$ for h^1 and $NSR^1 = 50\%$ (s_{R^1})		
(N, L)	$< R^1 >$ (s_{R^1})	$< R^2 >$ (s_{R^2})	$dm = 1$	$dm = 2$	$dm = 3$
$(250, 200)$	81.4044 (3.3442)	50.9736 (5.1603)	35.3211 (5.4111)	42.5436 (7.6653)	36.8192 (8.2089)
$(500, 100)$	86.1104 (2.2790)	58.7388 (3.2391)	36.7394 (4.3142)	44.5127 (6.0664)	40.4621 (5.7711)
$(1000, 50)$	89.1018 (1.5642)	61.7666 (3.8835)	36.1591 (2.9624)	45.6408 (4.7604)	41.2277 (3.860)
$(2500, 20)$	91.9465 (0.9764)	65.1723 (4.3370)	35.9484 (1.7779)	47.2789 (2.7879)	45.9071 (2.6735)
$(5000, 10)$	93.1878 (0.5395)	64.0106 (1.5160)	29.6034 (1.2317)	47.5461 (2.4870)	46.7443 (2.3428)

TABLE IV: *Efficacy of the dynamical coupling for Lorenz time series corrupted by highly heteroscedastic noises.*

Sample mean noise reduction level $< R^1 >$ and sample standard deviation s_{R^1} (in parenthesis) obtained using L noisy Lorenz time series of length N corresponding to the observable $h^1(y_1, y_2, y_3) = \frac{3y_1 + 2y_2 + 3y_3}{4}$ and $NSR^1 = 50\%$, with and without the dynamical coupling procedure. In the first case the noisy time series is processed together with two additional noisy time series with which the target time series is dynamically coupled. Such coupled time series correspond to the observables and levels of noise $h^2(y_1, y_2, y_3) = \frac{-y_1 + 2y_2 + 7y_3}{4}$ with $NSR^2 = 5\%$ and $h^3(y_1, y_2, y_3) = \frac{-y_1 + 6y_2 + 3y_3}{4}$ with $NSR^3 = 10\%$, respectively. In the second case the input time series is a 7-delay embedding of the noisy scalar time series.

	$< R^1 >$ and s_{R^1} for noisy Lorenz time series				
(N, L)	$(250, 200)$	$(500, 100)$	$(1000, 50)$	$(2500, 20)$	$(5000, 10)$
Dynamical coupling	58.9386	71.9751	76.7067	81.5473	84.2294
$NSR = (50\%, 5\%, 10\%)$	(4.2907)	(3.3092)	(2.8505)	(2.0410)	(1.1641)
Without dynamical coupling	47.4752	50.3395	51.9087	52.7861	53.4961
$NSR = 50\%;\ dm = 7$	(3.6892)	(3.4044)	(2.5399)	(1.8111)	(1.5803)

TABLE V: *Effectiveness of the dynamical coupling procedure using noise reduction measures other than that based on the pointwise distance.*

Mean noise reduction measures and sample standard deviations (in parentheses) obtained using $L = 10$ noisy time series corresponding to the Hénon map, the observable $h^1(y_1, y_2) = \ln(4 + y_1)$, $NSR^1 = 10\%$, and $N = 5000$. The results in the top part of the table are obtained by processing the target time series with one dynamically coupled noisy time series corresponding to the nonlinear observable $h^2(y_1, y_2) = (2y_1 + 5)^2$ and a noise level $NSR^2 = 5\%$. The results in the bottom part are obtained by processing the target time series with an additional dynamically coupled time series corresponding to the linear observable $h^3(y_1, y_2) = 2y_1 + 3$ and $NRS^3 = 7.5\%$.

	$< R^1 >$	$< R^1_{dyn} >$	$< \widehat{R}^1_{dyn} >$	$< R_h >$	$< R_{<h>} >$
Two coupled time series	70.0265	80.40958	96.7622	62.3924	75.8628
	(0.7226)	(1.0398)	(0.6965)	(20.2805)	(3.3847)
Three coupled time series	73.0497	81.9136	98.0892	78.6847	89.4506
	(1.2218)	(1.27277)	(0.7942)	(10.5651)	(0.9819)

FIG. 1: *Summary of the results reported in Tables I and II corresponding to the target time series* $\mathbf{X}^1$.

