Smart Meter Data-driven Enhancement of Behind-the-Meter Observability

Fankun Bu, Kaveh Dehghanpour, Yuxuan Yuan, Zhaoyu Wang, and Yifei Guo

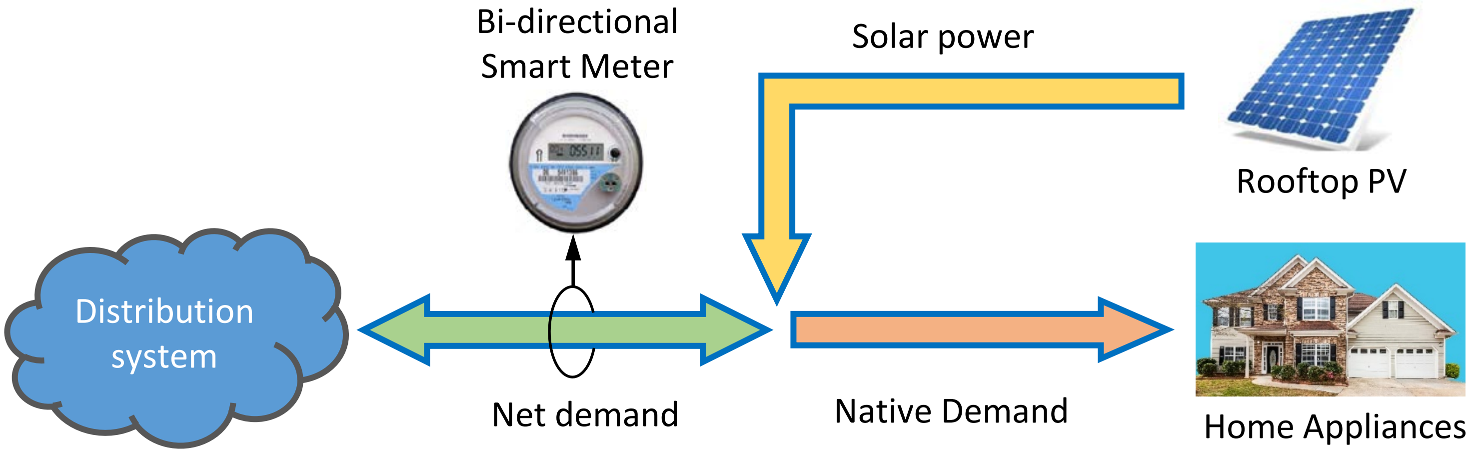
Iowa State University

Presentation Outline

- **Background and Motivation**
- **Two Findings Supporting Our Proposed Approach**
- **Overall Framework**
- **Formulating An MLE Optimization**
- **Conclusion**

