Load Restoration in Islanded Microgrids: Formulation and Solution Strategies[1]

Shourya Bose and **Yu Zhang**
Department of Electrical and Computer Engineering
University of California, Santa Cruz

Funding ACK: Hellman Fellowship, CITRIS Seed Fund, Faculty Research Grant of UCSC

[1] in IEEE Transactions on Control of Network Systems, doi: 10.1109/TCNS.2023.3337710.

Outline

- Background
- System Description
- Model Predictive Control (MPC) Approach
- Constrained Policy Optimization (CPO) Approach
- Simulation Results
- Conclusion

Background

- Extreme weather events (e.g., wildfires, heat waves, tornadoes, and hurricanes) pose threats to the reliable operation of electric distribution systems (DS).
- Primarily by damaging electric equipment such as overhead cables, thereby curbing the distribution of electricity.
- Traditionally, DS are designed to be **reliable** during nominal operations and in face of predictable off-nominal operating conditions.
- Nowadays, DS must be **resilient**: rapid recovery to a state of nominal operation post extreme events.

Figure: Examples of extreme weather events: 2020 CA Wildfires, 2021 TX Power Crisis, 2021 Hurricane Ida

Microgrids

- A microgrid (MG) is a local energy grid with control capability, which can be in *grid-connected* or *islanding* mode depending on the situation of the utility grid.
- A MG is equipped with distributed energy resources including *microturbines (MTs) renewable energy systems (RESs)*, and *energy storage systems (ESSs)*.
- MG's control of generation and storage resources is available to the *MG controller (MGC)*.
- MG benefits: enhanced power quality, reliability and resilience, reduced electricity cost and carbon emissions, as well as revenue generation.

Figure: MT, RES, ESS, and loads.

Literature Review

- **Load restoration problem**:
  - Risk-limiting strategies (Wang et. al, 2019)
  - Wide-area monitoring systems (Liu et. al, 2013)
  - Distributed Systems (Nejad and Sun, 2019)
  - Expert Systems (Tsai, 2008)
- **Model predictive control (MPC)** is widely used in power systems:
  - Voltage stability assurance (Jin et. al, 2010)
  - Industrial demand response (Zhang et. al, 2016)
  - Volt/VAR control (Dhulipala et. al, 2019)
  - PV storage scheduling (Wang et. al, 2014)
- **Reinforcement learning (RL)** has also been widely used:
  - Volt/var control (Wang et. al, 2020)
  - EV charging scheduling (Li et. al, 2020)
  - Power management in networked MG (Zhang et. al, 2021)
  - Optimal control of ESSs in MGs (Duan et. al, 2019)

Mathematical Modeling of MG

- MG is represented by a *directed graph* $\mathcal{G} \triangleq (\mathcal{N}, \mathcal{E})$.
  - $i \rightarrow j \implies (i,j) \in \mathcal{E}$.
  - $\mathcal{N} = \mathcal{N}^{\mathsf{L}} \cup \mathcal{N}^{\mathsf{MT}} \cup \mathcal{N}^{\mathsf{ESS}} \cup \mathcal{N}^{\mathsf{RES}}$.
- The problem is solved over time horizon $\mathcal{T} \triangleq \{1, 2, \cdots, T\}$ with time index t.
- Nodal and branch variables at time t
  - Voltage magnitude squared, real and reactive power injections: $v_{i,t} \in \mathbb{R}_+, s_{i,t} \in \mathbb{C}$.
  - Line current magnitude squared, real and reactive power flows: $l_{ij,t} \in \mathbb{R}_+, S_{ij,t} \in \mathbb{C}$.
- Goal: Maximize load restored while minimize MT fuel cost using convex cost functions.

$$\boxed{J = \sum_{t \in \mathcal{T}} \left(\sum_{i \in \mathcal{N}^{\mathsf{L}}} C^{\mathsf{L}}_{i,t}(\mathrm{Re}\,(s_{i,t})) + \sum_{i \in \mathcal{N}^{\mathsf{MT}}} C^{\mathsf{MT}}(\mathrm{Re}\,(s_{i,t})) \right)}$$

