On the Imputation of Power System Measurement Streams with Imperfect Wireless Communication

Theodoros A. Alexopoulos*, Charalampos Kalalas†, and George N. Korres*
*National Technical University of Athens, School of Electrical and Computer Engineering, Athens, Greece
†Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), Barcelona, Spain
Emails: talexo@mail.ntua.gr, ckalalas@cttc.es, gkorres@cs.ntua.gr

***Abstract*—Dependable measurement data are essential for the accuracy and integrity of power system state estimation and actuation. However, distribution grids empowered by wireless connectivity are subject to missing sensor observations due to channel stochasticity. In this paper, we leverage the potential of dynamical models to extract knowledge from the ambient measurement space under unfavorable and ideal network conditions. A rigorous assessment of various network configurations based on empirical evaluations reveals meaningful performance trends for model fitting and recovery of incomplete trajectories. In particular, while the imperfect connectivity may curtail the ability of dynamical models to exploit the intrinsic spatio-temporal structure, measurement imputation under ideal network configuration exhibits favorable performance. Our research outcomes highlight the impact of the underlying wireless connectivity on measurement data acquisition and handling in distribution grids.**

***Index Terms*—data imputation, distribution system, dynamical system, outage probability, wireless channel**

I. Introduction

In recent years, key investments in power systems have remarkably improved the overall quality of their associated information and communication infrastructure [1]. The availability of dependable measurement data streams, provided by phasor measurement units (PMUs) [2], in conjunction with expanded computational resources and the proliferation of effective techniques for knowledge extraction drive a paradigm shift in power system analysis and provide a fertile ground for enhanced monitoring and control capabilities, even in the presence of elusive governing dynamics due to stochasticity. Nevertheless, the utilization of emerging wireless communication (i.e., 5G and beyond) networks, mainly in distribution grids, comes inadvertently with challenges which involve distortions and missing data due to the unreliable nature of the shared wireless medium. Addressing the degradation of received data is essential for high fidelity estimation of state variables and informed system control.

The notion that power system data streams consist of spatially co-evolving time-series [3] elucidates the hypothesis that the exploitation of spatio-temporal cross-correlations among measurements holds the promise of extracting the underlying dynamics which govern their behavior in an effort to curtail their loss of quality attributed to communication channel imperfections. Such data are characterized by the key conceptual traits of high-order, physical context, and temporal smoothness. High-order stems from the fact that sensor installations exist at disparate spatial locations, reporting local electrical quantities. Moreover, the underlying power grid which comprises the physical process giving rise to the collected measurements induces contextual information, while temporal smoothness is associated with the inherent correlation among adjacent observations of the same measurement stream.

The problem of missing data recovery can be addressed at its foundational level by the following three schemes. The most intuitive case of data imputation is conventional linear, cubic, or nearest neighbor interpolation. Such an approach, albeit exploiting temporal smoothness pertaining to a single data stream and resonating computational efficiency, fails to incorporate information leveraged from correlations among different measurement sources [4]. This cogently justifies the utilization of approaches based on singular value decomposition (SVD). Their prototypical hallmark constitutes the ability to deduce linear correlations among measurement trajectories and, thus, reconstruct missing values in one data stream from the rest. Notably, missing values can be inferred by applying an iterative scheme based on consecutive low-rank decompositions [5], [6]. It is to be noted that the existence of invariances among low-rank embeddings in the data and the potential presence of transient content pose perils which may reduce the effectiveness of SVD-based schemes for occluded data recovery [7]. Dynamical systems, capable of exploiting spatio-temporal correlations among data streams, can be utilized in an iterative strategy for estimating missing values [4].

In the framework of power system measurements, several works have addressed the problem of missing data estimation. In [8], occluded data among PMU streams are imputed through the utilization of a low-rank matrix completion scheme, with the ability to be deployed as a real-time application. Simultaneous reconstruction of missing values and correction of bad data is proposed in [9] based on employing the Hankel matrix of PMU measurements and its low-rank feature. By modeling measurement data as samples from a Gaussian random process with known covariance matrix, Bayesian singular value thresholding is introduced in [10] as an efficient, iterative approach to data recovery in low voltage distribution systems, exhibiting the ability to estimate the optimal singular value threshold in each iteration. A real-time missing data estimation strategy based on matrix completion with penalization of singular values by means of Schatten-q quasi-norm is outlined in [11].