Background

- **Behind-the-meter (BTM) installation of rooftop PV**

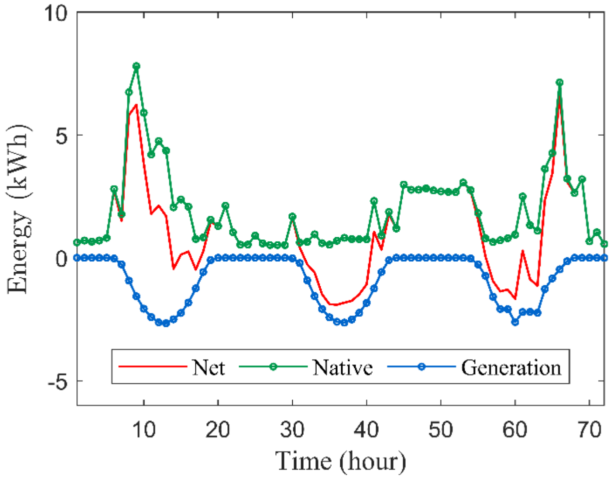


$$\text{Net demand} = \text{Native demand} \; - \text{solar generation}$$

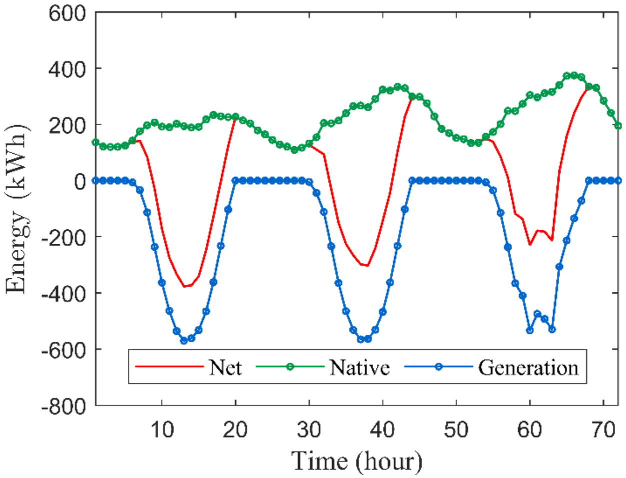
- In most cases, utilities *only* measures the net demand.
- BTM solar generation and native demand are usually *invisible* to utilities.

Motivation

- **Invisibility of solar generation and native demand**



An individual customer with PV

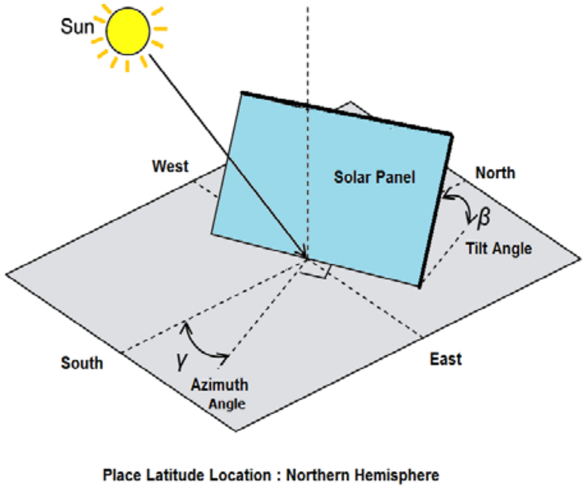


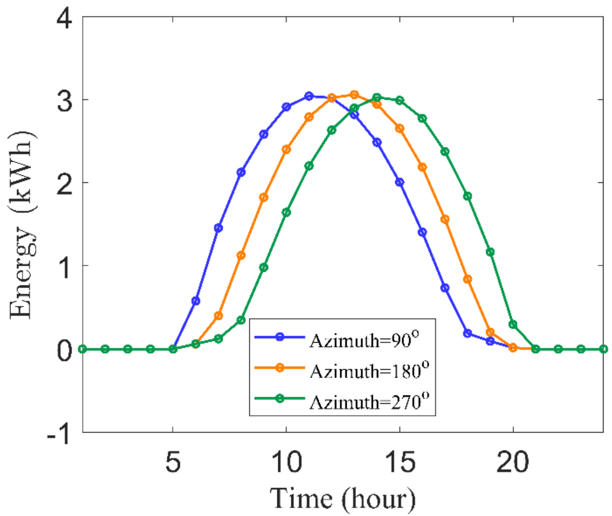
200 customers with PVs

Causes **Problems** for (1) solar generation and native load monitoring, (2) load forecasting, (3) service restoration, and (4) demand response, etc.

Goal: Disaggregate the unknown native demand and solar generation from the known net demand.

First Finding: Spatial Correlation between Generations



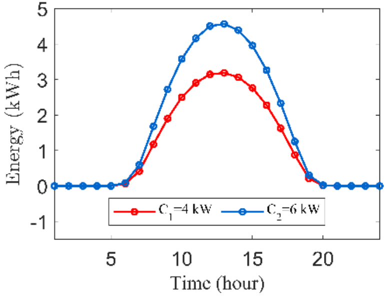


Exemplary solar generation curves

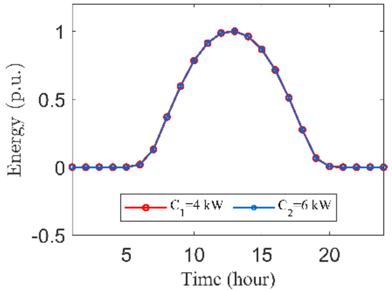
Primary factors: (1) Irradiance, (2) PV array capacity, (3) PV system loss, (4) PV array tilt, and (5) PV array azimuth

Conditioned on nearly identical irradiance input, could we estimate solar generation without using the complicated PV model and PV array parameters (generally unavailable)?