Power Flow Constraints

AC DistFlow model

- Complex injections: $s_{j,t} = \underbrace{\sum_{k:j\to k} S_{ij,t}}_{\text{downstream flow}} - \underbrace{\sum_{i:i\to j} (S_{ij,t} - z_{ij} l_{ij,t})}_{\text{upstream flow}},\ \forall j \in \mathcal{N}$

- Voltage drop: $v_{j,t} - v_{i,t} = \underbrace{-2\mathrm{Re}\left(\bar{z}_{ij} S_{ij,t}\right)}_{\text{due to flows}} + \underbrace{|z_{ij}|^2 l_{ij,t}}_{\text{due to line impedance}},\ \forall (i,j) \in \mathcal{E}$

- Relation of line flows and currents: $l_{ij,t} v_{i,t} = |S_{ij,t}|^2,\ \forall (i,j) \in \mathcal{E}$

Voltage/Current Constraints & MT Constraints

Voltage & current limits

$$\underline{v} \leq v_{i,t} \leq \bar{v},\ \forall i \in \mathcal{N}, \forall t \in \mathcal{T}$$

$$l_{ij,t} \leq \bar{l}_{ij},\ \forall (i,j) \in \mathcal{E}, \forall t \in \mathcal{T}$$

MT power constraints

- Generation limits:

$$\underline{P}_i^{\text{MT}} \leq \text{Re}\left(s_{i,t}\right) \leq \bar{P}_i^{\text{MT}},\ \forall i \in \mathcal{N}^{\text{MT}}, \forall t \in \mathcal{T}$$

- Ramping up/down:

$$\underline{P}_{\text{rd}}^{\text{MT}} \leq \text{Re}\left(s_{i,t}\right) - \text{Re}\left(s_{i,t-1}\right) \leq \bar{P}_{\text{ru}}^{\text{MT}}, \forall i \in \mathcal{N}^{\text{MT}}, t \in \mathcal{T} \backslash \{1\}$$

$$\text{Re}\left(s_{i,t}\right) \leq \bar{P}_{\text{ru}}^{\text{MT}},\ \forall i \in \mathcal{N}^{\text{MT}}, t = 1$$

- Fuel constraint (define $\varsigma_{i,t}$ as the remaining fuel in MT i at time t):

$$\varsigma_{i,t} = \varsigma_{i,t-1} + \tau_i \text{Re}\left(s_{i,t}\right), \forall i \in \mathcal{N}^{\text{MT}}, \forall t \in \mathcal{T}$$

$$\varsigma_{i,t} \geq 0,\ \forall i \in \mathcal{N}^{\text{MT}}, \forall t \in \mathcal{T}$$

$$\varsigma_{i,0} = E_i,\ \forall i \in \mathcal{N}^{\text{MT}}$$

ESS Constraints

ESS charge/discharge power constraints

$$
\begin{aligned}
0 \leq P_{i,t}^{\text{dis}} \leq \bar{P}_i^{\text{dis}},\ \forall i \in \mathcal{N}^{\text{ESS}}, \forall t \in \mathcal{T} \\
0 \leq P_{i,t}^{\text{ch}} \leq \bar{P}_i^{\text{ch}},\ \forall i \in \mathcal{N}^{\text{ESS}}, \forall t \in \mathcal{T} \\
P_{i,t}^{\text{ch}} P_{i,t}^{\text{dis}} = 0,\ \forall i \in \mathcal{N}^{\text{ESS}}, \forall t \in \mathcal{T} \\
\mathrm{Re}\left(s_{i,t}\right) = P_{i,t}^{\text{dis}} - P_{i,t}^{\text{ch}},\ \forall i \in \mathcal{N}^{\text{ESS}}, \forall t \in \mathcal{T}
\end{aligned}
$$

