Bayesian Hankel Matrix Completion for Synchrophasor Data Recovery

Ming Yi

Postdoctoral Research Scientist
Data Science Institute

Columbia University & Rensselaer Polytechnic Institute

NREL Autonomous Energy Systems Workshop
September 7, 2023

Acknowledgements: RPI: Meng Wang, EPRI: Evangelos Farantatos, Tapas Barik, ANL:Tianqi Hong (Now at UGA), Dongbo Zhao (Now at Eaton)

Outline

Outline

Phasor Measurement Units (PMU)

- PMUs provide synchronized measurements of bus voltage phasors, line current phasors, and frequency in power systems.
- High sampling rate of 30 or 60 samples per second per channel. (One sample every 2-4 seconds in the traditional SCADA systems)
- 2500+ PMUs in North America in 2017.
- Provide better visibility of dynamics of power system operations.

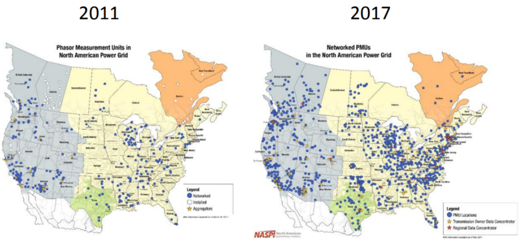


Figure 1: The comparison of installed PMUs in North America between 2011 and 2017 [1]

[1] https://www.sgsma2021.org/wp-content/uploads/2021/05/QTech-SGSMA-Silver-Sponsor-Slides-May-25-2021-Updated2.pdf

PMU Data Quality Issues

- Data losses and errors resulting from communication congestions and device malfunction.
- California Independent System Operator [CAISO 2011] reported that 10%-17% of data in 2011 had availability and quality issues.
- Limited incorporation into the real-time operations

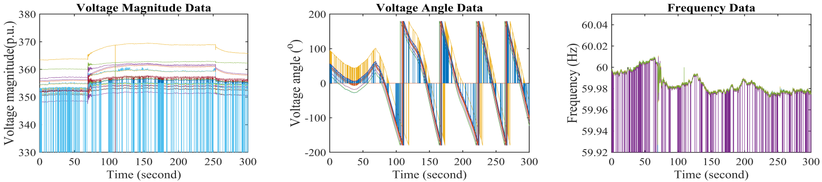


Figure 2: Missing data examples. (Data from New York Independent System Operator (NYISO))

Problem Formulation

Y: ground-truth measurements with m channels and n time instants,

$$Y = [y_1, y_2, ..., y_n] \in \mathbb{R}^{m \times n}$$

Matrix $Y^o \in \mathbb{R}^{m \times n}$: the observed data

$$Y^o_{i,j} = Y_{i,j} + E_{i,j} + N_{i,j} \quad (i,j) \in \Omega$$

$E_{i,j}$: bad data, $N_{i,j}$: noise, Ω: observed entries.

The **objective** of robust matrix completion is to recover Y from partial observations $Y^o_{i,j}$ that contain bad data and noise.

Low-Rank Matrix Completion

The **objective** of matrix completion is to recover data matrix from partial observations.

$$\begin{bmatrix} 5 & ? & 7 & ? \\ 10 & 12 & ? & 16 \\ ? & 18 & 21 & ? \\ 20 & ? & ? & 32 \end{bmatrix}$$

Without a prior knowledge of the matrix, this problem is ill-conditioned.

Low-Rank Matrix Completion

The **objective** of matrix completion is to recover data matrix from partial observations.

$$\begin{bmatrix} 5 & ? & 7 & ? \\ 10 & 12 & ? & 16 \\ ? & 18 & 21 & ? \\ 20 & ? & ? & 32 \end{bmatrix}$$

Without a prior knowledge of the matrix, this problem is ill-conditioned.

Luckily, many data matrices exhibit low-dimensional structures.

Low Dimensionality of PMU Data

- 6 PMUs measure 37 voltage/current phasors. 30 samples/second. (Data source: New York Power Authority (NYPA))
- A generator trip event in New York state.

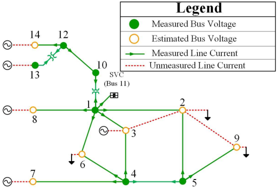


Figure 3: PMUs in Central NY Power Systems [Gao et al. 2016]

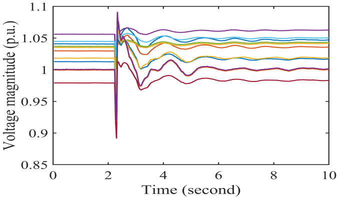


Figure 4: Voltage magnitudes of PMU data [Hao et al. 2018]

The standard low-rank methods cannot handle the simultaneous and consecutive missing data, i.e., M2 and M3 modes, even when the data matrix is low rank.