An online scheme for parameter learning of hidden Markov models with contextual information has been employed in [3] with the purpose of imputing missing PMU data.

There are two main desiderata in the framework of this paper. First, in an effort to regain perspective on the mechanisms which causally induce occlusions, we investigate the lingering influence of communication impairments to the emergence of missing data among measurement streams at the level of a fusion center. In particular, we thoroughly explore the impact of various network configurations on the quality of the received measurement data with the aid of a tractable expression for the outage probability. Second, we examine the ability of interpretable dynamical systems *i*) to fit the observed data at the fusion center in the presence of less favorable network connectivity conditions; and *ii*) to impute missing values under ideal decoding of received measurements. It should be underlined that this work is not inclined towards the identification and correction of bad data among decoded measurement values; it is rather a first step towards the design of appropriate connectivity enablers which allow for decoded values with meaningful spatio-temporal correlations.

Material in this manuscript is organized as follows. The employed network model and the respective probability of sensor connectivity outage are described in Section II. The utilized scheme for data imputation is outlined in Section III. Results pertaining to the performance of model fitting and quality of missing data reconstruction based on numerical simulations under imperfect and ideal decoding conditions, respectively, are presented in Section IV. Section V is reserved for conclusion and discussion of the path forward.

II. Network Model

We consider a power grid monitoring scenario composed of a fusion center and a number of deployed monitoring sensors. Each spatially allocated sensor captures a multitude of electrical quantities depending on its functionality. For instance, a PMU is typically capable of reporting voltage and current magnitudes and phases, frequency, and rate of change of frequency, while a remote terminal unit (RTU) associated with conventional supervisory control and data acquisition (SCADA) infrastructure may provide measurements of voltage and current magnitudes, as well as power flows and injections.

A cellular network topology is considered where the sensors, equipped with radio interfaces, transmit their measurements to the fusion center which is co-located with their associated base station (BS). A closest BS association scheme is adopted in this paper. We employ a stochastic geometry framework for the positions of the BSs and the sensors to facilitate the modeling of the network topology and the tractable analysis of the uplink sensor data transmission [12]. In particular, we assume that the BSs are distributed on the plane according to a homogeneous Poisson point process (PPP) Ψ_b of intensity λ_b, while the sensors are spatially distributed according to a homogeneous PPP $\Psi_u = \{u_k; i = 1, 2, 3, \dots\}$ with intensity λ_u. We further assume saturated uplink traffic conditions where all sensors always have data to transmit while λ_u is considered high enough such that each BS will serve at least one sensor per channel.

A. Channel Model and Power Control

A generic power-law path-loss model is considered where the signal power decays with the propagation distance r at the rate $r^{-\alpha}$, where α denotes the path-loss exponent. In addition to standard path-loss attenuation, we assume a single-parameter Rayleigh fast fading model for a tractable representation of the random channel effects[1]. In particular, the channel power gains, denoted by g, are assumed to be independent of each other and of the spatial locations, and identically distributed (i.i.d), following an exponential distribution with unit mean. Let p_k denote the transmit power of sensor k. All sensors are considered to have equal maximum transmit power denoted by $p_{\max}$ and employ full channel-inversion power control to compensate for the path loss. That is, each sensor controls its p_k such that the average received signal power at the serving BS is equal to a threshold ρ. The threshold ρ constitutes a network design parameter and takes values greater than the receiver sensitivity $\rho_{\min}$.

B. Outage Probability

An uplink sensor transmission is considered to be successfully decoded at the serving BS if and only if the signal-to-interference-plus-noise ratio (SINR) of the useful signal is greater than a certain threshold γ_{th}. Based on the stationarity property of the PPP [13], the SINR experienced at the origin BS for a target sensor can be expressed as

$$\text{SINR} = \frac{\rho g_0}{\sigma^2 + \sum_{n=1}^{N_s} \rho g + \sum_{u_k \in \tilde{\Psi}_u} p_k g_k \left\| u_k \right\|^{-\alpha}}, \qquad (1)$$