ESS state-of-charge (SoC) evolution

$$
\begin{aligned}
\mathcal{S}_{i,t+1} = \mathcal{S}_{i,t} + \eta^{\text{ch}} P_{i,t}^{\text{ch}} \Delta t - \frac{1}{\eta^{\text{dis}}} P_{i,t}^{\text{dis}} \Delta t,\ \forall i \in \mathcal{N}^{\text{ESS}}, \forall t \in \mathcal{T} \backslash \{T\} \\
\mathcal{S}_{i,1} = \mathcal{S}_i^{\text{init}},\ \mathcal{S}_{i,t} \in [\underline{\mathcal{S}}, \bar{\mathcal{S}}],\ \forall i \in \mathcal{N}^{\text{ESS}}
\end{aligned}
$$

RES, Load, and Reactive power Constraints

RES power outputs (define $\kappa_{i,t}$ as curtailment ratio)

$$P_{i,t} = (1 - \kappa_{i,t}) \underbrace{\hat{P}^{\mathrm{RES}}_{i,t}}_{\text{RES forecast}}, \ \forall i \in \mathcal{N}^{\mathrm{RES}}, \ \forall t \in \mathcal{T}$$

Load constraint (define $\rho_{i,t}$ as pickup ratio and assume constant power factor)

$$\mathrm{Re}\,(s_{i,t}) = \rho_{i,t} \underbrace{\tilde{P}^{\mathrm{L}}_{i,t},}_{\text{real demand forecast}} \quad \mathrm{Im}\,(s_{i,t}) = \rho_{i,t} \underbrace{\tilde{Q}^{\mathrm{L}}_{i,t},}_{\text{reactive power forecast}} \quad \forall i \in \mathcal{N}^{\mathrm{L}}, \forall t \in \mathcal{T}$$

$$\underbrace{\rho_{i,t} - \rho_{i,t-1} \geq -\epsilon}_{\text{near-monotonic load restoration constraint}}, \ \forall i \in \mathcal{N}^{\mathrm{L}}, \forall t \in \mathcal{T} \backslash \{1\}$$

Droop constraint (droop buses $\mathcal{N}^{\mathrm{droop}} \subseteq \{\mathcal{N} \setminus \mathcal{N}^{\mathrm{L}}\}$)

$$\underbrace{\omega_t = \omega^*_i - k_P(\mathrm{Re}\,(s_{i,t}) - P^*_i),}_{\text{distributed global frequency control}} \quad \underbrace{\sqrt{v_{i,t}} = \sqrt{v^*_i} - k_Q(\mathrm{Im}\,(s_{i,t}) - Q^*_i),}_{\text{local droop voltage control}}$$

$$\forall i \in \mathcal{N}^{\mathrm{droop}}, \forall t \in \mathcal{T}$$

Reactive power constraint

$$|Q_{i,t}| \leq \bar{Q}_i, \ \forall i \in \mathcal{N} \setminus \mathcal{N}^{\mathrm{L}}, \ \forall t \in \mathcal{T}$$

Load Restoration Optimization Problem

- Define sets of the optimization variables

$$\mathcal{X}_t \triangleq \Big\{ \{s_{i,t}, v_{i,t}\}_{i\in\mathcal{N}}, \{S_{ij,t}, l_{ij,t}\}_{i\to j}, \{\zeta_{i,t}\}_{i\in\mathcal{N}^{\mathsf{MT}}},$$
$$\left\{\mathcal{S}_{i,t}, P^{\mathsf{ch}}_{i,t}, P^{\mathsf{dis}}_{i,t}\right\}_{i\in\mathcal{N}^{\mathsf{ESS}}}, \{\kappa_{i,t}\}_{i\in\mathcal{N}^{\mathsf{RES}}}, \{\rho_{i,t}\}_{i\in\mathcal{N}^{\mathsf{L}}}, \omega_t \Big\}.$$

- The **LR optimization problem** is given as

$$\max_{\{\mathcal{X}_t\}_{t\in\mathcal{T}}} \quad J$$
$$\text{subject to.} \quad \text{constraints,}$$
$$\left(\kappa_{i,t} \in [0,1], \forall i \in \mathcal{N}^{\mathsf{RES}};\ \rho_{i,t} \in [0,1], \forall i \in \mathcal{N}^{\mathsf{L}}\right), \forall t \in \mathcal{T}$$

- Problem may run into **curse of dimensionality** as $|\mathcal{T}|$ increases. We need to find an efficient method that yields suboptimal solutions.