Sample mean noise reduction levels $< R^1 >$ and sample standard deviations s_{R^1} (with error bars) obtained for the target time series $\mathbf{X}^1$:

(a) using two (in blue) versus three (in red) dynamically coupled time series (second and fourth columns of Table I, respectively).

(b) without using the dynamical coupling procedure, for different values of dm (second, third, and fourth columns of Table II).

(c) using three dynamically coupled time series (in red) versus without using the dynamical coupling procedure and $dm = 2$ (in blue) (fourth column of Table I and third column of Table II, respectively).
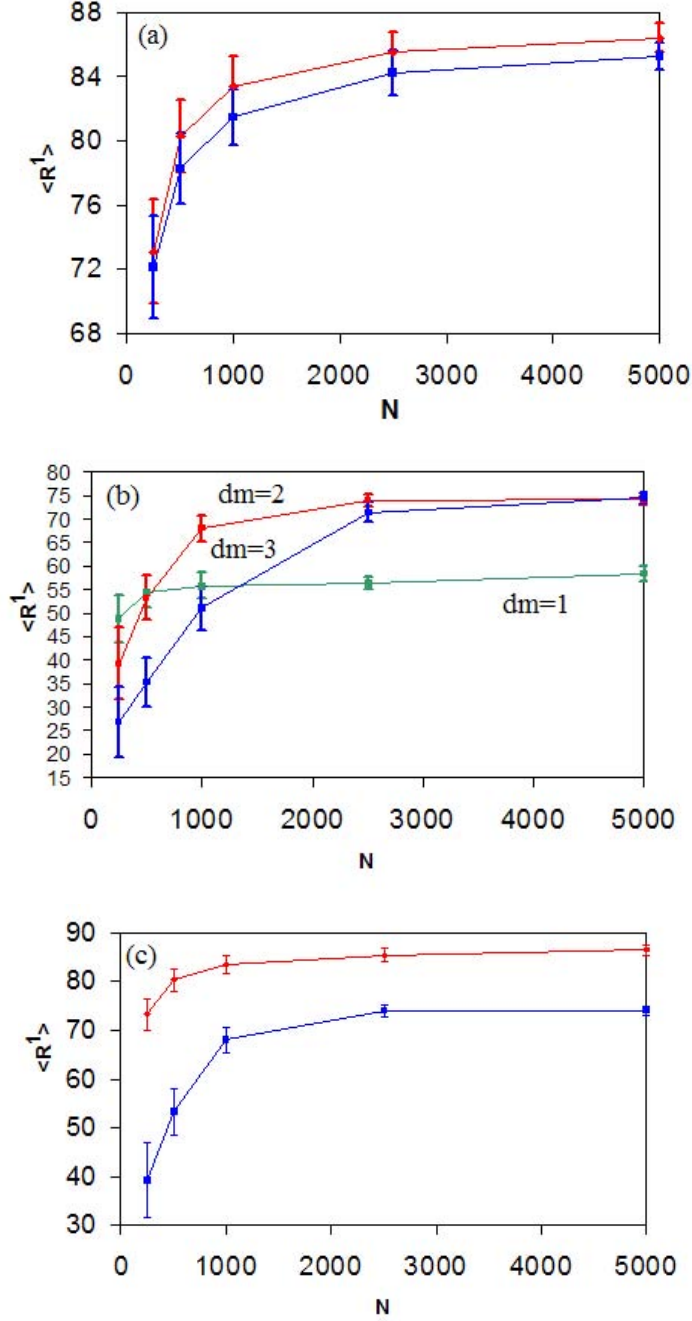
FIG. 2: *Summary of the results reported in Table III corresponding to the target time series* $\mathbf{X}^1$.