where the numerator corresponds to the useful signal power and σ^2 is the noise power. The second term in the denominator of Eq. (1) denotes the aggregate intra-cell interference conditioned on the number of neighboring sensors N_s while the third term pertains to the inter-cell interference from other uplink sensor transmissions on the same time-frequency resource[2]. Based on the SINR analysis conducted in [14] and for $\alpha = 4$, the SINR outage probability, $P_o = \mathbb{P}(\text{SINR} < \gamma_{\text{th}})$, for a sensor transmission can be calculated as

$$\text{P}_o \approx 1 - \left(\frac{\mu(1+\gamma_{\text{th}})}{\mu+\gamma_{\text{th}}} \right)^c \exp\left\{ -\frac{\sigma^2 \gamma_{\text{th}}}{\rho} - \frac{\tilde{\lambda}_u \sqrt{\gamma_{\text{th}}} \arctan \sqrt{\gamma_{\text{th}}}}{\lambda_b} \right\}, \quad (2)$$

where $\mu = \frac{c\lambda_b}{c\lambda_b + \tilde{\lambda}_u}$ and $c = 3.575$ is a constant related to the Voronoi tessellation in the $\mathbb{R}^2$. The impact of γ_{th} and λ_b on the SINR outage probability is illustrated in Fig. 1. It can be observed that a high SINR threshold may severely degrade the achieved performance leading to a high number of missing sensor measurements. In addition, a lower BS intensity further exacerbates this detrimental effect.

[1] We note that more general fading/shadowing distributions can be considered at the expense of tractability loss and without significant deviation on the system design insights [12].

[2] For analytical tractability, the locations of the interfering sensors are considered to follow a PPP with intensity $\tilde{\lambda}_u$ and their transmit powers p_k are independent.

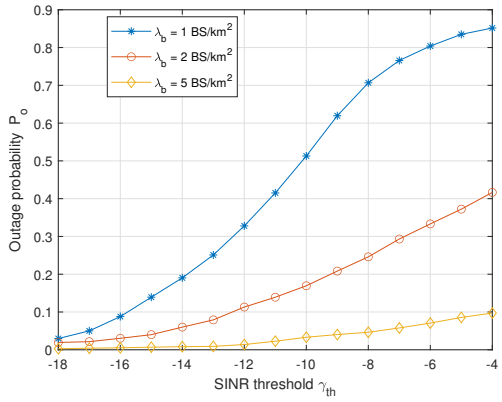

Fig. 1. Sensor outage probability for various SINR threshold values and BS intensities.

Due to the imperfect communication, the received sensor measurements at the fusion center, co-located with the serving BS, can be represented by a partially observable time sequence $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_T]$, where each vector $\mathbf{z}_t$ contains the received measurements at time-step t from the deployed sensors[3]. By incorporating the outage probability P_o (given by Eq. (2)), the entries $z_{t,k}$ of $\mathbf{z}_t$ can be expressed as

$$z_{t,k} = (1 - P_o)x_{t,k} + \upsilon_{t,k}, \tag{3}$$

where $x_{t,k}$ denotes the transmitted measurement from sensor k at time-step t and $\upsilon_{t,k}$ is a zero-mean white Gaussian noise with covariance matrix Ω. It is noted that Eq. (3) takes into account the impact of the wireless channel lossy nature on the sensor transmissions which may result in a received vector $\mathbf{z}_t$ with intermittent measurements. In order to impute the missing sensor observations, the conditional expectation of the missing measurements with respect to the observed ones over the time-frame from $t = 1$ to T needs to be estimated.

III. Model-based Handling of Measurement Data Streams

A. Dynamical Systems

The utilization of dynamical systems as an interpretable scheme for gaining insight into the behavior of measurement streams stems from their ability to extract knowledge from the spatio-temporal synergy among the respective trajectories. A linear dynamical system (LDS) [15] constitutes the foundational block upon which more sophisticated models, such as the switching linear dynamical system (SLDS) [16] and the recurrent switching linear dynamical system (rSLDS) [17], are built. Each of these systems can be graphically described by a Dynamic Bayesian Network (DBN), as shown in Fig. 2.

