Physics Letters A 384 (2020) 126300

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Robust prediction of complex spatiotemporal states through machine learning with sparse sensing

G.D. Barmparis [a,b,c,*], G. Neofotistos [a,b,c,d], M. Mattheakis [d], J. Hizanidis [b], G.P. Tsironis [a,b,c,d], E. Kaxiras [d,e]

[a] Institute of Electronic Structure and Laser, Foundation for Research and Technology-Hellas, 71110 Heraklion, Crete, Greece
[b] Department of Physics, University of Crete, Heraklion, Greece
[c] National University of Science and Technology, MISiS, Leninsky Prospekt 4, Moscow, Russia
[d] School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
[e] Department of Physics, Harvard University, Cambridge, MA 02138, USA

ARTICLE INFO

Article history:
Received 21 December 2019
Received in revised form 25 January 2020
Accepted 27 January 2020
Available online 30 January 2020
Communicated by B. Malomed

Keywords:
Prediction
Machine learning
Sparse sensor placement
Chimera state

ABSTRACT

Complex spatiotemporal states arise frequently in material as well as biological systems consisting of multiple interacting units. A specific, but rather ubiquitous and interesting example is that of "chimeras", existing in the edge between order and chaos. We use Machine Learning methods involving "observers" to predict the evolution of a system of coupled lasers, comprising turbulent chimera states and of a less chaotic biological one, of modular neuronal networks containing states that are synchronized across the networks. We demonstrated the necessity of using "observers" to improve the performance of Feed-Forward Networks in such complex systems. The robustness of the forecasting capabilities of the "Observer Feed-Forward Networks" versus the distribution of the observers, including equidistant and random, and the motion of them, including stationary and moving was also investigated. We conclude that the method has broader applicability in dynamical system context when partial dynamical information about the system is available.

1. Introduction

The very purpose of natural science is to make reliable predictions about the future. These predictions, while possible in the Newtonian framework of several "simple" systems, become practically impossible in the presence of nonlinearity and larger number degrees of freedom. Chaotic systems with three degrees of freedom and above are notoriously impossible to predict in a long-time horizon. Similarly, spatiotemporal complex systems involve a large degree of unpredictability that in most cases must be handled in a statistical manner. Artificial Intelligence (AI), on the other hand, also aims at forecasting the future, although in a way different from that of natural science; it uses past experience and training so that an AI-system first "learns" and then predicts. Is it possible to merge the two approaches and produce a method that will increase the predictability of complex systems? Although there is no general answer, the very recent study of various important examples shows that this approach is very promising [1–11].

One could question the value of predicting a fully chaotic system, if this was possible. Let us give a familiar example: On the surface of the sea we have formation of smaller and larger waves that are in constant state of change. Depending of the prevailing conditions of wind, pressure, temperature, etc these waves can appear to be stochastic. In these cases, a statistical picture is more useful rather than the knowledge of the evolution of each little wave, its collisions and transformations. On the other hand, there are instances in spatiotemporal systems where partial order and stochasticity or chaos coexist. Predicting the evolution of complex structures in an otherwise stochastic system is both challenging and relevant. Chimera states provide such complex states where order and chaos coexists in spatiotemporal dynamics and, as a result, provide interesting and relevant modes to investigate from the AI point of view.

* Corresponding author at: Department of Physics, University of Crete and Institute of Electronic Structure and Laser, Foundation for Research and Technology-Hellas (FORTH), FORTH Campus, Heraklion, Crete 71110, Greece.
E-mail address: barmparis@physics.uoc.gr (G.D. Barmparis).

https://doi.org/10.1016/j.physleta.2020.126300

In the AI motivated work we need to select appropriate Machine Learning (ML) methods that in our case will be based solely on Artificial Neural Networks (ANN) [1,8]. Subsequently, we need to train the ANNs with "ground truth" data, i.e. simulated spatiotemporal data that describe precisely the phenomena we investigate and then we let the trained ANNs to predict the future evolution. In order to facilitate the ANNs in their performance we introduce additionally sparse sensors that we term "observers". These observers give partial system information to the AI units *even after their training* and thus "guide" them towards a better prediction. This observer-based approach was initially introduced in the context of the Reservoir computing method [3, 12] but it has been shown to be quite efficient more generally [13].

While reconstructing and predicting complex nonlinear dynamical systems using partial information through a small number of sparse sensors is a very challenging problem of great importance, a critical aspect of this problem is the optimum placement of the sparse observers, in order to achieve efficient prediction when the cost of acquiring and processing data is high [14–17]. The main motivation thus of the present work is to investigate how observers "interact" with the AI predictive units. More specifically we look into the effect that the various distributions of static observes as well as observers in a dynamical state have in the predictability of the complex spatiotemporal systems we analyze.