Sample mean noise reduction levels $< R^1 >$ and sample standard deviations s_{R^1} (with error bars) obtained for the target time series $\mathbf{X}^1$ using the dynamical coupling procedure versus without using the dynamical coupling procedure with $dm = 2$ (second and fifth columns of Table III).
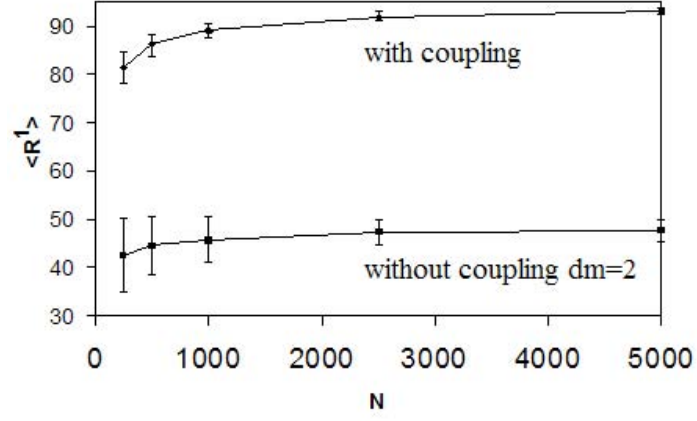
FIG. 3: *Output time series for a noisy logistic time series with and without the dynamical coupling procedure.*

Two dimensional delay embeddings of the target time series for the clean, noisy, and output time series obtained with and without the dynamical coupling procedure. The target noisy time series corresponds to the logistic map, the observable $h^1(y) = \sqrt{5y+4}$, and the noise level $NSR^1 = 50\%$. The time series in (a) are the clean and the