Detailed derivation of the submodular inequality

If $E(z)$ is submodular in z, z is binary variables, define a set $V := \{1,2,...,p\}$, with an indicator set of z as $M(z) := \{1 \le m \le p : z_m = 1\}$, and a set function $o(M(z)) = E(z)$ for all z. Then,

$$o(M \cup \{l\}) - o(M) \le o(N \cup \{l\}) - o(N), \quad N \subseteq M \subseteq V, l \in V \setminus M, \tag{1}$$

holds [1].

In Section IV-B, we establish the following definition: n is the dimension of $\boldsymbol{x}$, L is the set $\{1,2,\ldots,2n\}$, $S(\tilde{\boldsymbol{x}})$ is the indicator set $\{1 \le k \le 2n$: $\tilde{x}_k = 1\}$ for $\tilde{\boldsymbol{x}}$, and the set function $\varphi(\cdot)$ is defined such that $\varphi(S(\tilde{\boldsymbol{x}})) = \tilde{\Phi}(\tilde{\boldsymbol{x}})$ for all $\tilde{\boldsymbol{x}}$, marginal gain $\rho(S,k) := \varphi(S \cup \{k\}) - \varphi(S)$ for all $S \subseteq L$ and $k \in L \setminus S$. Moreover, according to Section 4.2, the lower-level value function $\tilde{\Phi}(\tilde{\boldsymbol{x}})$ is submodular in $\tilde{\boldsymbol{x}}$. Therefore,

$$\rho(S,k) = \varphi(S \cup \{k\}) - \varphi(S) \le \varphi(N \cup \{k\}) - \varphi(N) = \rho(N,k), \quad N \subseteq S, \tag{2}$$

holds.

For any subset $S \cup A \setminus D$ of the set L, where $D = \{d_1, d_2, ..., d_m\} \subseteq S, A = \{a_1, a_2, ..., a_n\} \subseteq L \setminus S$, the inequality

$$\varphi(S \cup A \setminus D) \le \varphi(S) - \sum_{d_i \in D} \rho(L \setminus \{d_i\}, d_i) + \sum_{a_i \in A} \rho(S, a_i), \tag{3}$$

holds. The proof is as follows:

$$\begin{aligned}
\varphi(S \cup A) - \varphi(S) &= \varphi(S \cup \{a_2, a_3, ..., a_n\} \cup \{a_1\}) - \varphi(S \cup \{a_2, a_3, ..., a_n\}) + \varphi(S \cup \{a_3, a_4, ..., a_n\} \cup \{a_2\}) \\
&\quad - \varphi(S \cup \{a_3, a_4, ..., a_n\}) + ... + \varphi(S \cup \{a_n\}) - f(S) \\
&= \rho(S \cup \{a_2, a_3, ..., a_n\}, a_1) + \rho(S \cup \{a_3, a_4, ..., a_n\}, a_2) + ... + \rho(S, a_n) \\
&\le \rho(S, a_1) + \rho(S, a_2) + ... + \rho(S, a_n) = \sum_{a_i \in A} \rho(S, a_i)
\end{aligned}, \tag{4}$$

$$\begin{aligned}
\varphi(S \cup A) - \varphi(S \cup A \setminus D) &= \varphi(S \cup A) - \varphi(S \cup A \setminus \{d_1\}) + \varphi(S \cup A \setminus \{d_1\}) - \varphi(S \cup A \setminus \{d_1, d_2\}) \\
&\quad + ... + \varphi(S \cup A \setminus \{d_1, d_2, ..., d_{m-1}\}) - \varphi(S \cup A \setminus D) \\
&= \rho(S \cup A \setminus \{d_1\}, d_1) + \rho(S \cup A \setminus \{d_1, d_2\}, d_2) + ... + \rho(S \cup A \setminus D, d_m) \\
&\ge \rho(L \setminus \{d_1\}, d_1) + \rho(L \setminus \{d_2\}, d_2) + ... + \rho(L \setminus \{d_m\}, d_m) \\
&= \sum_{d_i \in D} \rho(L \setminus \{d_i\}, d_i)
\end{aligned}. \tag{5}$$

Combining (4) and (5), we obtain (3), thereby completing the proof.