Convex Relaxations of Nonconvex Constraints

- **Convex relaxation of droop voltage constraint** is nonconvex since it is a nonlinear equality. We propose a second-order cone relaxation of the same.

Lemma 1 (Convex relaxation of droop voltage equation)

The nonconvex drrop bus voltage constraint given as

$$\sqrt{v_{i,t}} = \sqrt{v_i^*} - k_Q(\mathrm{Im}\,(s_{i,t}) - Q_i^*)$$

admits a second-order cone relaxation given as

$$\left\| \left[\sqrt{v_i^*} - k_Q(\mathrm{Im}\,(s_{i,t}) - Q_i^*) \quad v_{i,t} \quad \tfrac{1}{2} \right] \right\|_2 \leq v_{i,t} + \frac{1}{2}.$$

- **Convex relaxation of DistFlow equations:** The non-convex DistFlow equation $l_{ij,t} v_{i,t} = |S_{ij,t}|^2$ can be relaxed to the well-known second-order cone $l_{ij,t} v_{i,t} \geq |S_{ij,t}|^2$.

Convex Relaxations of Nonconvex Constraints (contd.)

- Convex relaxation of ESS complementarity constraint: Replace by the convex hull of nonconvex feasibility region.

Lemma 2 (Convex hull of nonconvex ESS feasible region)

Define the set $\mathcal{P}_i$ as

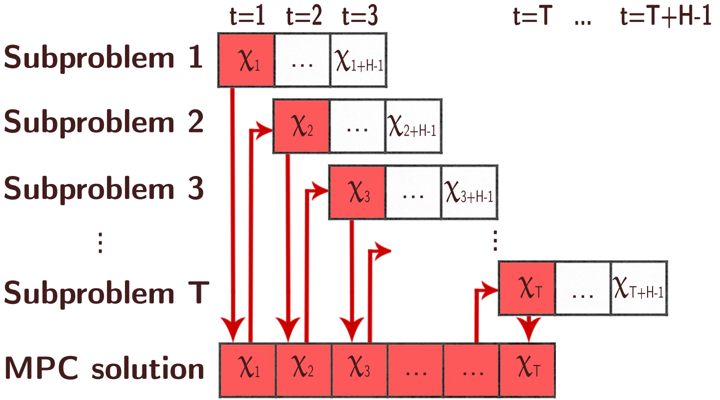
$$\mathcal{P}_i \triangleq \{(P^{\text{ch}}, P^{\text{dis}}) \,|\, P^{\text{ch}}, P^{\text{dis}} \text{ satisfy bounds and complementarity constraints}\}.$$

Then $\mathcal{P}_i^{\text{conv}} \triangleq \text{conv}(\mathcal{P}_i)$ is given as

$$\mathcal{P}_i^{\text{conv}} = \left\{ (P^{\text{ch}}, P^{\text{dis}}) \,\middle|\, P^{\text{ch}} \geq 0, P^{\text{dis}} \geq 0, \frac{P^{\text{ch}}}{\bar{P}_i^{\text{ch}}} + \frac{P^{\text{dis}}}{\bar{P}_i^{\text{dis}}} \leq 1 \right\}.$$

Model Predictive Control (MPC) Optimization Problem

The convexified problem may run into **curse of dimensionality** as T becomes large. Thus, we use MPC to solve over sub-horizons of size H, and implement only the solution of the first time step.



Constrained Reinforcement Learning

- Reinforcement learning (RL): decide the best sequence of **actions** for an agent interacting with its **environment**. The choice of actions is made by maximizing the **reward**.

- Constrained RL leverages *Constrained Markov Decision Process* (CMDP) to describe the underlying system. For this work, we adopt the **Constrained Policy Optimization** (CPO) proposed by Achiam et al.