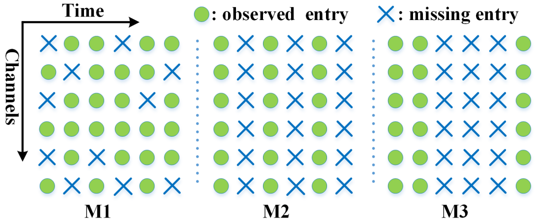


Figure 5: Three different missing data modes

Y: m channels and n time instants,

$$Y = [y_1, y_2, ..., y_n] \in \mathbb{R}^{m \times n}$$

Hankel structure:

$$\mathcal{H}_{n_2}(Y) = \begin{bmatrix} y_1 & y_2 & \cdots & y_{n_1} \\ y_2 & y_3 & \cdots & y_{n_1+1} \\ \vdots & \vdots & \ldots & \vdots \\ y_{n_2} & y_{n_2+1} & \cdots & y_n \end{bmatrix}$$

$\mathcal{H}_{n_2}(Y) \in \mathbb{R}^{mn_2 \times n_1}$ and $n_1 + n_2 = n + 1$

Hankel low-rank property:

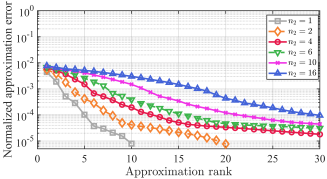


Figure 6: Normalized approximation errors of Hankel matrices

Outline

Related Work

- Low-rank matrix completion [Gao et al. 2016; Hao et al. 2018; Zhang et al. 2018]
  - Lack confidence measure for the returned results
  - Provide theoretical guarantee but the bound underestimates the methods' capabilities
- Bayesian matrix completion [Babacan et al. 2012, Chen et al. 2021]
  - Fail when simultaneous and consecutive data is missing and/or corrupted across all channels
  - Have no **uncertainty modeling** for the returned results

Methodology [2]

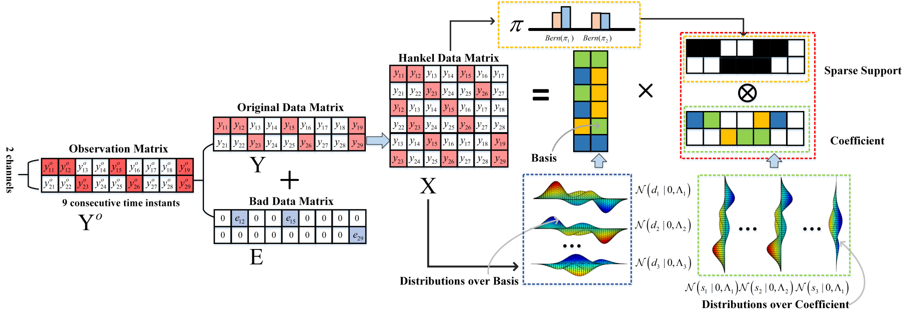


Figure 7: An overall illustration of the proposed approach.

[2] Ming Yi, Meng Wang, Evangelos Farantatos, and Tapas Barik. "Bayesian Robust Hankel Matrix Completion with Uncertainty Modeling for Synchrophasor Data Recovery" *ACM Energy Informatics Review*, 2022

Bayesian Hankel Matrix Completion [3]

Hierarchical model:

$$Y^o_{i,j} = (\mathcal{H}^\dagger X)_{i,j} + E_{i,j} + N_{i,j} \quad (i,j) \in \Omega,$$

$$X = UV \Leftrightarrow x_{.q} = UV_{.q},$$

$(\mathcal{H}^\dagger X)_{i,j}$: the inverse of Hankel matrix
$V_{.q}$: qth column in V, $x_{.q}$: qth column in X
$U_{p.}$: pth row in U, I_K: $K \times K$ identity matrix

$$\begin{aligned}
&U_{p.} \sim \mathcal{N}(\mathbf{0}, \lambda_d^{-1} I_K), && \\
&V_{.q} \sim \mathcal{N}(\mathbf{0}, \gamma_s^{-1} I_K) && \gamma_s \sim \Gamma(c_0, d_0) \\
&E_{i,j} \sim \mathcal{N}(0, \beta_{i,j}^{-1}) \quad (i,j) \in \Omega && \beta_{i,j} \sim \Gamma(g_0, h_0) \\
&N_{i,j} \sim \mathcal{N}(0, \gamma_\epsilon^{-1}) && \gamma_\epsilon \sim \Gamma(e_0, f_0)
\end{aligned}$$