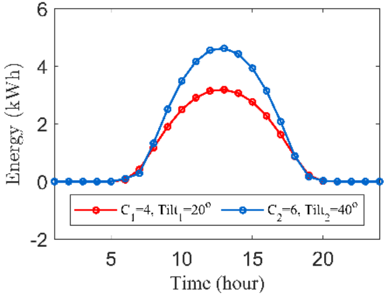
First Finding: Spatial Correlation between Generations



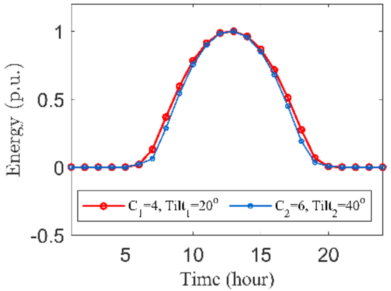
Normalizing



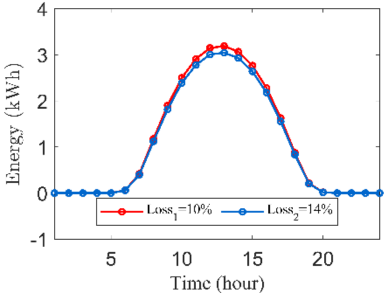
Different capacities



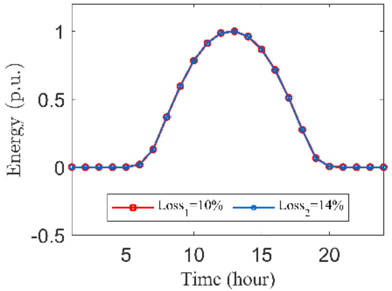
Normalizing



Different tilts



Normalizing



Different losses

Summary: (1) PVs with the same azimuth generates highly correlated power; (2) capacity, tilt and loss determine the magnitude of a solar generation curve.

First Finding: Spatial Correlation between Generations

- **An unknown solar curve can be represented using known solar power exemplars**

Nearly identical for all PVs and its effect can be inferred by known solar generation exemplars

Determine an unknown solar generation curve's magnitude which needs to be estimated

Can be captured by choosing solar generation exemplars from PVs with typical azimuths

Primary factors: (1) Irradiance, (2) PV array capacity, (3) PV system loss, (4) PV array tilt, and (5) PV array azimuth

Summary: Since the solar generation curves have a strong ***spatial correlation***, estimating the unknown PV generation curve comes down to estimating the ***weights*** assigned to solar generation exemplars.

Formulation:

Assume there are *N* solar generation exemplars:

$$\boldsymbol{G}_{h,i}^{E},\ i = 1, \dots, N.$$

The unknown weights to be estimated are denoted as:

$$\widehat{\boldsymbol{w}} = [\widehat{w}_1, \dots, \widehat{w}_N]^T.$$

The unknown solar generation series can be estimated as:

$$\widehat{\boldsymbol{G}}_h = \left[\boldsymbol{G}_{h,1}^{E}, \dots, \boldsymbol{G}_{h,N}^{E}\right] * \widehat{\boldsymbol{w}} = \mathbf{G}_h^{E} * \widehat{\boldsymbol{w}}.$$

The unknown native demand series can be expressed as:

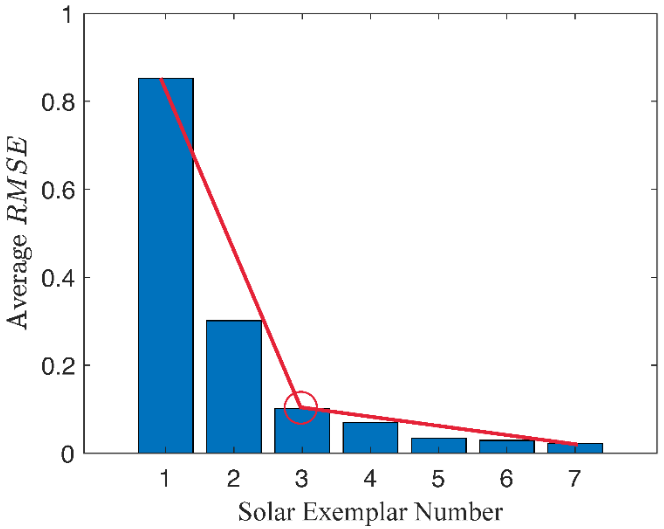
$$\widehat{\boldsymbol{P}}_h = \boldsymbol{P}_h' - \widehat{\boldsymbol{G}}_h,$$

where, $\boldsymbol{P}_h'$ denotes the measured net demand series.

How Many Solar Generation Exemplars are Needed?

Candidate Sets of PVs with Typical Azimuths

Number of PVs	Azimuths
1	$A_1 = 180°$
2	$A_1 = 135°, A_2 = 225°$
3	$A_1 = 90°, A_2 = 180°, A_3 = 270°$
4	$A_1 = 90°, A_2 = 150°, A_3 = 210°, A_4 = 270°$
5	$A_1 = 90°, A_2 = 135°, A_3 = 180°, A_4 = 225°, A_5 = 270°$
6	$A_1 = 90°, A_2 = 126°, A_3 = 162°, A_4 = 198°, A_5 = 234°, A_6 = 270°$
7	$A_1 = 90°, A_2 = 120°, A_3 = 150°, A_4 = 180°, A_5 = 210°, A_6 = 240°, A_7 = 270°$



(Step – I): Construct an explanatory matrix: $\mathbf{X}_k^e = [\boldsymbol{G}_1^e, \dots, \boldsymbol{G}_k^e]$, where $\boldsymbol{G}_i^e$ denotes the i'th solar generation exemplar in the k'th set, and $i = 1, \dots, k$. Please note that $\boldsymbol{G}_i^e$ is an 8760-by-1 vector.

(Step – II): For each PV generation curve to be estimated, $\boldsymbol{G}_j$, perform a linear regression estimation:

$$\widehat{\boldsymbol{w}}_j = \left({\mathbf{X}_k^e}^T \mathbf{X}_k^e\right)^{-1} {\mathbf{X}_k^e}^T \boldsymbol{G}_j, \quad j = 1, \dots, N_{PV}.$$

Compute the estimated PV generation vector:

$$\widehat{\boldsymbol{G}}_j = \mathbf{X}_k^e \widehat{\boldsymbol{w}}_j.$$

Compute the RMSE for each PV generation vector:

$$RMSE_{j,k} = \sqrt{\Sigma_{t=1}^{8760} \frac{\left(\widehat{\boldsymbol{G}}_j(t) - \boldsymbol{G}_j(t)\right)^2}{8760}}, \quad j = 1, \dots, N_{PV}.$$

(Step – III): Compute the average RMSE for each set of solar generation exemplars:

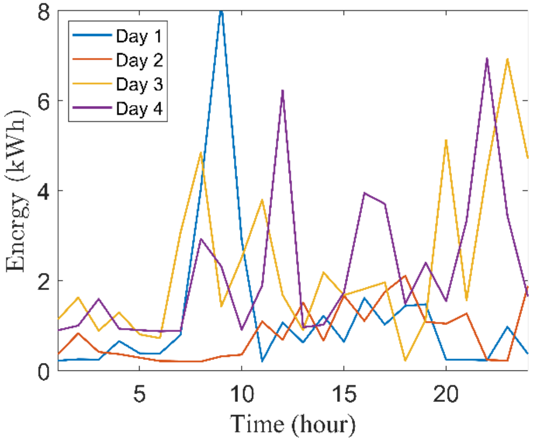
$$Average\ RMSE_k = \frac{1}{N_{PV}} \Sigma_{j=1}^{N_{PV}} RMSE_{j,k}$$

