Reply to the comments from Dr. Zhixin Lu:

This paper is in general well written with sufficient simulations that support its conclusions.

However, two main issues need to be addressed.

1. In Sec. 2.2, the authors introduce the reservoir computing method (Lu et al., 2017) for the variable reconstruction problem. However, I find this introduction very confusing. It seems that different constructions of reservoir computers for different tasks (for reservoir observer or for predicting future of time series) are introduced as different layers for a single reservoir. (lines 144-150). It is also confusing why one would need the so-called prediction reservoir as a layer for this reservoir observer task. (lines 175-178) Does this closed-loop reservoir really being used in the simulation in this paper? If so, why is it necessary? A reservoir observer does not need to feedback its own output to its input, as it is simply trying to estimate variable b(t) based on the measured a(t), rather than predicting the future of both a(t) and b(t).

Response: Thanks for your comments and suggestions**.** By means of the first two components shown in Figure 1*, the $a(t)$ is trained and then $\psi[r^*(t)]$ is obtained. In this procedure, the value of $\psi[r^*(t)]$ is already very close to the value of $b(t)$.

Then, if $\psi[r^*(t)]$ is feedback to function "*f*" and "ψ", this repetitive operation might make the value of $\psi[r^*(t)]$ more close to the value of $b(t)$. Actually we also found this repetitive operation no longer influenced the results. This is to say, that the third component shown in Figure 1* might be redundant in this reconstruction framework, and the first two components are enough. In the revised manuscript, we will carefully modify the diagram and the introduction of Reservoir computer according to the introduction in *Lu et al.* 2017 [1].

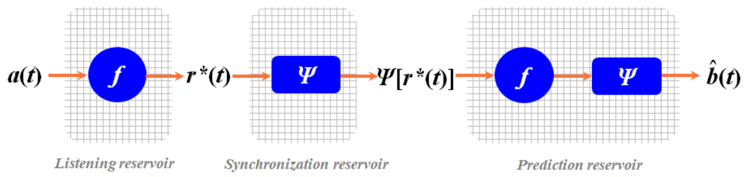

Figure 1* The schematic of Reservoir computer in the previous manuscript (we will revised this figure in the revised manuscript).

[1] Lu Z, Pathak J, Hunt B, Girvan M, Brockett R, Ott E. Reservoir observers: Model-free inference of unmeasured variables in chaotic systems. Chaos 27(4), 041102 (2017).

2. The authors in Sec. 3.2.1-3.2.2 discuss the nonlinear coupling relation, which is essentially the nonlinear observability in the control theory, as being pointed out in (Lu et al., 2017). This direction dependence can be explained by the nonlinear observability. For example, in the Lorenz 63 model, due to the symmetry of that ODE system, both (x(t), y(t), z(t)) and (-x(t),-y(t), z(t)) are solutions on the same chaotic attractor. Thus, one can not construct any nonlinear state-observer that estimates the value of x or y given the time series of variable z. However, a state observer can estimate z(t) given either x(t) or y(t). It was also shown that xˆ2(t) and yˆ2(t) can be estimated given z(t) as it is nonlinearly observable. The authors employ CCM to quantify the "nonlinear coupling relation" and show that it is better than a linear coupling relation. It is the reviewer's opinion that a brief discussion of the relation between the CCM and the nonlinear observability should be given. Is CCM essentially the same as nonlinear-observability? If not, what is the difference?

Response: Thanks for your comments and suggestions**.** Referred to the literature [1-6], we found that the meanings of "nonlinear observability" and "CCM" are partially close to each other:

"Nonlinear observability":

For two variables x_0 and x_1 , their time series follows that: $x_0(t) \in U$ and $x_1(t) \in \Sigma$. If they are from nonlinear systems, it is a general fact that Σ restricted to U is not necessarily complete [1]. Hence, *Hermann and Krener* 1977 [1] demonstrated that $x_0(t)$ might be not able to totally recover the values of $x_1(t)$. Then, the asymmetry reconstruction between $x_0(t)$ and $x_1(t)$ is common for nonlinear systems, which is also called "estimability" and is discussed in the previous papers [2-3].

"CCM":

The convergent cross mapping (CCM) coefficient is a kind of causality index [4]. Takens 1981 [5] proposed that: for two variables *x* and *y*, if *x* does influence *y* in the dynamical system, the value of *x* can be recovered from the records of *y*.

Further, *Sugihara et al.* 2012 [4] demonstrated this theorem of Takens determines the reconstruction between two variables: for two variables *x* and *y*, if *x* does influence *y* in the dynamical system (but *y* does not influence *x*), the information of *x* will be transferred into *y*, and so that the records of *y* will be able to recover the values of *x*. However, this information transfer between *x* and *y* is asymmetry, and then the reconstruction between *x* and *y* will be also asymmetry. Hence, the CCM index is proposed to measure such asymmetry information transfer between the observational variables [4, 6].