- A CMDP is defined as the 6-tuple $\{\mathbb{X}, \mathbb{U}, p, R, \mathbf{C}, \gamma\}$, where:
  - $\mathbb{X}$ is the state space
  - $\mathbb{U}$ is the action space
  - $p : \mathbb{X} \times \mathbb{U} \times \mathbb{X} \mapsto [0, 1]$ is the state transition probability
  - $R : \mathbb{X} \times \mathbb{U} \mapsto \mathbb{R}$ is the reward function
  - $\mathbf{C} : \mathbb{X} \times \mathbb{U} \mapsto \mathbb{R}^M$ is the constraint function (assume the CMDP actions have M constraints)
  - $\gamma \in (0, 1]$ is the discount factor

Choice of State, Action, and Reward

- For the load restoration problem, **state** $\mathbf{x}_t$ is defined as

$$\mathbf{x}_t \triangleq \Big\{ \{\mathcal{S}_{i,t-1}\}_{i\in\mathcal{N}^{\mathsf{ESS}}}, \{\varsigma_{i,t-1}\}_{i\in\mathcal{N}^{\mathsf{MT}}}, \{\rho_{i,t-1}\}_{i\in\mathcal{N}^{\mathsf{L}}}, \Big[\Big\{ \hat{P}^{\mathsf{L}}_{i,t'}, \hat{Q}^{\mathsf{L}}_{i,t'} \Big\}_{i\in\mathcal{N}^{\mathsf{L}}}, \Big\{ \hat{P}^{\mathsf{RES}}_{i,t'} \Big\}_{i\in\mathcal{N}^{\mathsf{RES}}} \Big]_{t'=t}^{t+H-1} \Big\}.$$

- **Action** $\mathbf{u}_t$ is defined as

$$\mathbf{u}_t \triangleq \Big\{ \Big[\{s_{i,t'}\}_{i\in\mathcal{N}^{\mathsf{RES}}}, \{s_{i,t'}\}_{i\in\mathcal{N}^{\mathsf{MT}}}, \{\rho_{i,t'}\}_{i\in\mathcal{N}^{\mathsf{L}}} \Big\{ P^{\mathsf{ch}}_{i,t'}, P^{\mathsf{dis}}_{i,t'}, \mathrm{Im}\,(s_{i,t'}) \Big\}_{i\in\mathcal{N}^{\mathsf{ESS}}} \Big]_{t'=t}^{t+H-1} \Big\}.$$

- The **reward** at time step t is given as:

$$R(\mathbf{x}_t, \mathbf{u}_t) = \sum_{t'=t}^{t+H} \gamma^{(t-t')} \Big[\sum_{i\in\mathcal{N}^{\mathsf{L}}} C^{\mathsf{L}}_{i,t'}(\mathrm{Re}\,(s_{i,t'})) + \sum_{i\in\mathcal{N}^{\mathsf{MT}}} C^{\mathsf{MT}}_{i,t'}(\mathrm{Re}\,(s_{i,t'})) \Big].$$

Policy

- A **policy** $\pi_{\boldsymbol{\theta}} : \mathbb{X} \times \mathbb{U} \mapsto [0, 1]$ parameterized by $\boldsymbol{\theta} \in \mathbb{R}^h$ and denoted as $\pi_{\boldsymbol{\theta}}(\mathbf{u}|\mathbf{x})$ gives the probability of taking action **u** on current state **x**.
- We model the policy as a **multivariate Gaussian distribution** whose mean vector and covariance matrix are generated by a **feedforward neural network** (FNN). Let d denote the dimension of the action variable,

$$\pi_{\boldsymbol{\theta}}(\mathbf{u}|\mathbf{x}) = \frac{1}{\sqrt{|\boldsymbol{\Sigma}_{\mathbf{x}}|(2\pi)^d}} e^{-\frac{1}{2}(\mathbf{u}-\boldsymbol{\mu}_{\mathbf{x}})^\top \boldsymbol{\Sigma}_{\mathbf{x}}^{-1}(\mathbf{u}-\boldsymbol{\mu}_{\mathbf{x}})},$$