IEEE PES Power & Energy Society

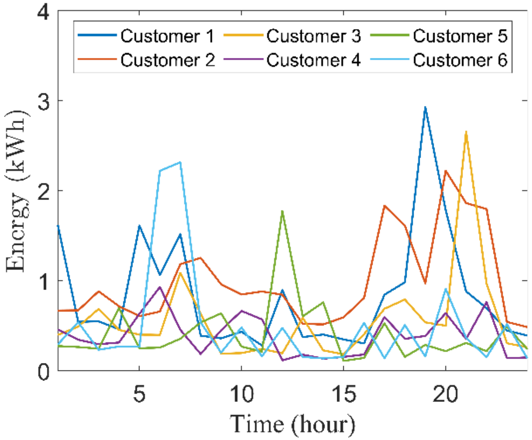
IEEE

Challenge: How to Estimate the Weights?

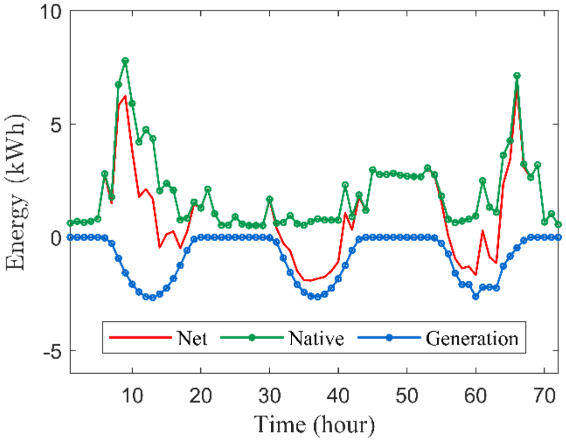
If the native demand follows a fixed pattern, we might estimate the weights by forcing the estimated native demand, $\widehat{P}_h$, to follow that pattern. However, the hourly native demand shows a high uncertainty, which brings challenges for BTM solar generation estimation.



Temporal uncertainty of the native demand

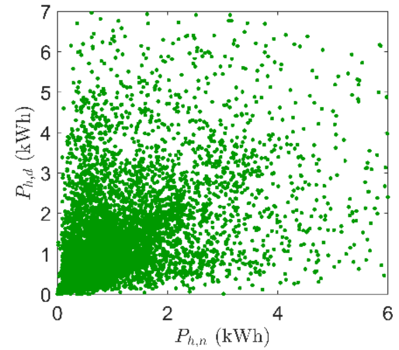


Spatial uncertainty of the native demand

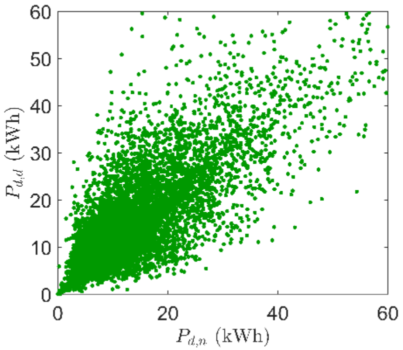


Distortion caused by uncertainty of the native demand

Second Finding: Correlation of Monthly Native Demand

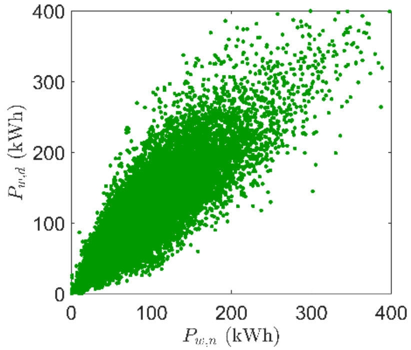


Customers w/o PV (Hourly)