Leading sparse sensing methods typically exploit either spatial or temporal correlations, but rarely both [14]. Our work is motivated on the exploitation of both spatial and temporal correlations by means of AI-ML methods and aims in addition to finding optimum sensor placement and dynamics for the prediction of complex spatiotemporal behavior. We target primarily a distinct prototypical phenomenon, viz. partially coherent chimera states. Chimeras are archetypical complex phenomena appearing in many physical systems; their intrinsic dynamics delineate nontrivial cases in the complexity of wave motion and present severe challenges in predicting chaotic behavior in extended physical systems. What is especially interesting and, in a sense, "paradigmatic" about chimera states is that they embody simultaneously both faces of nonlinearity, viz. order and chaos. Predicting order is unnecessary while predicting chaos is futile; the possibility of predicting a "mixed" complex state such as chimeras is thus very challenging since it demarks two completely different predictability regimes.

In the present work, we focus particularly on the robustness of the long-term forecasting capability of fully connected Feed-Forward Neural Networks (FNN) with Observers (OFNNs), as applied to the spatiotemporal evolution of two distinct types of systems that support quite different complex spatiotemporal states. The first is an optical system and involves turbulent chimeras in simulated arrays of coupled semiconductor lasers [13]. The second is biological relates to complex dynamics in modular neuronal networks based on the simulated connectome of the *C.elegans* soil worm [18]. This connectome is organized into six interconnected communities, with neurons obeying chaotic bursting dynamics. Both systems are of practical importance and they have been studied extensively. The OFNNs method, we employ, assigns one network to each one of the system's nodes except for the "observer" nodes which provide continual "ground truth" measurements as input. The reason for using OFNNs is that they train much faster than similar (e.g. OLSTM) methods exhibiting better forecasting performance [13]. Here, we use the same implementation of the OFNNs method as in [13] and we provide further evidence of the necessity of the "observers" in order the FNNs models to achieve satisfactory forecasting capabilities. Then, in order to investigate the robustness of the prediction capabilities of the OFFNs method in respect of different observers placements we use the following observers' schemes: **a)** uniformly distributed (equidistant) observers at all time steps; **b)** randomly assigned observers at each time step; **c)** uniformly distributed (equidistant) observers moving over time in phase (performing oscillatory periodic motion around their positions, with amplitude equal to 5 nodes distance), both in training and testing; **d)** randomly assigned observers, (at fixed positions for all time steps); **e)** randomly assigned observers at fixed positions for all time steps, moving (in phase) around their fixed positions in an oscillatory periodic motion (amplitude equal to 5 node distance); **f)** uniformly distributed (equidistant) observers, stationary during training, but moving (in phase) during testing (oscillatory periodic motion around their positions, with amplitude equal to 5 node distance); and **g)** uniformly distributed (equidistant) observers during training, but moving (with random phase) during testing. Although there are many other observed distributions, the static and moving observers we chose to investigate practically exhaust the interesting modes of interaction of observes with the AI-ML system. From these modes we arrive at quantitative conclusions on observer-assisted AI predictability of spatiotemporal systems.

The outline of this article is the following. In the next Section, we demonstrate the necessity of using "observers" in order to achieve satisfactory predictability in spatiotemporal dynamical systems. Section 3 and 4, provide the central results of this work for coupled semiconductor lasers and biological organized states respectively; in these sections we employ several sparse sensing schemes with different spatial and kinematic distributions and demonstrate the applicability of the OFFNs method for chimera predictions. In Section 5, we provide a quantitative comparison of the various schemes that gives through the prediction error the predictability horizon. In Section 6, we conclude and in Appendix A, we show the full evolution of the observers' positions for each one of the schemes investigated.

2. The necessity of using observers in feed-forward neural networks

We study the forecasting capability of fully connected FNNs in the spatiotemporal evolution of multi-clustered turbulent chimera states in coupled lasers system, characterized by self-organized patterns of coexisting coherence and incoherence, as a function of the number of observers in the system.

Our initial approach comprises a simple model with a single FNN assigned to each system node, which is independent of all other FNNs in the system nodes. However, this simple model found to produce very large prediction error (see Fig. 1, blue line). This is because of the fact that chimera states are collective phenomena (i.e. the nodes of the system are not isolated but correlated to each other) and thus one FNN model per node cannot learn the correlations between the nodes since it has no information from the surrounding environment.