[3] Ming Yi, Meng Wang, Evangelos Farantatos and Tapas Barik. "Bayesian Robust Hankel Matrix Completion with Uncertainty Modeling for Synchrophasor Data Recovery" *ACM Energy Informatics Review*, 2022

Variational Inference

Given Y^o_Ω, we aim to compute the posterior $P(\Theta, Y|Y^o_\Omega)$. From the Bayes' rule,

$$P(\Theta, Y|Y^o_\Omega) = \frac{P(\Theta, Y, Y^o_\Omega)}{P(Y^o_\Omega)}$$

Θ denotes all the latent variables. The posterior distribution is intractable.

Variational Inference

Given Y^o_Ω, we aim to compute the posterior $P(\Theta, Y|Y^o_\Omega)$. From the Bayes' rule,

$$P(\Theta, Y|Y^o_\Omega) = \frac{P(\Theta, Y, Y^o_\Omega)}{P(Y^o_\Omega)}$$

Θ denotes all the latent variables. The posterior distribution is intractable. The mean field variational inference [Bishop 2006] is employed to approximate $P(\Theta, Y|Y^o_\Omega)$ by the variational distribution $q(\Theta)$.

Mean field theory assumes that elements in Θ are mutually independent

$$q(\Theta) = q(U)q(V)q(E)q(\beta)q(\gamma_s)q(\gamma_\epsilon)$$

Variational Inference

Variational inference: find the closest approximation $q(\Theta)$ to $P(\Theta, Y_\Omega | Y_\Omega^o)$

$$\begin{aligned} q(\Theta) &= \operatorname*{argmin}_{q(\Theta)} \mathbb{KL}(q(\Theta) || P(\Theta, Y | Y_\Omega^o)) \\ &= \operatorname*{argmax}_{q(\Theta)} \mathbb{E}[\ln P(\Theta, Y, Y_\Omega^o)] - \mathbb{E}[\ln q(\Theta)] \end{aligned}$$

The optimal $q(U_{p\cdot})$ which maximizes the objective function is

$$\ln q(U_{p\cdot}) = \mathbb{E}_{q(\Theta \backslash U_{p\cdot})}[\ln \ P(\Theta, Y, Y_\Omega^o)] + \text{constant}$$

$\mathbb{E}_{q(\Theta \backslash U_{p\cdot})}$: expectation with all the latent variables except U_p.

Uncertainty Modeling

Goal: estimate the distribution of $Y_{i,j}$

Mean: the estimation of $Y_{i,j}$
Variance: model uncertainty of the estimation

It is **intractable** to obtain the closed-form of the distribution of $Y_{i,j}$

$$\begin{aligned}
\mathbb{E}[Y_{i,j}] &= \int p(Y_{i,j}|Y_{\Omega}^{o})Y_{i,j}dY_{i,j} \\
&= \int (\int p(Y_{i,j}|\theta)p(\theta|Y_{\Omega}^{o})d\theta)Y_{i,j}dY_{i,j} \\
&= \int \mathbb{E}_{p(Y_{i,j}|\theta)}[Y_{i,j}]p(\theta|Y_{\Omega}^{o})d\theta \\
&= \int f^{\theta}(Y_{i,j})p(\theta|Y_{\Omega}^{o})d\theta \\
&\approx \frac{1}{L}\sum_{l=1}^{l=L} f^{\theta_l}(Y_{i,j}) \quad \theta_l \sim q(\theta|Y_{\Omega}^{o})
\end{aligned}$$

Uncertainty Modeling

Goal: estimate the distribution of $Y_{i,j}$

Mean: the estimation of $Y_{i,j}$
Variance: model uncertainty of the estimation

It is **intractable** to obtain the closed-form of the distribution of $Y_{i,j}$

$$
\begin{aligned}
\mathbb{E}[Y_{i,j}] &= \int p(Y_{i,j}|Y_\Omega^o) Y_{i,j} dY_{i,j} \\
&= \int (\int p(Y_{i,j}|\theta) p(\theta|Y_\Omega^o) d\theta) Y_{i,j} dY_{i,j} \\
&= \int \mathbb{E}_{p(Y_{i,j}|\theta)}[Y_{i,j}] p(\theta|Y_\Omega^o) d\theta \\
&= \int f^\theta(Y_{i,j}) p(\theta|Y_\Omega^o) d\theta \\
&\approx \frac{1}{L} \sum_{l=1}^{l=L} f^{\theta_l}(Y_{i,j}) \quad \theta_l \sim q(\theta|Y_\Omega^o)
\end{aligned}
$$