  where $\boldsymbol{\mu}_x \in \mathbb{R}^d$ and $\boldsymbol{\Sigma}_x \in \mathbb{R}^{d\times d}$ are generated by using FNN $f_\theta(\mathbf{x})$ as

$$\boldsymbol{\mu}_{\mathbf{x}} = L_{\boldsymbol{\mu}}(f_{\boldsymbol{\theta}}(\mathbf{x})), \quad \boldsymbol{\Sigma}_{\mathbf{x}} = M(L_{\boldsymbol{\Sigma}}(f_{\boldsymbol{\theta}}(\mathbf{x}))),$$

  wherein the function $M(\mathbf{A}) = \mathbf{A}\mathbf{A}^\top$ converts the outputs of the FNN into a positive semidefinite covariance matrix.
- The vector $\boldsymbol{\theta}$ contains the **weights** and **biases** which parameterize the FNNs.

CPO Formulation

- Let vectorized notation $\mathbf{C}(\mathbf{x}_t, \mathbf{u}_t) \leq \mathbf{0}$ denote the constraints on the variables in action $\mathbf{u}_t$. Define the reward and constraint function at time t as

$$J^R(\pi_{\boldsymbol{\theta}}, \mathbf{x}_t) \triangleq \mathbb{E}_{\mathbf{u}_t \sim \pi_{\boldsymbol{\theta}}} \left[R(\mathbf{x}_t, \mathbf{u}_t) \,|\, \mathbf{x}_t \right]$$
$$J^{\mathbf{C}}(\pi_{\boldsymbol{\theta}}, \mathbf{x}_t) \triangleq \mathbb{E}_{\mathbf{u}_t \sim \pi_{\boldsymbol{\theta}}} \left[\mathbf{C}(\mathbf{x}_t, \mathbf{u}_t) \,|\, \mathbf{x}_t \right]$$

- The CPO approach solves the problem by determining a policy $\boldsymbol{\theta}^*$ such that for any state s, $\pi_{\boldsymbol{\theta}^*}$ maximizes $J^R(\pi_{\boldsymbol{\theta}^*}, s)$ while respecting $J^{\mathbf{C}}(\pi_{\boldsymbol{\theta}^*}, s) \leq \mathbf{0}$.
- Such a policy $\boldsymbol{\theta}^*$ can be found in an *episodic fashion* by determining a sequence of policies $\{\boldsymbol{\theta}_t\}$ such that

$$\begin{aligned} \boldsymbol{\theta}_{t+1} = \arg\max_{\boldsymbol{\theta}} \; & J^R(\pi_{\boldsymbol{\theta}}, \mathbf{x}_t) \\ \text{s.t. } & J^{\mathbf{C}}(\pi_{\boldsymbol{\theta}}, \mathbf{x}_t) \leq \mathbf{0} \\ & D_{\mathrm{KL}}\left(\pi_{\boldsymbol{\theta}}(\,.\,|\mathbf{x}_t) \| \pi_{\boldsymbol{\theta}_t}(\,.\,|\mathbf{x}_t)\right) \leq \delta \end{aligned}$$

for some trust region parameter $\delta > 0$, the last constraint ensures that successive policies do not have large variations.

Approximating the CPO Problem

- Directly solving the CPO problem can be challenging due to the highly **nonlinear** and **nonconvex** nature of the FNNs.

- Instead, as proposed by Achiam et al, we solve the following convex nonlinear program to approximate the original CPO

$$\begin{aligned} \boldsymbol{\theta}_{t+1} = \arg\max_{\boldsymbol{\theta}} \; & \mathbf{a}_t^\top (\boldsymbol{\theta} - \boldsymbol{\theta}_t) \\ \text{s.t. } & \mathbf{B}_t^\top (\boldsymbol{\theta} - \boldsymbol{\theta}_t) + \mathbf{c}_t \le \mathbf{0} \\ & (\boldsymbol{\theta} - \boldsymbol{\theta}_t)^\top \mathbf{F}_t (\boldsymbol{\theta} - \boldsymbol{\theta}_t) \le \delta, \end{aligned}$$

where $\mathbf{F}_t$ is a positive (semi)definite matrix.