This limitation of the FNNs models can be overcome by using OFNNs models [13] that can capture the correlations between the nodes. The advantage of using models with observers, providing real time data to the model, is that in this case the model is informed about the surrounding environment and is capable, using this information, to react in real time in any change of the system. Fig. 1, presents the performance of the OFNNs model (RMSE) as a function of the number of the observers, for the coupled lasers system chimeras. The blue line (in the top of the figure) depicts the RMSE (calculated over all the system nodes at each predicted time-step) for an FNN with no observers. As can be seen,

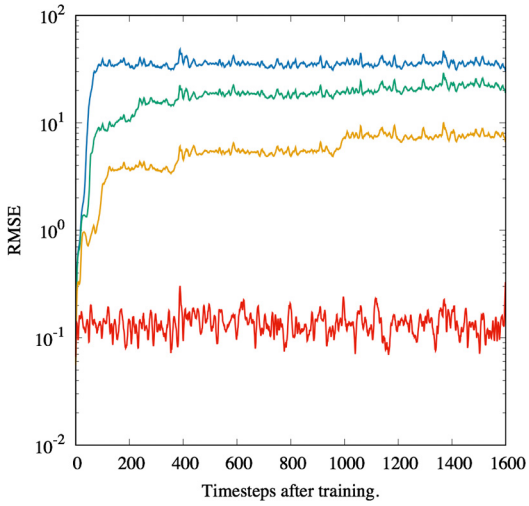

Fig. 1. RMSE of OFNNs with no observers (blue line), 1 observer (green line), 3 observers (orange line), and 14 observers (red line), calculated over all the system nodes at each predicted time-step in a system of an array of 200 coupled lasers (turbulent chimeras). (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

the prediction error for this model is very large, the largest of all cases studied. The error reduces by adding observers in the FNNs (making them OFNNs) as can be seen in the OFNNs with 1 observer (green line), 3 observers (orange line), and 14 observers (red line). The root mean square error (RMSE) is calculated as in [3], Eq. (21).

3. Predicting turbulent chimeras in coupled lasers array using various observers' schemes

The dynamics of the coupled lasers system have been initially studied from the point of view of ML-based predictability in [13]. Here in Fig. 2, we present the spatiotemporal plot of this turbulent chimera state. Subplot a) presents the simulated results (as generated in the 1-dimensional semiconductor Class B laser array in [13], see Fig. A.1 (a), studied numerically for a non-local coupling scheme). Subplot b) presents results for the same placement and motion cases as in Table 1. It should be noted that the thick vertical black line represents the size of the training dataset used for training the OFNNs model (comprising 400 time steps; 20% of the total number of time steps).

4. Predicting synchronized states in modular neuronal networks using various observers' schemes

Hizanidis et al. [18] considered a neuronal network inspired by the connectome of the *C. elegans* soil worm, organized into six interconnected communities. The neurons a) obey chaotic bursting dynamics, b) are connected with electrical synapses within their communities and with chemical synapses across them, and c) the neuron dynamics are modeled in terms of the Hindmarsh-Rose system. As their numerical simulations reveal, the coaction of these two types of coupling can shape the dynamics in such a way that various organized or partially organized states can happen. One type consists of a fraction of synchronized neurons, which belong to the larger communities, and a fraction of desynchronized neurons that are part of smaller communities. Since this state is structurally similar to the one addressed in the context of the optical system previously, we focus on a different state where there is intermittent synchronization. In addition to the Kuramoto order parameter, they also employ other measures of coherence, such as the chimera-like χ and metastability λ indices, which quantify the degree of synchronization among communities and along time, respectively. The Kuramoto order parameter, which measures synchronization, is bounded in the interval $[0, 1]$ and is equal to 1 when neurons in the considered population are completely synchronized and 0 when they are totally desynchronized. The relevant equations are presented in detail in [18].

In Fig. 3, subplot a) presents the spatiotemporal plot of the modular neuronal networks' synchronized state (see also Fig. A.2).