noisy time series. The cleaned time series shown in (b) is obtained using our noise reduction algorithm with a dynamically coupled noisy logistic time series corresponding to the observable $h^2(y) = (y+2)^2$ and corrupted by a Gaussian noise of 5%. The cleaned time series shown in (c) is obtained without using the dynamical coupling procedure and taking as the input time series for the noise reduction algorithm the two-dimensional delay embedding of the target time series. The noise reduction levels are $R^1 = 93.9378$ for the cleaned time series represented in (b) and $R^1 = 50.1631$ for the cleaned time series represented in (c).
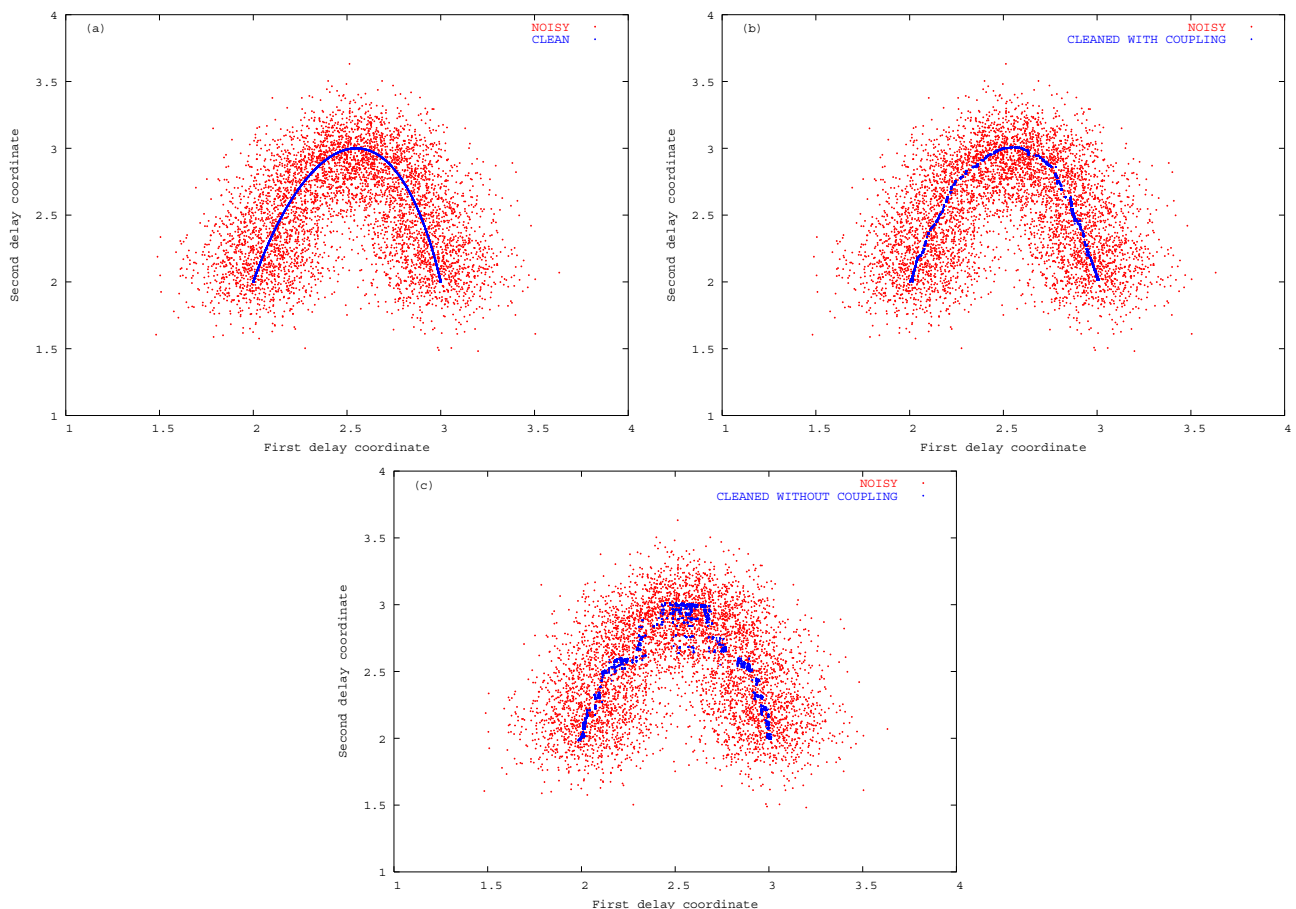
FIG. 4: *Distribution, according to the length N, of the outputs of the algorithm for two dynamically coupled time series.*