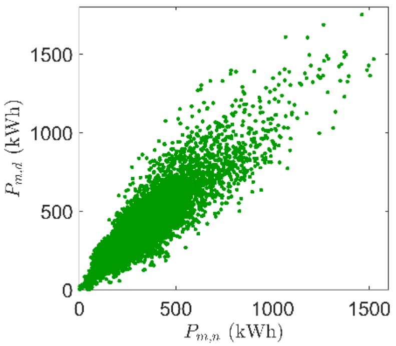


Customers w/o PV (Daily)

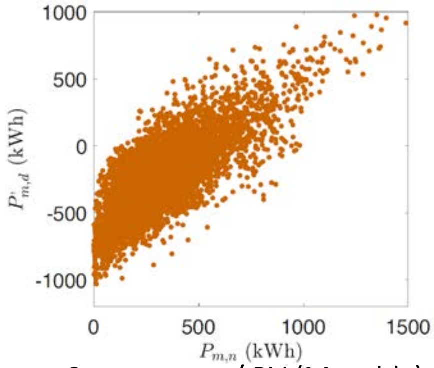
- As the timescale increases from one hour to one month, the diurnal native demand is highly correlated with the nocturnal native demand.
- The strong correlation on the timescale of one month inspires us to first estimate the monthly BTM generation of customers with PVs, and then decompose it into hourly solar power.



Customers w/o PV (Weekly)



Customers w/o PV (Monthly)



Customers w/ PV (Monthly)

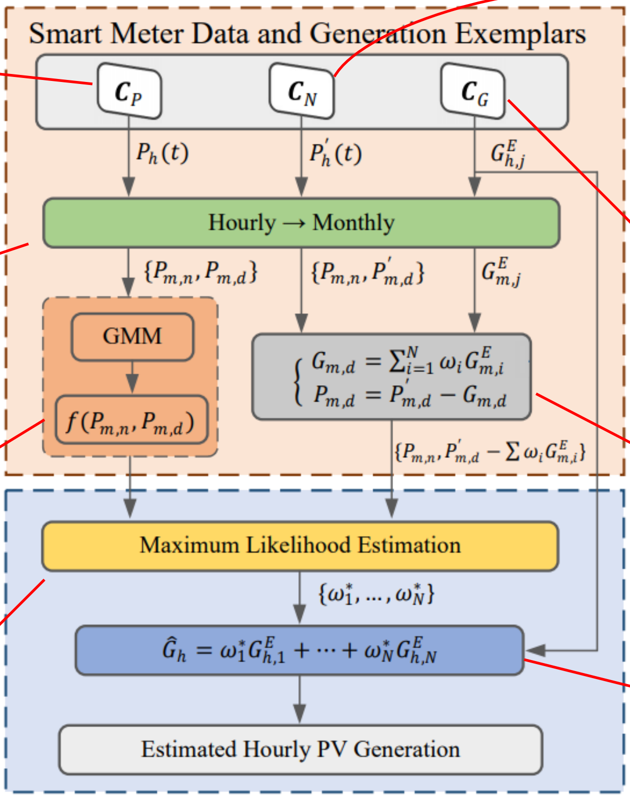
Structure of the Proposed Approach

C_P: The set of customers without PVs and whose native demand is recorded by smart meters.

$P_{m,n} = \sum_{t \in I_n} P_h(t)$, $P_{m,d} = \sum_{t \in I_d} P_h(t)$, $G^E_{m,j} = \sum_{t \in I_d} G^E_{h,j}(t)$: Converting hourly demand/generation to monthly nocturnal and diurnal demand/generation.

GMM is used to construct the joint distribution of known monthly nocturnal and diurnal native demands for representing their high correlation.

Using the maximum likelihood estimation (MLE) to optimize the weights assigned to solar generation exemplars.