Monte-Carlo integration [Paisley et al. 2012] is employed to compute the mean and variance approximately

Uncertainty Modeling

$$f^{\theta}(Y_{i,j}) = \mathcal{H}^{\dagger}(UV)_{i,j} \quad \theta = \{U, V, \gamma_\epsilon\}$$

Predictive mean

$$\hat{Y}_{i,j} = \mathbb{E}[Y_{i,j}] \approx \frac{1}{L}\sum_{l=1}^{l=L} f^{\theta_l}(Y_{i,j}) \quad \theta_l \sim q(\theta|Y^o_\Omega)$$

L : number of Monte-Carlo samples
Predictive variance

$$\begin{aligned} Var[Y_{i,j}] &= \mathbb{E}[Y^2_{i,j}] - \mathbb{E}[Y_{i,j}]^2 \\ &\approx \frac{1}{L}\sum_{l=1}^{l=L}\frac{1}{\gamma_\epsilon} + \frac{1}{L}\sum_{l=1}^{l=L} f^{\theta_l}(Y_{i,j})^2 - (\frac{1}{L}\sum_{l=1}^{l=L} f^{\theta_l}(Y_{i,j}))^2 \quad \theta_l \sim q(\theta|Y^o_\Omega) \end{aligned}$$

$\mathbb{E}[Y_{i,j}]$: an estimate $\hat{Y}_{i,j}$ of $Y_{i,j}$, $Var[Y_{i,j}]$: uncertainty index of the estimation.

Block Processing

Handling streaming data in real-time:

- Truncate the measurements into blocks and process each time block separately.
- Use a sliding window with length n and step size s.

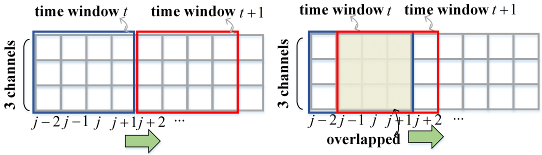


Figure 8: Non-overlapping and overlapping sliding windows

Numerical Experiments (Practical PMU Data)

- Case 1: 20% M2 missing data. Additional Gaussian noise $\mathcal{N}(0, 0.003^2)$ is added during time 5.6 to 6.6 seconds
- Case 2: 20% M1 and 15% B1 bad data

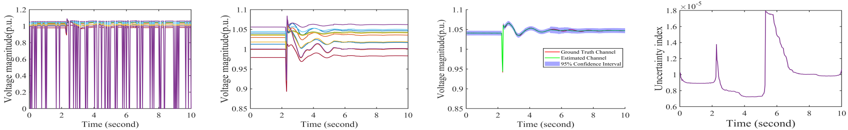


Figure 9: The recovery performance on case 1, 4 figures are (a) the observed data, (b) the estimated data, (c) the estimated data in one channel (d) the uncertainty index for the channel in (c)

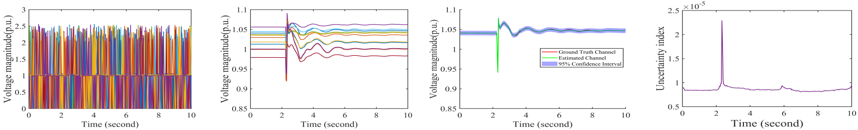


Figure 10: The recovery performance on case 2

- Recovery performance degrades significantly when the power system is experiencing nonlinear dynamics during significant events.

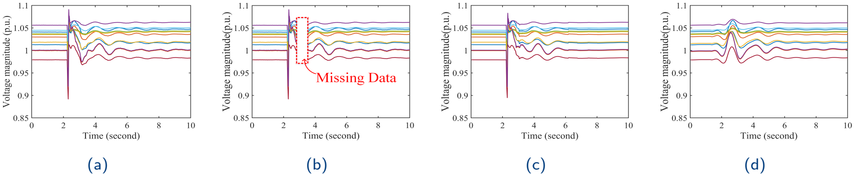


Figure 11: (a) the ground truth data, (b) the observed data, the estimated data by (c) the proposed method, (d) the deterministic low-rank Hankel data recovery method [Zhang et al. 2019].