[3]It should be underlined that this work is based on the assumption that measurements exhibit synchronization among them in each time-step. Although this is valid for measurements stemming from PMUs, synchronization is not guaranteed in practice for SCADA-related measurements. Gross errors in time alignment may perniciously affect the intrinsic low-rank representation of data. However, examining this phenomenon is outside the scope of this work.

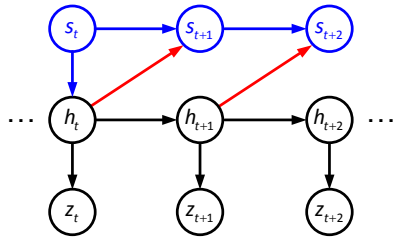

Fig. 2. Dynamic Bayesian Network for LDS (black), SLDS (black, blue), and rSLDS (black, blue, red).

In the case of LDS, the governing dynamics are captured by latent variables h_t, which exhibit smaller dimensionality than z_t, through linear mappings. According to the mathematical formulation outlined in [4], the equations which describe an LDS are shown in Eqs. (4) and (5), as follows

$$h_{t+1} = Fh_t + e_h, \; e_h \sim \mathcal{N}(0, \Lambda), \tag{4}$$

$$z_t = Gh_t + e_z, \; e_z \sim \mathcal{N}(0, \Sigma), \tag{5}$$

where the former expresses temporal correlations among latent variables of measurement streams and the latter captures spatial interactions among different measurements in the same time-step. Gaussian transition and observation noises are assumed, in order to account for the variation in observed and latent variables caused by sundry factors which cannot be predicted from the deterministic transitions and emissions. Based on the factorized form of Fig. 2, the joint distribution of latent and observation variables is expressed as

$$P\left(\{h_t, z_t\}_{t=1}^{T}\right) = P(h_1) \prod_{t=1}^{T-1} P(h_{t+1}|\, h_t) \prod_{t=1}^{T} P(z_t|\, h_t). \tag{6}$$

The parameters which need to be estimated consist of the entries in linear mappings F and G, as well as the noise covariances Λ and Σ. The latter are assumed to constitute diagonal matrices. Term $p(h_{t+1}|\, h_t)$ in Eq. (6) can be expressed as

$$p(h_{t+1}|\, h_t) \propto \det(\Lambda)^{-1/2} \times \exp\left(-\frac{1}{2}(h_{t+1} - Fh_t)^T \Lambda^{-1} (h_{t+1} - Fh_t)\right). \tag{7}$$

The remaining terms can be straightforwardly deduced. Learning parameters $\theta = \{F, \Lambda, G, \Sigma\}$ can be achieved by means of Expectation-Maximization algorithm [15] or in a Bayesian framework using blocked Gibbs sampling by setting conjugate prior distributions over all parameters [18].

Since it is plausible that measurement trajectories may not be sufficiently summarized by a single LDS due to temporal structural changes, the use of SLDS which switches among a predetermined number of LDS can be made. Discrete latent variable s_t serves as the switching enabler for each respective LDS, which is now characterized by its unique parameters F_{s_t}, G_{s_t}, Λ_{s_t}, and Σ_{s_t}. It is common that parameter sharing pertaining to emission matrix G_{s_t} and covariance matrix Σ_{s_t} among the different LDS is imposed. Finally, SLDS is

extended to rSLDS through the explicit incorporation of a dependence of discrete hidden variable s_{t+1} on continuous latent variable h_t, which introduces non-Gaussian potentials $\psi_{t,t+1}(h_t, s_{t+1}) = p(s_{t+1}|h_t)$ with further implications in the sampling of hidden variables and Bayesian parameter learning [17]. It should be stated that the increased expressiveness of SLDS and rSLDS and their capability of globally summarizing the underlying non-linear dynamics come with the inadvertent cost of elevated computational complexity which poses a challenge for real-time applications.

B. Iterative Scheme for Missing Measurement Imputation

We adopt the general framework of [4] for iterative missing values imputation. The main differentiation is that, instead of employing an LDS with parameter learning by means of Expectation-Maximization algorithm, we substitute it with an LDS based on Bayesian parameter learning, as well as an SLDS and an rSLDS. The main steps of iterative missing value imputation are outlined as follows. First, the missing values in the fusion center are initally reconstructed based on linear interpolation and parameters are initialized with respect to their prior distributions. Subsequently, the posterior expectations of latent variables with respect to the aggregation of observed and interpolated measurements, as well as prior parameter values, are computed by leveraging belief propagation. Then, the estimated values of parameters are revised via conjugate updates. Finally, the missing values are imputed by computing their conditional expectation with respect to the values of observed measurements, the posterior expectations of latent variables, and the updated parameter values. The process continues from the step of updating the posterior expectations for latent variables based on the aggregation of observed measurements and the newly imputed values of missing variables until convergence.