Two-dimensional delay embeddings of L noisy, clean, and cleaned time series of length N of each of two dynamically coupled time series. The input noisy time series correspond to the Hénon map, the observables $h^1(y_1, y_2) = 2y_1 + y_2$, $h^2(y_1, y_2) = y_1 + 2y_2$, and the noise levels $NSR^1 = 50\%$ and $NSR^2 = 5\%$ respectively. The cleaned time series are obtained using the dynamical coupling procedure.

(a) Observable h^1 with $NSR^1 = 50\%$, $N = 250$ and $L = 200$. $< R^1 > = 70.8096$ and $s_{R^1} = 3.7620$.

(b) Observable h^1 with $NSR^1 = 50\%$, $N = 5000$ and $L = 10$. $< R^1 > = 87.5648$ and $s_{R^1} = 1.1196$.

(c) Observable h^2 with $NSR^2 = 5\%$, $N = 250$ and $L = 200$. $< R^2 > = 22.1851$ and $s_{R^2} = 8.1841$.

(d) Observable h^2 with $NSR^2 = 5\%$, $N = 5000$ and $L = 10$. $< R^2 > = 54.1915$ and $s_{R^2} = 2.7755$.
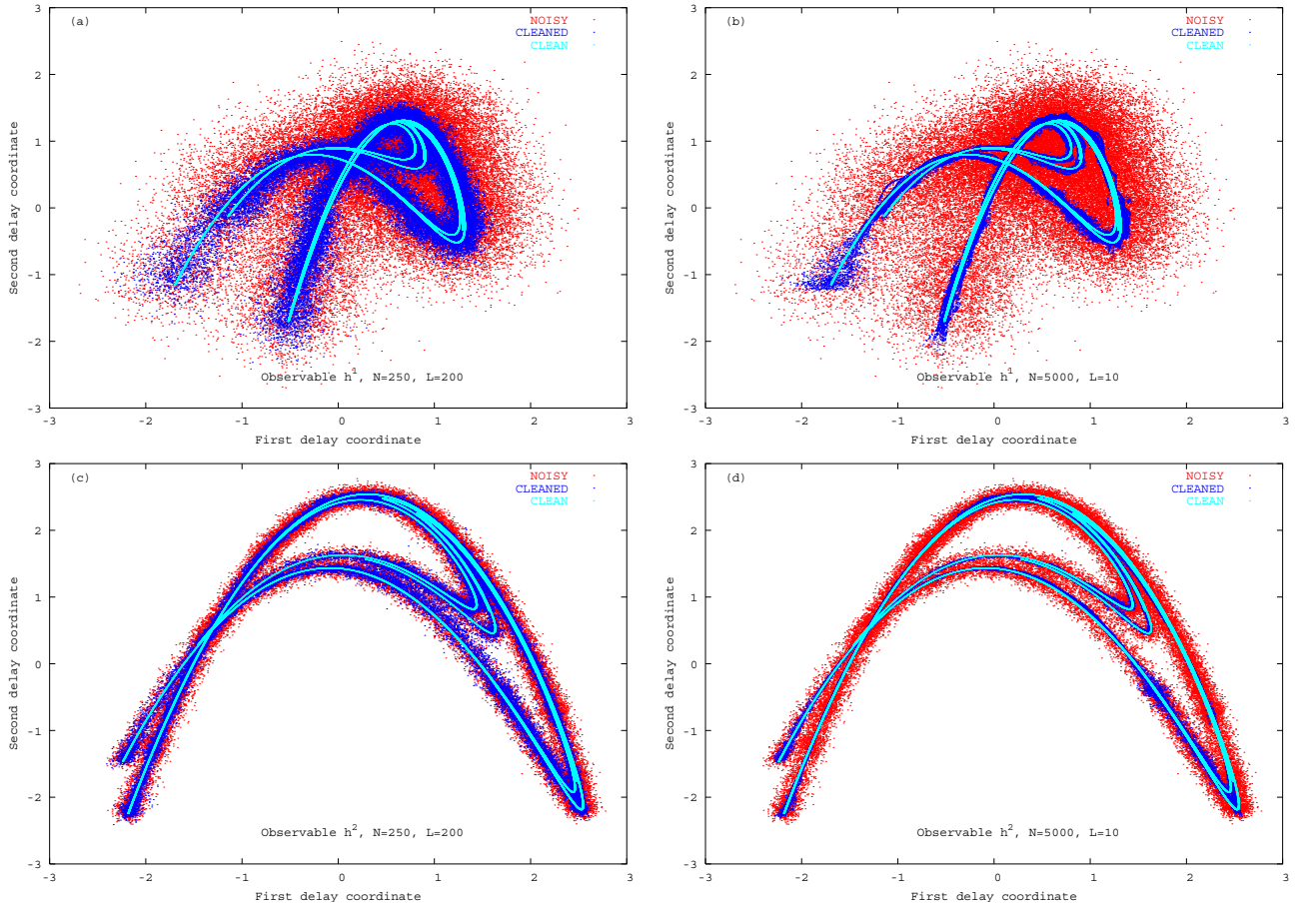
FIG. 5: *Output using the dynamical coupling scheme in a short noisy Lorenz time series.*

Three-dimensional delay embedding of a noisy (a), clean (b), and cleaned (c) time series using the dynamical coupling technique. The target noisy time series corresponds to the Lorenz dynamics, $N = 250$ data points, the observable $h^1(y_1, y_2, y_3) = \frac{-y_1 + 6y_2 + 3y_3}{4}$, and a noise level of $NRS^1 = 50\%$. The dynamical coupling procedure is made using two additional noisy time series, corresponding to the observables $h^2(y_1, y_2, y_3) = \frac{-y_1 + 2y_2 + 7y_3}{4}$, $h^3(y_1, y_2, y_3) = \frac{-y_1 + 6y_2 + 3y_3}{4}$, and the initial noise levels are $SNR^2 = 5\%$ and $SNR^3 = 10\%$. The level of noise reduction is $R^1 = 68.7497$.