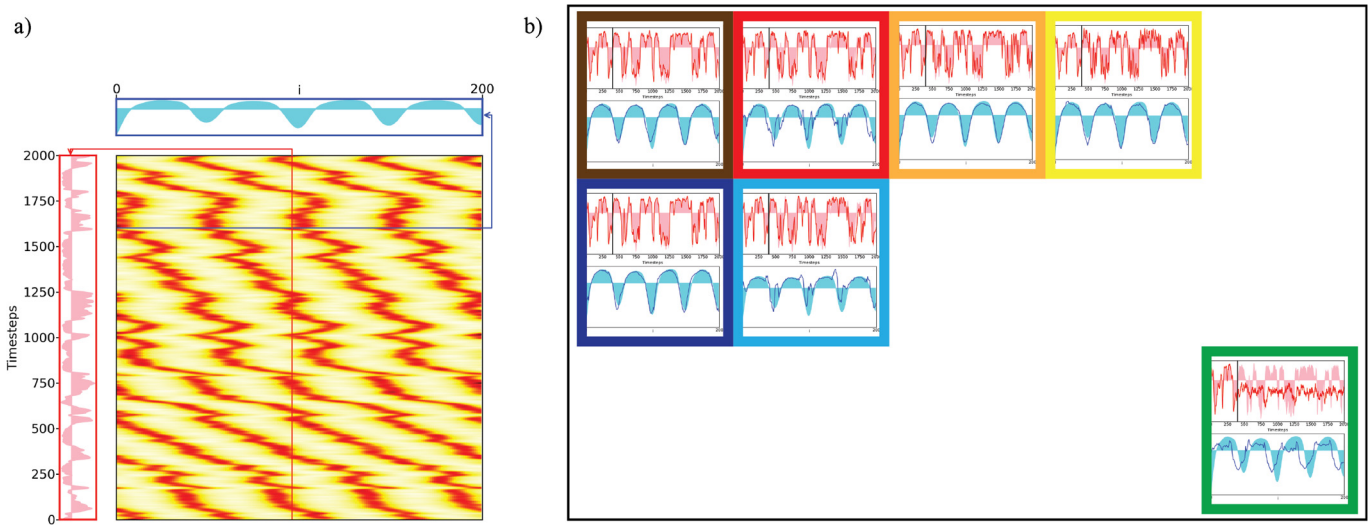

Fig. 2. a) Spatiotemporal plot of the turbulent chimera state as studied numerically in [13], data also in Fig. A.1a. Right, vertical inset: filled pink area presents the actual (ground-truth) time series for laser #96. Top inset: filled cyan area presents the snapshot of the spatial profile of the ground truth local curvatures for the entire array at time step $t_n = 1{,}600$ (four times the training time). **b)** *Top*: presents the predicted time series for laser #96: red line depicts the predicted data by OFNNs and the filled pink area the actual (ground-truth) data. The thick vertical black line represents the size of the training dataset used for training the OFNNs model (comprising 400 time steps; 20% of the total number of time steps). *Bottom*: snapshot of the spatial profile of the predicted local curvatures for the entire array at time step $t_n = 1{,}600$ (four times the training time): blue line depicts the predicted data by OFNNs and the filled cyan area the actual (ground-truth) data. Colored boxes code as in Table 1.

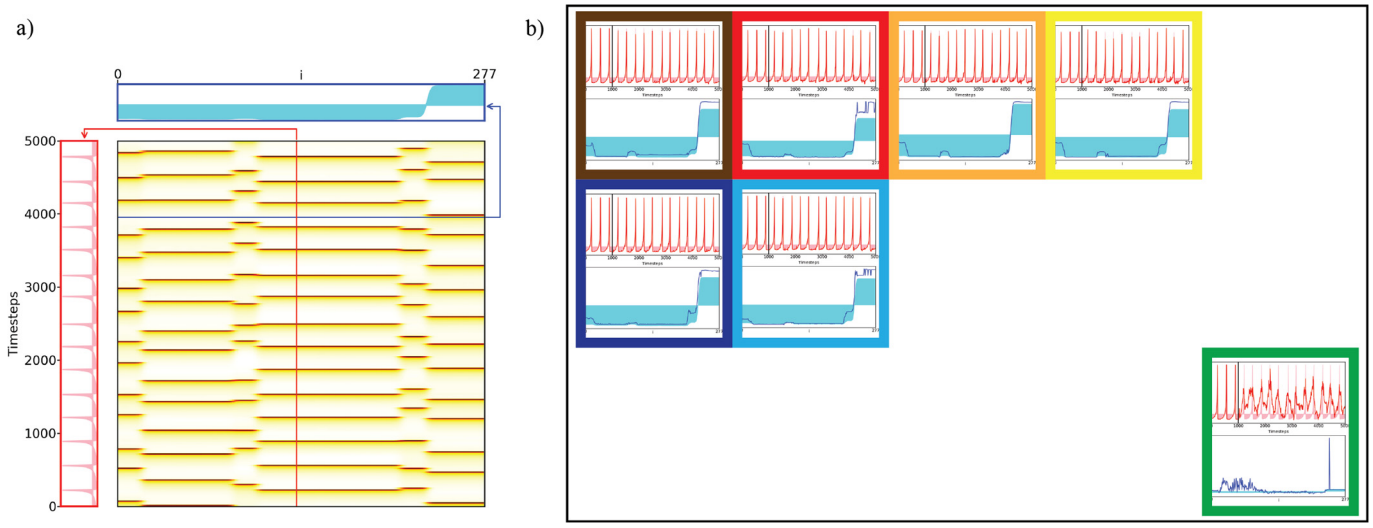

Fig. 3. a) Spatiotemporal plots of the modular partially synchronized state generated in Hizanidis et al. [18] Figure 3a. Right, vertical inset: filled pink area presents the actual (ground-truth) time series for node #135. Top inset: filled cyan area presents the snapshot of the spatial profile of the ground truth local curvatures for the entire array at time step $t_n = 4{,}000$ (four times the training time). **b)** *Top:* presents the predicted time series for node #135: red line depicts the predicted data by OFNNs and the filled pink area the actual (ground-truth) data. The thick vertical black line represents the size of the training dataset used for training the OFNNs model (comprising 1000 time steps; 20% of the total number of time steps). *Bottom:* snapshot of the spatial profile of the predicted local curvatures for the entire array at time step $t_n = 4{,}000$ (four times the training time): blue line depicts the predicted data by OFNNs and the filled cyan area the actual (ground-truth) data. Colored boxes code as in Table 1.