IV. Simulation Results

A. Settings

Data provided by 6 PMUs as part of the EPFL Smart Grid project [19] are utilized as input to this work. From each PMU, we retain 3 measurement streams which correspond to three-phase voltage magnitudes. For the simulation setup of the network topology, we consider a PPP for the BSs and their co-located phasor data concentrators (PDCs) with intensity λ_b while the PMUs are randomly spread over the area until all PDCs are associated with at least one PMU, i.e., to achieve traffic saturation conditions. Missing data stem from either the lossy PMU transmissions which may result in data decoding failures at the PDC, as indicated by Eq. (3), or the creation of synthetic occlusions due to a denial-of-service attack or PMU hardware failure. For all experiments, we have utilized the respective codes which are provided in [20] and [21] for LDS, SLDS, and rSLDS. Each experiment was repeated 20 times and we averaged the achieved performance. It is to be noted that reported results in this section stem from the best average performance achieved by the union of the aforementioned dynamical systems in the examined case scenarios.

To quantify the quality of model fitting and data imputation, the root mean squared error (RMSE) is employed as

$$\mathrm{RMSE}\left(\{z_{t,k}, \hat{z}_{t,k}\}\right) = \sqrt{\frac{1}{|\mathcal{O}|}\sum_{(t,k)\in\mathcal{O}}\left(z_{t,k} - \hat{z}_{t,k}\right)^2}, \quad (8)$$

where $z_{t,k}$ and $\hat{z}_{t,k}$ pose a dual role; in the case of examining model fitting, they correspond to the decoded measurements and predicted values by the dynamical model, respectively, for each i-th time-step where observed values exist, while for quantifying the quality of data imputation they pertain to the ground truth of occluded values and the imputed values, respectively, at each i-th time-step where missing values exist.

For a given number of measurement streams M and time-series segment duration T, the dimensionality of continuous latent variables h_t and the highest integer value of discrete latent variable s_t comprise hyper-parameters which need to be judiciously tuned. The performance of dynamical models is contingent upon the selection of suitable values for the aforementioned quantities. Considering computational resources, we carry out a limited hyper-parameter search where the dimensionality of h_t and the maximum value of s_t may take values in $\{8, 10, 12, 14, 16\}$ and in $\{2, 3, 4\}$, respectively, for 10 randomly selected time-series segments.

B. Model Fitting Performance with Lossy PMU Transmissions

In Figs. 3–6, we illustrate the model fitting performance in terms of RMSE for different network configurations. Our goal is to shed light on the impact of various network parameters, as defined in Section II, with respect to the ability of dynamical models to summarize the decoded measurement stream. It is worth noting that, besides the detrimental effect of missing observations due to the PMU connectivity outage, the imperfect decoding accuracy of the observed PMU data also affects the model fitting performance.

Fig. 3 illustrates the RMSE performance with respect to the SINR threshold γ_{th}. It can be observed that the RMSE levels increase with increasing γ_{th}, as the aggregate interference leads to significant signal degradation and results in a high number of PMU links in SINR outage (as also shown in Fig. 1). Subsequently, the increased outage probability in the high γ_{th} regime, impacts the achieved RMSE and leads to degraded performance in the model fitting process.

Fig. 4 depicts the evolution of the RMSE performance for varying BS intensity λ_b. Note that each PDC is considered to be co-located with its respective BS, therefore λ_b also represents the PDC intensity. We observe that model fitting performance improves (i.e., the RMSE decreases) with increasing λ_b. This can be intuitively explained as follows. When the BSs/PDCs are dense enough, the distance between a generic PMU and its corresponding serving BS/PDC decreases, resulting in higher quality of the received measurements and lower outage probability levels.