Outline

Related Work

- Bayesian synchrophasor data recovery [Yi et al. 2022]
  - Time-consuming for block processing
  - The linear dynamical system assumption becomes inaccurate when handling nonlinear dynamics
- High-rank matrix completion [Ongie et al. 2017; Fan et al. 2018; Fan et al. 2019]
  - Lacks confidence measure for the returned results
  - Cannot handle simultaneous and consecutive missing data across all channels
  - Cannot handle the bad data
- Nonlinear synchrophasor data recovery [Hao et al. 2019]
  - Lacks confidence measure for the returned results
  - Sensitive to the parameter selection

Kernel Idea

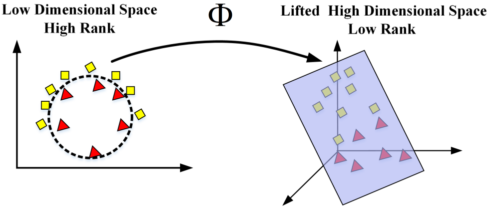


Figure 12: An overall illustration of the kernel method.

$$y_{1i}^2 + y_{2i}^2 = 1$$

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{21} & \cdots & y_{2n} \end{bmatrix}$$

full rank

$$\Rightarrow$$

$$z_{1i} + 0 \times z_{2i} + z_{3i} = 1$$

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{21} & \cdots & z_{2n} \\ z_{31} & z_{31} & \cdots & z_{3n} \end{bmatrix}$$

rank is two

Mapping function: $\phi : R^2 \longrightarrow R^3$

$$(y_{1i}, y_{2i}) \mapsto (z_{1i}, z_{2i}, z_{3i}) = (y_{1i}^2, \sqrt{2} y_{1i} y_{2i}, y_{2i}^2)$$

Lifted Low-rank Hankel Property of PMU Data

Lifted Hankel structure:

$$\mathcal{H}_{n_2}(Z) = \begin{bmatrix} z_1 & z_2 & \cdots & z_{n_1} \\ z_2 & z_3 & \cdots & z_{n_1+1} \\ \vdots & \vdots & \cdots & \vdots \\ z_{n_2} & z_{n_2+1} & \cdots & z_n \end{bmatrix}$$

where $z_i = \phi(y_i)$, $\mathcal{H}_{n_2}(Z) \in \mathbb{R}^{Mn_2 \times n_1}$

$$\mathcal{K}_{YY}(i,j) = \phi(y_i)^T \phi(y_j) = \exp(-\frac{1}{2c}||y_i - y_j||_2^2)$$

c: kernel parameter

Lifted Low-rank Hankel Property of PMU Data

Lifted Hankel structure:

$$\mathcal{H}_{n_2}(Z) = \begin{bmatrix} z_1 & z_2 & \cdots & z_{n_1} \\ z_2 & z_3 & \cdots & z_{n_1+1} \\ \vdots & \vdots & \cdots & \vdots \\ z_{n_2} & z_{n_2+1} & \cdots & z_n \end{bmatrix}$$

where $z_i = \phi(y_i)$, $\mathcal{H}_{n_2}(Z) \in \mathbb{R}^{Mn_2 \times n_1}$

$$\mathcal{K}_{YY}(i,j) = \phi(y_i)^T \phi(y_j) = \exp(-\frac{1}{2c}||y_i - y_j||_2^2)$$

c: kernel parameter

Lifted low-rank Hankel property:

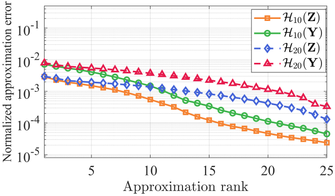


Figure 13: Normalized approximation errors of Hankel/lifted Hankel matrices

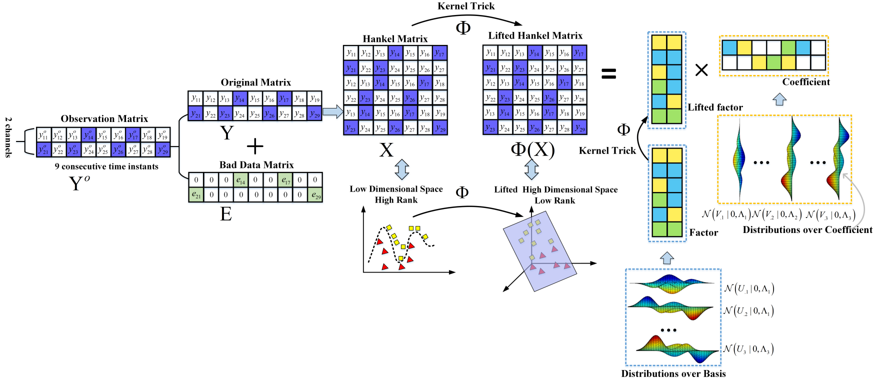


Figure 14: An overall illustration of the Bayesian high-rank matrix completion approach.