Table 1
On the left: Structure of each inner cell of this Table. On the right: Prediction errors (RMSEs) of the OFNNs methods calculated over all the predicted time-steps and over all the system nodes. **Upper left corner of each inner cell** (blue), turbulent chimera of the coupled lasers system; with 14 observers in a total of 200 nodes (7% of total). **Lower right corner of each inner cell** (black), synchronized states in modular neuronal networks; with 20 observers in a total of 277 nodes (∼7% of total). Colored boxes are used to easily distinguish the different observers' distribution.

Turbulent chimeras / Modular Neural Networks

Observer Placement \ Observer Motion	Stationary over time (training and testing)	Moving (in phase) over time (training and testing)	Stationary during training. Moving (in phase) during testing	Stationary during training. Moving (random phase) during testing	Random assignment at each time-step
Uniform (equidistant)	0.225 / 0.216	0.697 / 0.229	0.491 / 0.239	0.393 / 0.248	
Random assignment (fixed position)	0.238 / 0.335	0.706 / 0.327			
Random assignment at each time-step					1.255 / 1.980

Subplot b) presents results for the same placement and motion cases as in Table 1. The thick vertical black line represents the size of the training dataset used for training the OFNN model (comprising 400 time steps; 20% of the total number of time steps).

We emphasize that in both cases (of Fig. 2 and Fig. 3) these are very long-time predictions, not just short-term prediction of a few time steps beyond the training time.

5. Comparing prediction performance

Table 1, presents prediction errors for both the optical and the biological problem and for each observer-placement scheme, calculated by means of the RMSE over all nodes for all predicted time steps. The cell on the left of Table 1, presents the structure that each inner cell of Table 1 has. On the upper left corner of each inner cell of Table 1, we present the corresponding RMSE value for the turbulent chimeras system and on the lower right corner the corresponding RMSE value for the modular neural network system. As can be seen, the OFNNs method is robust for different observer placement schemes and different motions of the observers and fails only in the case of completely randomly moving observers. The stationary uniformly distributed observers demonstrate the best prediction performance with stationary and uniformly placed observers for both types of optical and biological states, followed by randomly placed observers (with training employing stationary observers) for the coupled lasers array chimera, and moving observers for the modular neuronal networks' synchronized state.

This effect can be explained as follows: when the underlying physical system involves spatiotemporal dynamics (exhibiting tur-

bulent chimeras), small movements (motions) of the observers can affect the predictability of the model since the observers may get values from a completely different environment (i.e. before the movement the observer can take ground-truth measurements from a coherent area and after the movement may take measurements from an incoherent area). This will affect the predictability of the model since the observer provides a very different value than the ones the nodes have been "trained" with. Stationary observers are the optimum ones in this case.

In contrast, when the system is not turbulent, small changes in the position of the observers will not affect much the ground-truth measurements taken by of the observer, and thus will not affect the predicted values of the model. Moving or stationary observers can both be good choices of observer motions.

6. Conclusions

Here, we address the issue of predicting complex spatiotemporal systems using ML techniques involving "observers" in various placement schemes. The two systems considered in this work are, a system of coupled lasers, comprising turbulent chimera states and a biological one, of modular neuronal networks containing states that are synchronized across the networks, being less chaotic than chimera states. We demonstrated the necessity of using "observers" in order to improve the performance of FNNs models in such complex systems. The robustness of the forecasting capabilities of the OFNNs models versus the distribution of the observers, including equidistant and random, and the motion of them, including stationary and moving was also investigated. Especially, we found that in case of the turbulent chimeras the predictability of the model does not show strong dependence on the spatial distribution of the observers; both equidistant and randomly distributed observers are able to predict the behavior of the system. When the observers are moving in phase, the prediction is less accurate but still satisfactory. In cases where the model is trained with stationary uniformly distributed observers but in the prediction stage the observers are moving around their initial position (oscillatory motion), the OFNNs is still able to predict the evolution of the system. However, the model fails in the case of random observers that are moving randomly at each step. In the case of the modular biological spatiotemporal state, the predictability of the model does not show strong dependence in either the spatial distribution of the observers or in their motion. As before, the model fails in the case of random observers that execute a random walk. We believe that the method has broader applicability in dynamical system context when partial dynamical information about the system is available.