The "nonlinear observability" is often measured for the nonlinear system with known mathematical equation. For the observational records from real-world system without known mathematical equation, the "nonlinear observability" might be hard to be measured. However, the CCM coefficient can be used to measure asymmetry information transfer between the observational variables in different real systems [4, 6].

Additionally, we also used CCM to analyze the "nonlinear observability" in the Lorenz 63 system. As Figure 2* and table 1* show, when using $z(t)$ to reconstruct $x(t)$, the reconstructed series largely deviates from the real $x(t)$. However, when using using $z(t)$ to reconstruct $[x(t)]^2$ or $x(t)*y(t)$, the reconstruction errors are much smaller. As Table 1* shows, we measured the CCM coefficient for $z(t)$ and $x(t)$, $z(t)$ and $[x(t)]^2$, and $z(t)$ and $x(t)*y(t)$ respectively, they are equal to 0.03, 0.95, and 0.91 respectively. Such results of CCM coefficient are really close to the analysis of "nonlinear observability".

We will discuss such association between "nonlinear observability" and "CCM" in the revised manuscript.

[1] Hermann R, Krener A. Nonlinear controllability and observability. IEEE Transactions on automatic control, 22(5), 728-740 (1977).

[2] Lu Z, Pathak J, Hunt B, Girvan M, Brockett R, Ott E. Reservoir observers: Model-free inference of unmeasured variables in chaotic systems. Chaos 27(4), 041102 (2017).

[3] Schumann-Bischoff J, Luther S, Parlitz U. Estimability and dependency analysis of model parameters based on delay coordinates. Phys Rev E, 94(3), 032221 (2016).

[4] Takens, F.: Detecting strange attractors in turbulence. Dynamical Systems and Turbulence, Lecture Notes in Mathematics, 898, 366–381 (Springer Berlin Heidelberg) (1981).

[5] Sugihara, G, May R, Ye H, Hsieh CH, Deyle E, Fogarty M, Munch S. Detecting causality in complex ecosystems. Science, 338(6106), 496-500 (2012).

[6] Tsonis AA, Deyle ER, Ye H, Sugihara G. Convergent cross mapping: theory and an example. In Advances in Nonlinear Geosciences (pp. 587-600), Springer, Cham., (2018).

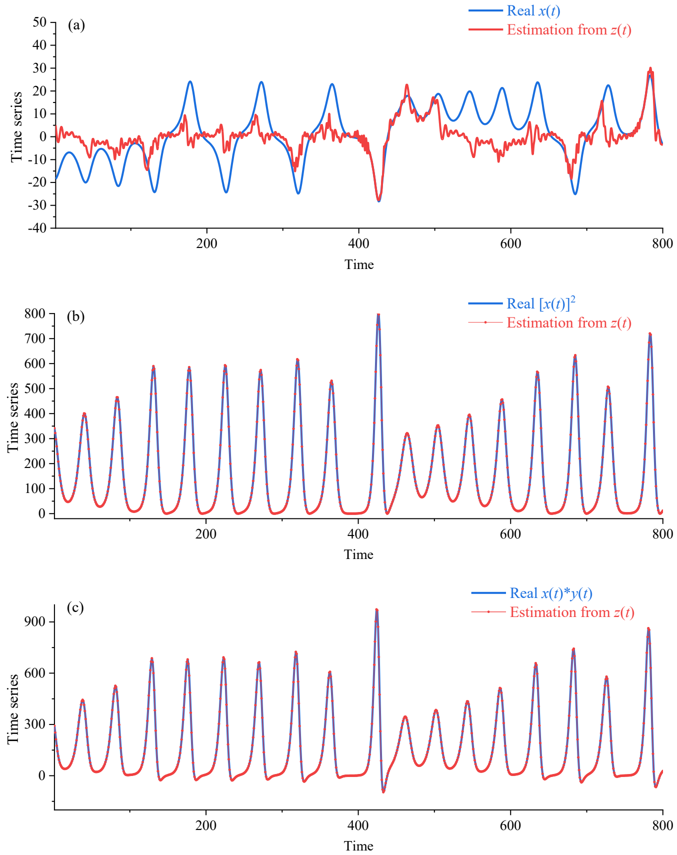

Figure 2* (a) The results of applying RC to reconstruct *x*(*t*) from *z*(*t*) (Lorenz 63 system). (b) The results of applying RC to reconstruct $[x(t)]^2$ from *z*(*t*). (c) The results of applying RC to reconstruct *x*(*t*)**y*(*t*) from *z*(*t*). The blue lines denote the real time series, and red lines represent the reconstructed series through the RC machine learning.

Table 1* Details of Lorenz63 system reconstruction

Input (*a*)	**Output (*b*)**	***CCM index*** $\rho_{a \to b}$	**Data length** (training/testing)	**Neural network**	**RMSE**
***Z*(*t*)**	***X*(*t*)**	**0.03**	2400/1600	**RC**	**1.13**
***Z*(*t*)**	$X(t)^2$	**0.95**	2400/1600	**RC**	**0.01**
***Z*(*t*)**	***X*(*t*)**Y*(*t*)**	**0.91**	2400/1600	**RC**	**0.01**