Closed form of Parameters

Theorem 1 (Parameters in the convex approximation)

The local second-order cone program approximation of the CPO problem contains parameters which may be computed as follows

$$\mathbf{a}_t = \left(\mathbb{E}_{\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d^{\mathrm{id}})} \left[\frac{\partial R_t}{\partial \mathbf{u}_t} \left((\boldsymbol{\epsilon}^\top \otimes \mathbf{I}_d^{\mathrm{id}}) \frac{\partial \mathbf{v}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{\mu}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}} \right) \middle| \mathbf{x}_t \right] \right)_{\boldsymbol{\theta} = \boldsymbol{\theta}_t}$$

$$\mathbf{B}_t = \left(\mathbb{E}_{\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d^{\mathrm{id}})} \left[\frac{\partial \mathbf{C}_t}{\partial \mathbf{u}_t} \left((\boldsymbol{\epsilon}^\top \otimes \mathbf{I}_d^{\mathrm{id}}) \frac{\partial \mathbf{v}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{\mu}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}} \right) \middle| \mathbf{x}_t \right] \right)_{\boldsymbol{\theta} = \boldsymbol{\theta}_t}$$

$$\mathbf{c}_t = J^{\mathbf{C}}(\pi_{\boldsymbol{\theta}_t}, \mathbf{x}_t)$$

$$\mathbf{F}_t(i,j) = \left[\frac{\partial \boldsymbol{\mu}_{\boldsymbol{\theta}}^\top}{\partial \boldsymbol{\theta}(i)} \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1} \frac{\partial \boldsymbol{\mu}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}(j)} + \frac{1}{2} \operatorname{Tr} \left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}(i)} \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}(j)} \right) \right]_{\boldsymbol{\theta} = \boldsymbol{\theta}_t}$$

where $R_t \triangleq R(\mathbf{x}_t, \mathbf{u}_t)$, $\mathbf{C}_t \triangleq \mathbf{C}(\mathbf{x}_t, \mathbf{u}_t)$, $\mathbf{v}_{\boldsymbol{\theta}} \triangleq \mathrm{vec}(\boldsymbol{\Sigma}_{\boldsymbol{\theta}})$, and $\mathbf{F}_t(i,j)$ and $\boldsymbol{\theta}(i)$ are the $(i,j)^{\mathrm{th}}$ and i^{th} element of $\mathbf{F}_t$ and $\boldsymbol{\theta}$, respectively. Furthermore, $\mathbf{I}^{\mathrm{id}}{}_d$ is the identity matrix of size $d \times d$.

Computing $\frac{\partial R_t}{\partial \mathbf{u}_t}$

- Partial derivatives of MT real power terms are easy to compute since

$$\frac{\partial P_{i,t}}{\partial P} = \begin{cases} 1, & \text{if } P \text{ is } P_{i,t} \\ 0, & \text{otherwise,} \end{cases} \quad \forall i \in \mathcal{N}^{\mathrm{MT}}.$$

- For the partial derivatives of load real power terms with respect to the action variables, we have to work across their implicit relations.

Lemma 2

The partial derivative $\frac{\partial P_{i,t}}{\partial P}$ for $P \in \mathbf{u}_t$ and $i \in \mathcal{N}^{\mathrm{L}}$ can be computed by setting $dP' = 0$ for all $P' \in \mathbf{u}_t$ such that $P' \neq P$ in the total derivative equations, and solving the resulting homogeneous linear system. A solution always exists. Then, $\frac{\partial P_{i,t}}{\partial P} = \frac{dP_{i,t}}{dP}$.

Proof by the *implicit function theorem*.