Methods

In both systems investigated in this work, the data were preprocessed so that they have zero mean and unit variance, and smoothed by means of Gaussian Smoothing Kernel with standard deviation $\sigma = 2.5$. For each one of the time series of each system one OFNN model is trained. Each model contains, an input layer of $N + 1$ nodes (taking values at time t), where N is the number of the Observers, plus one value from the given time series, a hidden layer of 400 nodes with ReLU activation function and an output layer with one node giving output at time $t + 1$. A validation set is used to determine the performance of the model. After each epoch the RMSE in the validation dataset is computed. Each model is trained for 200 epochs and the one with the smallest validation error is selected to avoid over-fitting. Each selected OFNN network is then used to forecast the node's state in the next time step in an iterative fashion. All the hyper-parameters of each OFNN model including the number of training epochs, the number of training batches, the number of neurons and the number of hidden layers are optimized so that the RMSE to converge. The *Adam* stochastic optimization method [19] with a learning rate of 0.001 as implemented in *Keras* [20] was used to optimize the OFNNs during training.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

G.D.B acknowledges support by the "HELLAS-CH" (MIS Grant No. 5002735) implemented under the "Action for Strengthening Research and Innovation Infrastructures," funded by the Operational Program "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund). G.N. and G.P.T. acknowledge support by the European Commission under project NHQWAVE (MSCA-RISE 691209). J.H. acknowledges support by the General Secretariat for Research and Technology (GSRT) and the Hellenic Foundation for Research and Innovation (HFRI) (Code: 203). This work was also partially supported by the Ministry of Education of the Russian Federation in the framework of the Increased Competitiveness Program of NUST "MISiS" (No. K2-2019-010), implemented by a governmental decree dated 16th March 2013, N211. G.D.B. and G.N. gratefully acknowledge the hospitality of the Laboratory for Superconducting Metamaterials, NUST "MISiS" where part of this work was performed.

Appendix A. Pictorial representation of observer dynamics and associated spatiotemporal predictions

In the Appendix we show the full evolution and predictions of the case of Fig. 2 and Fig. 3 respectively. In the specific representation the full movement of observes is visible as well as the predicted states.

The thick horizontal black line represents the time horizon of the dataset used for training the OFNNs model (comprising 400 time steps; 20% of the total number of time steps). Black dots represent the position of the observers at each time step. Right, vertical inset: The predicted time series for laser #96: cyan line depicts the predicted date by OFNNs and the dashed orange line the actual (ground-truth) data.

Bottom inset: Snapshot of the spatial profile of the predicted local curvatures for the entire array at time step $t_n = 1{,}600$ (four times the training time): color code is as in the right inset.

The thick horizontal black line represents the time horizon of the dataset used for training the OFNNs model (20% of the total number of time steps; 1,000 time steps). Black dots represent the position of the observers at each time step. Right, vertical inset: The predicted time series for laser #135: cyan line depicts the predicted date by OFNNs and the dashed orange line the actual (ground-truth) data. Bottom inset: Snapshot of the spatial profile of the predicted local curvatures for the entire array at time step $t_n = 4{,}000$ (four times the training time): color code is as in the right inset.