[4] Ming Yi, Meng Wang, Tianqi Hong and Dongbo Zhao, "Bayesian High-Rank Hankel Matrix Completion for Nonlinear Synchrophasor Data Recovery" *IEEE Transactions on Power System*, 2023

Bayesian High-Rank Hankel Matrix Completion

Hierarchical model:

$$Y^o_{i,j} = (\mathcal{H}^\dagger X)_{i,j} + E_{i,j} + N_{i,j} \quad (i,j) \in \Omega,$$

$$\Phi(X) = \Phi(U)V \Leftrightarrow \Phi(x_{.q}) = \Phi(U)V_{.q},$$

$(\mathcal{H}^\dagger X)_{i,j}$: the inverse of Hankel matrix, $V_{.q}$: qth column in V, $x_{.q}$: qth column in X.
$U_{.k}$: kth column in U, I_K: $K \times K$ identity matrix

$$U_{.k} \sim \mathcal{N}(0, \frac{1}{\gamma_u} I_m),$$

$$X_{.q} \sim \mathcal{N}(0, \frac{1}{\gamma_x} I_m)$$

$$V_{.q} \sim \mathcal{N}(0, \frac{1}{\gamma_v} I_K)$$

$$N_{i,j} \sim \mathcal{N}(0, \frac{1}{\gamma_y})$$

$$E_{i,j} \sim \mathcal{N}(0, \beta^{-1}_{i,j}) \quad (i,j) \in \Omega$$

$$\beta_{i,j} \sim \Gamma(g_0, h_0)$$

$$\gamma_y \sim \Gamma(e_0, f_0)$$

$$\mathcal{K}_{XU}(q,k) = \Phi(X_{.q})^T \Phi(U_{.k}) = \exp(-\frac{1}{2c_2}||X_{.q} - U_{.k}||^2_2)$$

$$\mathcal{K}_{UU}(i,j) = \Phi(U_{.i})^T \Phi(U_{.j}) = \exp(-\frac{1}{2c_3}||U_{.i} - U_{.j}||^2_2)$$

Technical Challenges

Variational inference: finds the closest approximation $q(\Theta)$ to $P(\Theta, Y_\Omega | Y_\Omega^o)$

$$q(\Theta) = \operatorname*{argmax}_{q(\Theta)} \mathbb{E}[\ln P(\Theta, Y, Y_\Omega^o)] - \mathbb{E}[\ln q(\Theta)]$$

The optimal $q(\theta_i)$ which maximizes the objective function is

$$q(\Theta_i) = \operatorname*{argmax}_{q(\Theta_i)} \Big(\int q(\Theta_i) \mathbb{E}_{q(\Theta \setminus \Theta_i)}[\ln p(\Theta, Y, Y_\Omega^o)] d(\Theta_i) - \int q(\Theta_i) \ln q(\Theta_i) d\Theta_i \Big)$$

Technical Challenges

$U_{.k}$ and $X_{.q}$ are lifted to a higher dimensional space via the kernel method, **no analytical solutions**.

Technical Challenges

$U_{.k}$ and $X_{.q}$ are lifted to a higher dimensional space via the kernel method, **no analytical solutions**.

We assume $U_{.k}$ and $X_{.q}$ are drawn from Gaussian distributions. Take $U_{.k}$ as an example,

$$q(U_{.k}) \sim \mathcal{N}(\mu_{U_{.k}}, \Sigma_{U_{.k}})$$

The problem is simplified to find the corresponding mean and the variance of each variable.

Technical Challenges

$U_{.k}$ and $X_{.q}$ are lifted to a higher dimensional space via the kernel method, **no analytical solutions**.

We assume $U_{.k}$ and $X_{.q}$ are drawn from Gaussian distributions. Take $U_{.k}$ as an example,

$$q(U_{.k}) \sim \mathcal{N}(\mu_{U_{.k}}, \Sigma_{U_{.k}})$$

The problem is simplified to find the corresponding mean and the variance of each variable.

How to differentiate and optimize the objective function with respect to the mean and the variance?