Fig. 5 depicts the RMSE performance with varying power threshold ρ. It can be observed that RMSE levels decrease with increasing ρ, as a higher received-power threshold at the BS

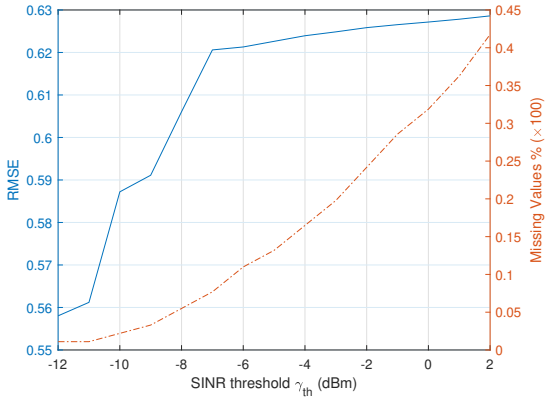

Fig. 3. Model fitting performance for varying levels of SINR threshold γ_{th} (parameter settings: $\lambda_b = 2\text{BS/km}^2$, $\rho = -60\text{dBm}$, $\sigma^2 = -90\text{dBm}$).

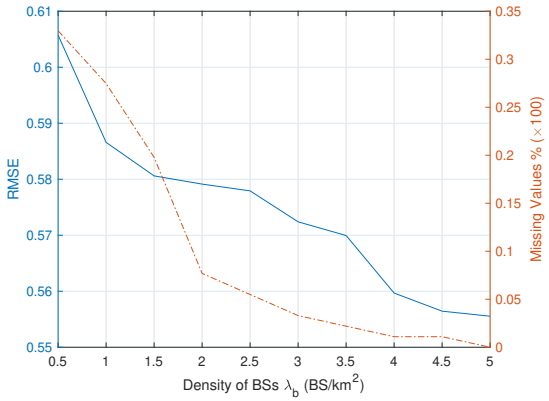

Fig. 4. Model fitting performance for varying levels of BS intensity λ_b (parameter settings: $\gamma_{\text{th}} = -8\text{dBm}$, $\rho = -60\text{dBm}$, $\sigma^2 = -90\text{dBm}$).

leads to lower SINR outage (for relatively low interference levels, as in our case) and lower probability for lost PMU observations. Finally, Fig. 6 shows the impact of the noise power σ^2 on the model fitting performance. As expected, when σ^2 values turn from negligible to significant, the received SINR levels of the PMU measurements deteriorate and RMSE slightly increases along with the SINR outage levels.

C. Reconstruction Performance with Denial-of-Service or PMU Failure

In contrast to the previous subsection, we now assume an ideal network configuration with high data-decoding accuracy for the PMU transmissions. Synthetic dropouts can be made by uniformly selecting space-time points for occlusion, which can be attributed to the effect of denial-of-service or PMU random hardware failure[4]. It is to be noted that the RMSE recorded for the mismatch between decoded measurements and their respective predictions by the dynamical systems exhibited values between 0.14 and 0.16 for all experiments conducted in this subsection; this is in stark contrast to the level of mismatch registered in subsection IV-B, which can be explained by the fact that the dynamical systems are capable of achieving a fairly good fit to the observed entries of measurement streams under ideally decoded transmitted data.

[4]A PMU hardware failure is often associated with the ability to capture or transmit a portion of its associated measurement streams.

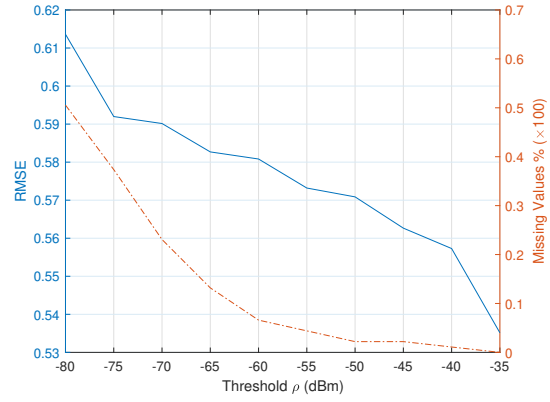

Fig. 5. Model fitting performance for varying levels of power threshold ρ (parameter settings: $\lambda_b = 2\text{BS/km}^2$, $\gamma_{\text{th}} = -8\text{dBm}$, $\sigma^2 = -90\text{dBm}$).