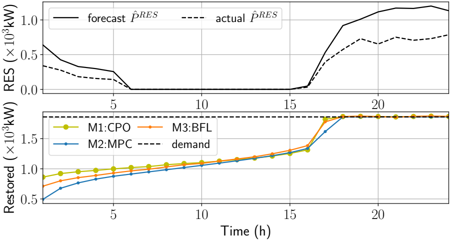
Simulation Setup (1/2)

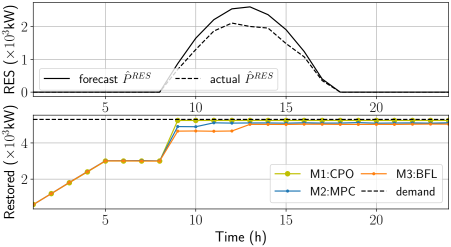
- All simulations carried out on 36- and 141- bus MGs adapted from MATPOWER . Former has (3 MT, 2 ESS, 2 ESS, 29 Load, 1 Droop) buses while latter has (7 MT, 6 ESS, 3 RES, 125 Load, 2 Droop) buses.
- We compare three methods: **M1:** CPO policy trained for 1000- and 3000- episodes for 36- and 141- bus systems respectively; **M2:** MPC with convex relaxations, **M3:** Learning MPC solutions with a FNN. All methods use lookahead $H = 5$.
- CPO parameters are as follows.

CPO Parameter	**36-bus MG**	**141-bus MG**
FNN f_θ layer sizes	$(344, 100, 75, 440)$	$(1403, 500, 250, 1630)$
Activation Function	tanh	tanh
Trust region δ	0.001	0.008
Batch size B	32	64
Loosening factor ξ	3	10

Simulation Results - Load Restoration Performance

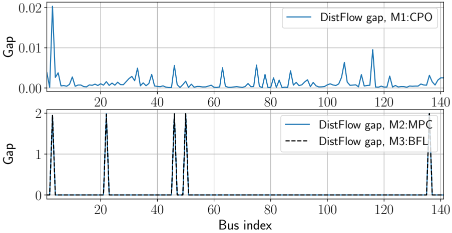
CPO performs better under uncertain RES forecasts. Results for 36- and 141- bus MGs.

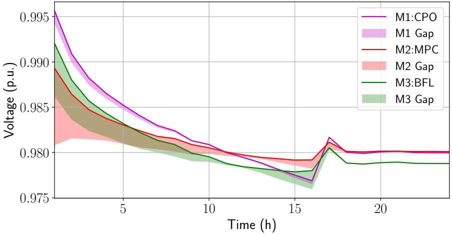




Simulation Results - Model Faithfulness

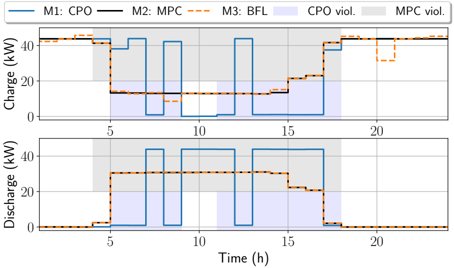
CPO solutions are more faithful to underlying model - both for DistFlow equations as well as voltage droop constraint. Results for 141- and 36- bus MGs.





Simulation Results - ESS Complementarity & Solve Times

CPO solutions are more faithful to ESS complementarity (results for 36- bus MG). Furthermore, they enjoy very small runtimes, although at the cost of extended offline training times.



Method	36-bus MG time		141-bus MG time	
	Training	Runtime	Training	Runtime
CPO	4310s	0.51s	33780s	0.76s
MPC	-	2.76s	-	7.59s
BFL	345s	0.38s	1215s	0.62s

Key Takeaways

- Due to increasing occurances of extreme weather events which lead to disruptions of DS operations, MGs have emerged as a potential candidate for increasing resilience of the DS. Thus, research into algorithms for automatic LR in MGs is timely.
- In this work, we formulated the load restoration problem for islanded MG with loads, ESSs, RESs, and MTs, and considered a relaxation which can be efficiently solved with MPC.
- We also considered a RL method called CPO for solving the LR problem, and compared its performance with MPC. Better LR performance was observed with CPO than MPC during simulations.
- Future work involves coordinated local restoration of multiple MGs, and consideration of privacy concerns that arise as a consequence of cooperation.

Thank You!