Reparameterization Trick

The reparameterization trick [Kingma et al. 2013] is employed here to make the Monte-Carlo estimation differentiable with respect to $U_{.k}$.

$$
\begin{aligned}
q(U_{.k}) &\overset{(a)}{=} \arg\max_{q(U_{.k})} \int q(U_{.k}) \mathbb{E}_{q(\Theta \setminus U_{.k})}[\ln p(\Theta, Y, Y_\Omega^o)] d(U_{.k}) - \int q(U_{.k}) \ln q(U_{.k}) dU_{.k}. \\
&\overset{(b)}{=} \arg\max_{q(U_{.k})} \int q(U_{.k}) \mathbb{E}_{q(\Theta \setminus U_{.k})}[\ln p(\Theta, Y, Y_\Omega^o | U_{.k})] d(U_{.k}) - \mathbb{KL}(q(U_{.k}) | p(U_{.k})). \\
&\overset{(c)}{\approx} \arg\max_{q(U_{.k})} \frac{1}{J} \sum_l^J \mathbb{E}_{q(\Theta \setminus U_{.k})}[\ln p(\Theta, Y, Y_\Omega^o | U_{.k}^{(l)})] - \mathbb{KL}(q(U_{.k}) | p(U_{.k})).
\end{aligned}
$$

$U_{.k}^{(l)} = \mu_{U_{.k}} + \Sigma_{U_{.k}} \epsilon^{(l)}$
ϵ: auxiliary noise variable, $\epsilon^{(l)} \sim \mathcal{N}(0, I_{mn_2})$.

Reparameterization Trick

The reparameterization trick [Kingma et al. 2013] is employed here to make the Monte-Carlo estimation differentiable with respect to $U_{.k}$.

$$
\begin{aligned}
q(U_{.k}) &\overset{(a)}{=} \arg\max_{q(U_{.k})} \int q(U_{.k})\mathbb{E}_{q(\Theta\setminus U_{.k})}[\ln p(\Theta, Y, Y_{\Omega}^{o})]d(U_{.k}) - \int q(U_{.k})\ln q(U_{.k})dU_{.k}. \\
&\overset{(b)}{=} \arg\max_{q(U_{.k})} \int q(U_{.k})\mathbb{E}_{q(\Theta\setminus U_{.k})}[\ln p(\Theta, Y, Y_{\Omega}^{o}|U_{.k})]d(U_{.k}) - \mathbb{KL}(q(U_{.k})|p(U_{.k})). \\
&\overset{(c)}{\approx} \arg\max_{q(U_{.k})} \frac{1}{J}\sum_{l}^{J} \mathbb{E}_{q(\Theta\setminus U_{.k})}[\ln p(\Theta, Y, Y_{\Omega}^{o}|U_{.k}^{(l)})] - \mathbb{KL}(q(U_{.k})|p(U_{.k})).
\end{aligned}
$$

$U_{.k}^{(l)} = \mu_{U_{.k}} + \Sigma_{U_{.k}}\epsilon^{(l)}$

ϵ: auxiliary noise variable, $\epsilon^{(l)} \sim \mathcal{N}(0, I_{mn_2})$.

Uncertainty Modeling

Predictive mean:

$$\hat{Y}_{i,j} = \mathbb{E}[Y_{i,j}] \approx \frac{1}{L}\sum_{l=1}^{L}(\mathcal{H}^{\dagger}X^{(l)})_{i,j} \quad X^{(l)} \sim q(X|Y^{o}_{\Omega})$$

L : number of Monte-Carlo samples
Predictive variance:

$$\begin{aligned} Var[Y_{i,j}] &= \mathbb{E}[Y_{i,j}^2] - \mathbb{E}[Y_{i,j}]^2 \\ &\approx \frac{1}{L}\sum_{l=1}^{L}\frac{1}{\gamma_y^{(l)}} + \frac{1}{L}\sum_{l=1}^{L}(\mathcal{H}^{\dagger}X^{(l)})_{i,j}^2 - (\frac{1}{L}\sum_{l=1}^{L}(\mathcal{H}^{\dagger}X^{(l)})_{i,j})^2 \end{aligned}$$

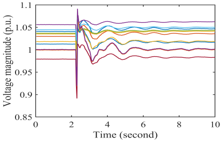
$\mathbb{E}[Y_{i,j}]$: an estimate $\hat{Y}_{i,j}$ of $Y_{i,j}$
Uncertainty index

$$U_{\text{index}} = (\sum_{i=1}^{m}\sum_{j=1}^{n} Var[Y_{i,j}])/(mn)$$

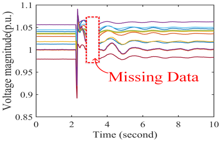
Numerical Experiments (Practical PMU Data)

Case study: 6.7% M3 missing data.