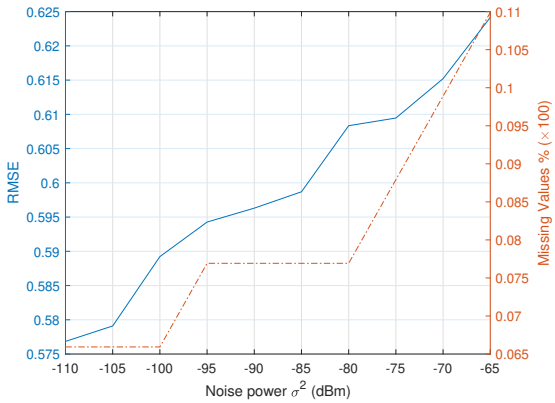

Fig. 6. Model fitting performance for varying levels of noise power σ^2 (parameter settings: $\lambda_b = 2\text{BS/km}^2$, $\gamma_{\text{th}} = -8\text{dBm}$, $\rho = -60\text{dBm}$).

TABLE I
IMPUTATION PERFORMANCE FOR VARYING PERCENTAGE OF MISSING PMU MEASUREMENTS UNDER IDEAL DECODING CONDITIONS

	Missing Values % (×100)				
	0.1	0.2	0.3	0.4	0.5
RMSE	0.1995	0.2054	0.20593	0.21594	0.23388

In Table I, the performance of randomly imputing missing values among measurement streams and time-steps is shown. In this case, dynamical systems are capable of drawing insight from the perfectly decoded measurement values to make valid inferences for the missing data. Imputation performance expectedly registers a decline with rising percentage of missing entries in the fusion center, albeit not at prohibitive levels.

Fig. 7 has been devised by conducting two types of experiments, investigating primarily the effect of ensuing occlusions: (*a*) we randomly opted for 1, 2, and 3 measurement streams to

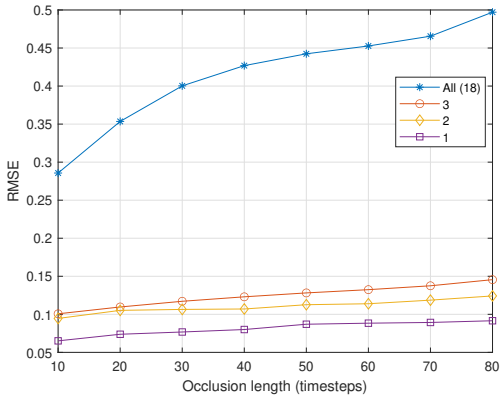

Fig. 7. Imputation performance for varying occlusion length.

exhibit consecutive missing values of varying length starting at a random point in time, and (*b*) we opted for all measurement streams to report missing values for different scenarios of length of ensuing occlusions.

The visual assessment of Fig. 7 suggests that reconstruction of missing values in the case of experiment (*a*) does not suffer from the limitations imposed by potentially unfavorable network configuration parameters. The imputation mechanism takes advantage of spatio-temporal correlations based on the perfectly decoded measurement streams which are consistent with the underlying physical process. Nevertheless, even under such ideal network conditions, the recovery of missing data in experiment (*b*) suffers from the fact that in the absence of any measurement stream for an extended period of time there is non-existent spatial information in each time-step for the dynamical model to exploit. Moreover, the temporal information is limited to the relatively distant time-steps with recorded measurement values. Hence, the predicted values are mainly influenced by the outcome of linear interpolation from the initial stage of data imputation process, registering significant mismatch with respect to the ground truth values.

V. Conclusion

In this paper, we explored the ability of dynamical systems to mine received measurement streams under imperfect and perfect data decoding conditions. It is deduced that when network parameters are such that significant stream-wise distortions are present among the decoded values, the utilized dynamical systems are not capable of capturing the underlying spatio-temporal interactions and, thus, the validity of estimating missing values is curtailed.

In the path forward, we will direct our efforts towards the design of appropriate connectivity enablers for high-quality received measurement streams whose potentially missing values are primed for reconstruction by means of dynamical systems. It is to be noted that since the process of imputation poses elevated merit as a real-time application, the design of network solutions will consider dynamical models which can handle measurement streams of short duration and high dimensionality with guarantees against overfitting.