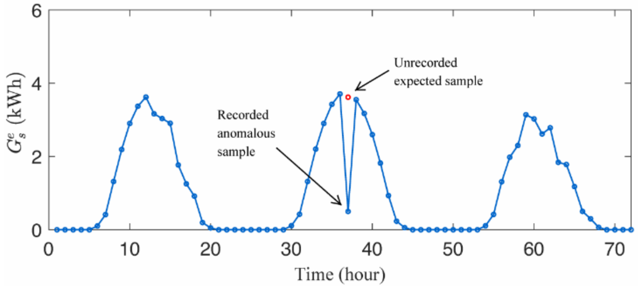
C_N: The set of customers with PVs and whose net demand is recorded by smart meter, while their native demand and PV generation are not separately measured and need to be ***estimated***.

C_G: The small group of customers with PVs and whose PV generation and native demand are both *observable* separately.

Representing the estimated monthly diurnal native demand using the known net demand and monthly exemplars.

Using the hourly solar generation exemplars and the optimized weights to estimate the unknown hourly solar generation.

Formulating An MLE Optimization



A solar exemplar with an anomalous sample.

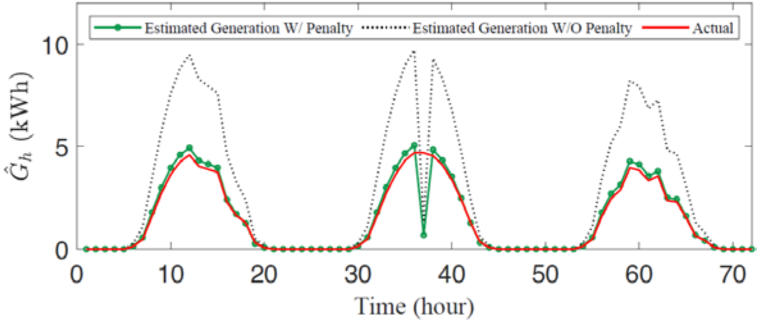
$$\boldsymbol{w}^* = \max_{\boldsymbol{w}} \left\{ \prod_{t=1}^{M} f\left(P_{m,n}(t), P'_{m,d}(t), \boldsymbol{G}_m^E(t) | \boldsymbol{w}\right) \right\} - \frac{1}{2}\lambda ||\boldsymbol{\beta}||_2^2$$

$$\text{s.t.} \quad \mathbf{G}_h^{E\,\mathrm{T}} * \boldsymbol{w} \geq \mathbf{0}$$

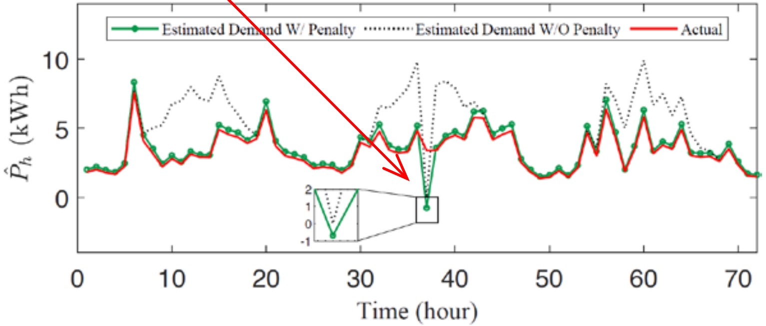
$$\boldsymbol{P}'_h + \mathbf{G}_h^{E\,\mathrm{T}} * \boldsymbol{w} \geq \boldsymbol{\beta}$$

$$\boldsymbol{\beta} \leq \mathbf{0}$$

A soft margin allows a limited number of negative native demand estimates and improves robustness



PV generation



Native demand

Conclusion

- A novel approach has been developed to disaggregate BTM PV generation and native demand from the net load for individual customers. The separated solar generation and native load can enhance distribution system observability, and thus to provide useful information for distribution system monitoring, operation, and planning, etc.
- The proposed approach does not require PV array parameters and weather data.
- The proposed approach does not rely on high-resolution data, and is designed to apply to the widely available hourly smart meter data.
- The proposed approach relies on a limited number of solar generation exemplars, which can be achieved by installing a couple of meters to record typical PV arrays.