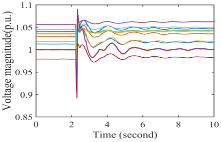
BHMC-S: Bayesian low-rank Hankel method [Yi et al. 2022]; AM-FIHT: deterministic low-rank Hankel method [Zhang et al 2019]; SDR-K: deterministic nonlinear streaming method [Hao et al. 2019]



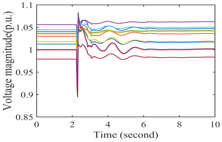
(a)



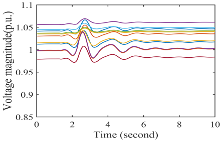
(b)



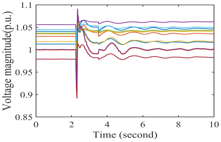
(c)



(d)



(e)



(f)

Figure 15: (a) ground truth, (b) the observed data, the estimated data by (c) the proposed method, (d) the BHMC-S method, (e) the AM-FIHT method. (f) the SDR-K method.

Numerical Experiments (Practical PMU Data)

Table 1: The recovery performance of recorded PMU data on 6.7% M3 mode

Method	Proposed	BHMC-S	AM-FIHT	SDR-K
NEE	8.3×10^{-4}	3.0×10^{-3}	6.0×10^{-3}	2.1×10^{-3}

Uncertainty index

Table 2: The recovery error and the uncertainty index on 5% B1 with varying missing data percentage of M2

Missing rate	5	15	25	35	45
NEE	0.0019	0.0037	0.0057	0.0060	0.18
$\mathrm{U}_{\mathrm{index}}$	2.6×10^{-5}	4.8×10^{-5}	1.5×10^{-4}	4.5×10^{-4}	1.1×10^{-2}

Table 3: The recovery error and the uncertainty index on 5% M2 with varying bad data percentage of B1

Bad rate	5	15	25	35	45
NEE	0.0019	0.0091	0.016	0.017	0.032
$\mathrm{U}_{\mathrm{index}}$	2.6×10^{-5}	5.8×10^{-5}	7.0×10^{-5}	8.3×10^{-4}	6.2×10^{-3}

Summary

- Snchrophasor data recovery with uncertainty modeling
  - Incorporate the Hankel structure to handle simultaneous and consecutive data loss/corruption
  - Provide an uncertainty index to evaluate the confidence of each estimation
- Nonlinear snchrophasor data recovery
  - Exploit the kernel method to recover data during nonlinear dynamics

Future Work

1. Synchrophasor data recovery:
   - Incorporate other system information to further improve the estimation performance
   - Extend the Hankel structure to more general Bayesian tensor recovery problems

Thank you!

REFERENCES

CAISO, “Five year synchrophasor plan,” https://www.caiso.com/Documents/FiveYearSynchrophasorPlan.pdf(AccessedNov.20,2019), 2011.

P. Gao, M. Wang, J. H. Chow, S. G. Ghiocel, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky, “Identification of successive “unobservable” cyber data attacks in power systems.” *IEEE Transactions on Signal Processing*, vol. 64, no. 21, pp. 5557–5570, Nov. 2016.

Y. Hao, M. Wang, and J. H. Chow, “A study of likelihood analysis of cyber data attacks to power systems,” pp. 3191–3202, 2018.

S. Zhang, M. Wang, and A, “Correction of simultaneous bad measurements by exploiting the low-rank hankel structure,” in *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2018, pp. 646–650.

S. D. Babacan, M. Luessi, R. Molina, and A. K. Katsaggelos, “Sparse bayesian methods for low-rank matrix estimation,” *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 3964–3977, 2012.

X. Chen, L. Sun, and A, "Bayesian temporal factorization for multidimensional time series prediction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

C. M. Bishop, "Pattern recognition," *Machine learning*, vol. 128, no. 9, 2006.

J. Paisley, D. M. Blei, and M. I. Jordan, "Variational bayesian inference with stochastic search," in *International Conference on Machine Learning*, 2012, pp. 1363–1370.

M. Yi, M. Wang, E. Farantatos, and T. Barik, "Bayesian robust hankel matrix completion with uncertainty modeling for synchrophasor data recovery," *ACM SIGENERGY Energy Informatics Review*, vol. 2, no. 1, pp. 1–19, 2022.

G. Ongie, R. Willett, R. D. Nowak, and L. Balzano, "Algebraic variety models for high-rank matrix completion," in *International Conference on Machine Learning*. PMLR, 2017, pp. 2691–2700.

J. Fan, T. W. Chow, and A, "Non-linear matrix completion," *Pattern Recognition*, vol. 77, pp. 378–394, 2018.

J. Fan, M. Udell, and A, "Online high rank matrix completion," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 8690–8698.

Y. Hao, M. Wang, and J. H. Chow, "Modeless streaming synchrophasor data recovery in nonlinear systems," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1166–1177, 2019.

D. P. Kingma, M. Welling, and A, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.