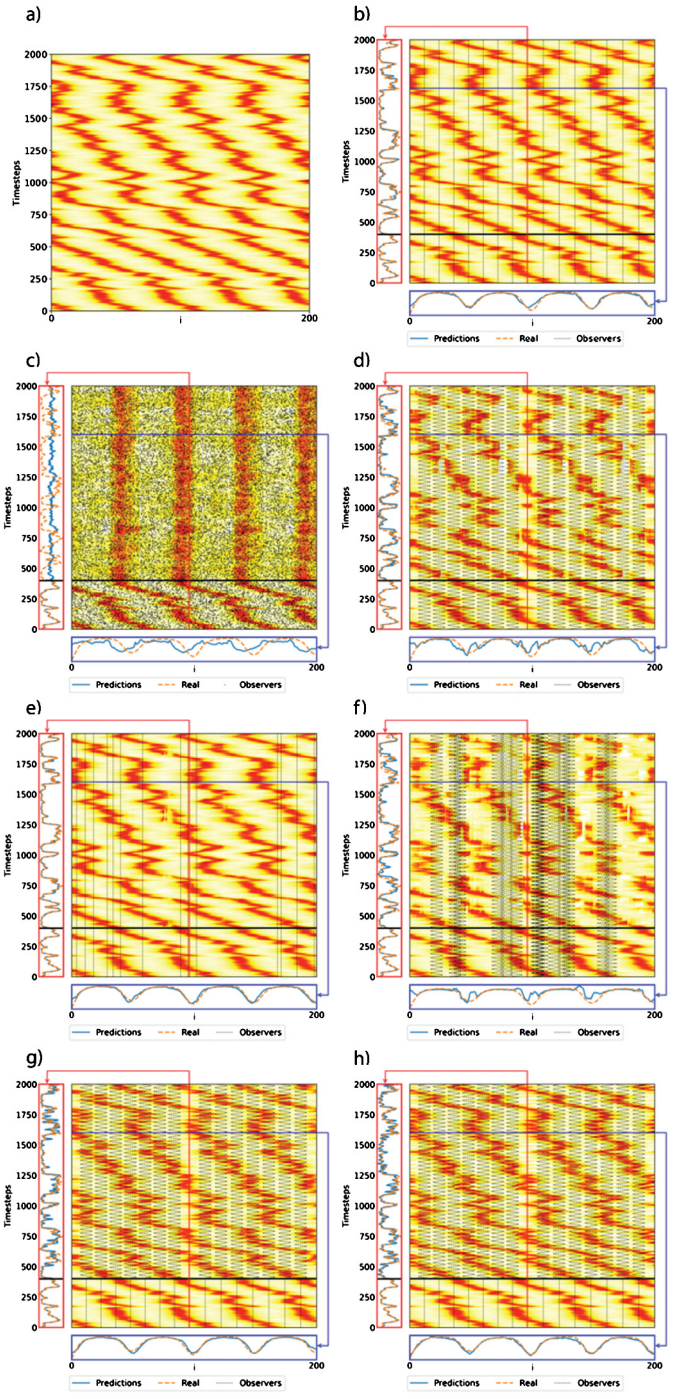

Fig. A.1. Spatiotemporal plots of the turbulent chimera state generated in [13] Figure 4V. **a)** the evolution of the chimera state as studied numerically in [13] for a non-local coupling scheme; **b)** results obtained by incorporating uniformly distributed (equidistant) observers at all time steps; **c)** results obtained by incorporating randomly assigned observers at each time step; **d)** results obtained by incorporating uniformly distributed (equidistant) observers moving over time in phase (performing oscillatory periodic motion around their positions, with amplitude equal to 5 nodes distance), both in training and testing; **e)** results obtained by incorporating randomly assigned observers, (at fixed positions for all time steps); **f)** results obtained by incorporating randomly assigned observers at fixed positions for all time steps, moving (in phase) around their fixed positions in an oscillatory periodic motion (amplitude equal to 5 node distance); **g)** results obtained by incorporating uniformly distributed (equidistant) observers, stationary during training, but moving (in phase) during testing (oscillatory periodic motion around their positions, with amplitude equal to 5 node distance); **h)** results obtained by incorporating uniformly distributed (equidistant) observers during training, but moving (with random phase) during testing.