Acknowledgment

This work has been partially supported by the CHIST-ERA FIREMAN project funded by the Spanish National Foundation (PCI2019-103780), and by the Generalitat de Catalunya under Grant 2017 SGR 891.

References

[1] S. K. Salman, *Introduction to the Smart Grid: Concepts, Technologies and Evolution*, ser. IET Energy Engineering. Institution of Engineering and Technology, 2017.

[2] G. Phadke and J. S. Thorp, *Synchronized phasor measurements and their applications*. New York: Springer Science and Business Media, 2008.

[3] Y. Zhou, R. Arghandeh, and C. J. Spanos, "Online learning of Contextual Hidden Markov Models for temporal-spatial data analysis," in *2016 IEEE 55th Conf. on Decision and Control*, Las Vegas, NV, 2016, pp. 6335–6341.

[4] L. Li, J. McCann, N. S. Pollard, and C. Faloutsos, "DynaMMo: mining and summarization of coevolving sequences with missing values," in *Proc. of the 15th ACM Int. Conf. on Knowledge Discovery and Data Mining*, pp. 507–515, 2009.

[5] M. Brand M, "Incremental singular value decomposition of uncertain data with missing values," in *Proc. of the 7th Eur. Conf. on Computer Vision*, pp. 707–720, 2002.

[6] N. Srebro and T. Jaakkola, "Weighted low-rank approximations," in *Proc. of the 20th Int. Conf. on Machine Learning*, pp. 720–727, 2003.

[7] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2019.

[8] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing Data Recovery by Exploiting Low-Dimensionality in Power System Synchrophasor Measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1006–1013, Mar. 2016.

[9] Y. Hao, M. Wang, J. H. Chow, E. Farantatos, and M. Patel, "Modelless Data Quality Improvement of Streaming Synchrophasor Measurements by Exploiting the Low-Rank Hankel Structure," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6966–6977, Nov. 2018.

[10] C. Genes, I. Esnaola, S. M. Perlaza, L. F. Ochoa, and D. Coca, "Robust Recovery of Missing Data in Electricity Distribution Systems," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4057–4067, July 2019.

[11] R. Pourramezan, H. Karimi, J. Mahseredjian, and M. Paolone, "Real-Time Processing and Quality Improvement of Synchrophasor Data," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3313–3324, July 2020.

[12] J. G. Andrews, A. K. Gupta, and H. S. Dhillon, "A Primer on Cellular Network Analysis Using Stochastic Geometry," *arXiv:1604.03183*, 2016.

[13] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2012.

[14] M. Gharbieh, H. ElSawy, A. Bader, and M. Alouini, "Tractable Stochastic Geometry Model for IoT Access in LTE Networks," *in Proc. of the 2016 IEEE Global Commun. Conf.*, pp. 1–7, 2016.

[15] Z. Ghahramani and G. Hinton. Parameter estimation for linear dynamical systems. Technical Report, Dept. Comp. Sci., Univ. Toronto, 1996.

[16] V. Pavlovic, J. M. Rehg, and J. MacCormick, "Learning Switching Linear Models of Human Motion," in *Proc. of Advances in Neural Information Processing Systems*, pp. 981–987, 2001.

[17] S. Linderman, M. Johnson, A. Miller, R. Adams, D. Blei, and L. Paninski, "Bayesian Learning and Inference in Recurrent Switching Linear Dynamical Systems," in *Proc. of the 20th Int. Conf. on Artificial Intelligence and Statistics*, pp. 914–922, 2017.

[18] A. Wills, T. B. Schön, F. Lindsten, and B. Ninness, "Estimation of linear systems using a Gibbs sampler," *in Proc. of the 16th IFAC Symposium on System Identification*, pp. 203–208, 2012.

[19] EPFL Smart Grid. [Online]. Available: http://nanotera-stg2.epfl.ch/data/. Accessed: 25 May 2020.

[20] PyLDS: Bayesian inference for linear dynamical systems. [Online]. Available: https://github.com/mattjj/pylds.

[21] SSM: Bayesian learning and inference for state space models. [Online]. Available: https://github.com/slinderman/ssm.