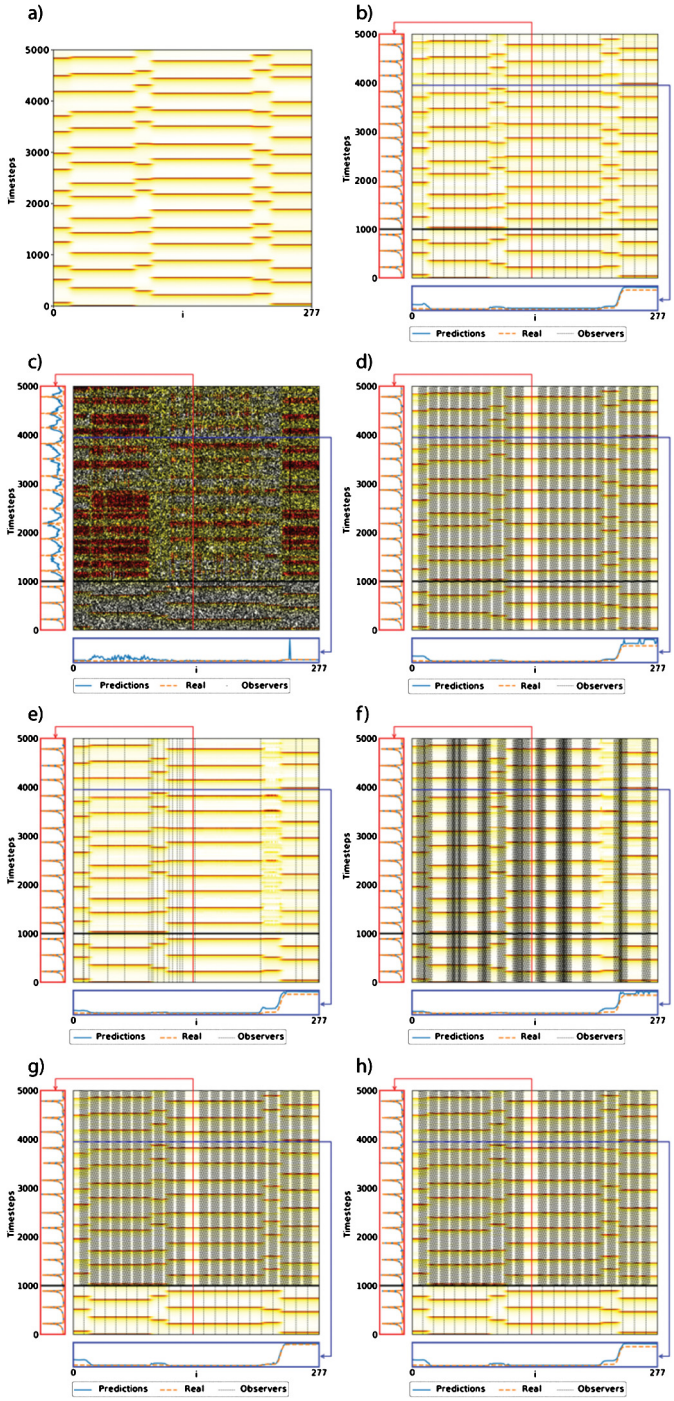

Fig. A.2. Spatiotemporal plots of the modular spatiotemporally synchronized states generated in Hizanidis et al. [18] Figure 3a. **a)** the evolution of the state as studied numerically in [18]; **b)** results obtained by incorporating uniformly distributed (equidistant) observers at all time steps; **c)** results obtained by incorporating randomly assigned observers at each time step; **d)** results obtained by incorporating uniformly distributed (equidistant) observers moving over time in phase (performing oscillatory periodic motion around their positions, with amplitude equal to 5 nodes distance), both in training and testing; **e)** results obtained by incorporating randomly assigned observers, (at fixed positions for all time steps); **f)** results obtained by incorporating randomly assigned observers at fixed positions for all time steps, moving (in phase) around their fixed positions in an oscillatory periodic motion (amplitude equal to 5 node distance); **g)** results obtained by incorporating uniformly distributed (equidistant) observers, stationary during training, but moving (in phase) during testing (oscillatory periodic motion around their positions, with amplitude equal to 5 node distance); **h)** results obtained by incorporating uniformly distributed (equidistant) observers during training, but moving (with random phase) during testing.

References

[1] G. Hinton, L. Deng, D. Yu, G.E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath, B. Kingsbury, Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups, IEEE Signal Process. Mag. 29 (2012) 82–97.

[2] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. Van Den Driessche, T. Graepel, D. Hassabis, Mastering the game of Go without human knowledge, Nature 550 (2017) 354.

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

[4] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780.

[5] A. Graves, A.R. Mohamed, G. Hinton, Speech recognition with deep recurrent neural networks, in: Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference, 2013, pp. 6645–6649.

[6] Y. Wu, M. Schuster, Z. Chen, Q.V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, Google's neural machine translation system: bridging the gap between human and machine translation, arXiv:1609.08144, preprint posted October 8, 2016.

[7] H. Maathuis, L. Boulogne, M. Wiering, A. Sterk, Predicting chaotic time series using machine learning techniques, in: B. Verheij, M. Wiering (Eds.), Preproceedings of the 29th Benelux Conference on Artificial Intelligence, BNAIC 2017, Groningen, Netherlands, 2017, pp. 326–340.

[8] P.R. Vlachas, W. Byeon, Z.Y. Wan, T.P. Sapsis, P. Koumoutsakos, Data-driven forecasting of high-dimensional chaotic systems with long short-term memory networks, Proc. R. Soc. A 474 (2018) 20170844.

[9] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, S. Savarese, Learning to predict human behaviour in crowded spaces, in: V. Murino, M. Cristani, S. Shah, S. Savarese (Eds.), Group and Crowd Behavior for Computer Vision, 1st edition, Elsevier, Academic Press, Cambridge, MA, 2017, pp. 183–204.

[10] J. Pathak, B. Hunt, M. Girvan, Z. Lu, E. Ott, Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach, Phys. Rev. Lett. 120 (2017) 024102.

[11] J. Pathak, Z. Lu, B. Hunt, M. Girvan, E. Ott, Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data, Chaos 27 (2017) 121102.

[12] H. Jaeger, H. Haas, Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication, Science 304 (5667) (2004) 78–80, https://doi.org/10.1126/science.1091277.

[13] G. Neofotistos, M. Mattheakis, G.D. Barmparis, J. Hizanidis, G.P. Tsironis, E. Kaxiras, Machine learning with observers predicts complex spatiotemporal behavior, Front. Phys. 7 (2019) 24, https://doi.org/10.3389/fphy.2019.00024.

[14] T.L. Mohren, T.L. Daniel, S.L. Brunton, B.W. Brunton, Neural-inspired sensors enable sparse, efficient classification of spatiotemporal data, Proc. Natl. Acad. Sci. 115 (42) (2018) 10564–10569, https://doi.org/10.1073/pnas.1808909115.

[15] K. Manohar, B.W. Brunton, J.N. Kutz, S.L. Brunton, Data-driven sparse sensor placement, IEEE Control Syst. Mag. 38 (2018) 63–86.

[16] B.W. Brunton, S.L. Brunton, J.L. Proctor, J.N. Kutz, Sparse sensor placement optimization for classification, SIAM J. Appl. Math. 76 (2016) 2099–2122.

[17] J. Annoni, T. Taylor, C. Bay, K. Johnson, L. Pao, P. Fleming, K. Dykes, Sparse-sensor placement for wind farm control, J. Phys. Conf. Ser. 1037 (2018) 032019.

[18] J. Hizanidis, N.E. Kouvaris, G. Zamora-López, A. Díaz-Guilera, C. Antonopoulos, Chimera-like states in modular neural networks, Sci. Rep. 6 (2016) 19845.

[19] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv:1412.6980v9, preprint posted January 30, 2017.

[20] F. Chollet, Keras, GitHub, https://github.com/fchollet/keras, 2015.