Based on the relationship between the indicator set and $\tilde{\boldsymbol{x}}$, Eq. (3) can be written as follows:

$$\tilde{\Phi}(\tilde{\boldsymbol{x}}) \le \varphi(S) - \sum_{k \in S} \rho(L \setminus \{k\}, k)(1 - \tilde{x}_k) + \sum_{k \in L \setminus S} \rho(S, k)\tilde{x}_k. \tag{6}$$

Owing to the high-precision approximation of the lower-level value function by the ISNN, it can be assumed that $\tilde{\Phi}(\tilde{\boldsymbol{x}}) = \Phi(\boldsymbol{x})$. Combining this with Eq. (10d), we obtain the following:

$$f(\boldsymbol{x}, \boldsymbol{y}) \le \varphi(S) - \sum_{k \in S} \rho(L \setminus \{k\}, k)(1 - \tilde{x}_k) + \sum_{k \in L \setminus S} \rho(S, k)\tilde{x}_k. \tag{7}$$

[1] Nemhauser G L, Wolsey L A. Maximizing submodular set functions: formulations and analysis of algorithms[J], North-Holland Mathematics Studies, 1981, 59: 279-301.

Pseudocode of the data-driven BMINLP algorithm

Data-driven BMINLP algorithm

Input sample size N_s, initial lower bound $\hat{F}^{\mathrm{lb}}$

Set $k_s \leftarrow 0, \Omega_{\mathrm{sample}} \leftarrow \varnothing, C \leftarrow \varnothing, I_{\mathrm{RSC}} \leftarrow \varnothing, \mathrm{flag}_1 \leftarrow 0, \mathrm{flag}_2 \leftarrow 0, i \leftarrow 1, j \leftarrow 1$

Sample integer variables in the upper level and train ISNN

while $k_s \leq N_s$ **do**

  Randomly generate positive semi-definite matrix Q and vector h.

  Solve the sampling model in Eq. (16) and store the sampled solution $\hat{\boldsymbol{x}}$.

  if $\hat{\boldsymbol{x}} \notin \Omega_{\mathrm{sample}}$ **then**

    Fix $\boldsymbol{x}$ as $\hat{\boldsymbol{x}}$ in the lower-level problem (15), solve it, and store the value function $\hat{\Phi}(\hat{\boldsymbol{x}})$ and optimal solution $\hat{\boldsymbol{y}}$.

    Expand $\Omega_{\mathrm{sample}} \leftarrow \Omega_{\mathrm{sample}} \cup \{\hat{\boldsymbol{x}}\}$ and Update $k_s \leftarrow k_s + 1$.

  end if

  if $F(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) > \hat{F}^{\mathrm{lb}}$ **then**

    Update $\hat{F}^{\mathrm{lb}} \leftarrow F(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$.

  end if

end while

Train ISNN using $(\tilde{\boldsymbol{x}}, -\hat{\Phi})$ for all $\hat{\boldsymbol{x}} \in \Omega_{\mathrm{sample}}$.

Solve MISOCP model (19) by the AHCP algorithm

Solve the relaxed model (23) and store solution $\boldsymbol{x}^{(1)}, \boldsymbol{y}^{(1)}$.

while $\mathrm{flag}_1 = 0$ **do**

  if $\mathrm{flag}_2 = 0$ **do**

    Compute Φ_i^{RSC} according to (22b) and expand $C \leftarrow C \cup \{\Phi_i^{\mathrm{RSC}}\}$.

    Expand $I_{\mathrm{RSC}} \leftarrow I_{\mathrm{RSC}} \cup \{i\}$ and update $i \leftarrow i+1$.

    Solve the approximate model (24) and store solution $\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}$.

  else

    Compute Φ_i^{SC} according to (21b) and expand $C \leftarrow C \cup \{\Phi_i^{\mathrm{SC}}\}$.

    Update $i \leftarrow i+1$.

    Solve the approximate model (25) and store solution $\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}$.

  end if

  if $f(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}) \leq \tilde{\Phi}_i(\tilde{\boldsymbol{x}})$ **do**

    if $\mathrm{size}(I_{\mathrm{RSC}}) = 0$ **then**

      Update $\mathrm{flag}_1 \leftarrow 1$.

      break;

    else if $\mathrm{size}(I_{\mathrm{RSC}}) < 2$ **then**

      Update $\mathrm{flag}_2 \leftarrow 1$.

    end if

    for j in I_{RSC} **then**

      Compute Φ_j^{SC} according to Eq. (21b) and do the set operation $C \leftarrow C \setminus \{\Phi_j^{\mathrm{RSC}}(\tilde{\boldsymbol{x}})\} \cup \{\Phi_j^{\mathrm{SC}}(\tilde{\boldsymbol{x}})\}$.

    end for

    do the set operation $I_{\mathrm{RSC}} \leftarrow \varnothing$.

  end if

end while

Correct the solution from MISOCP model

Fix $\boldsymbol{x}$ at $\boldsymbol{x}^*$ then get the solution $\boldsymbol{y}^*$ from the lower-level problem (15).

Calculate $F(\boldsymbol{x}^*, \boldsymbol{y}^*)$ and $f(\boldsymbol{x}^*, \boldsymbol{y}^*)$.

Output: $\boldsymbol{x}^*, \boldsymbol{y}^*, F(\boldsymbol{x}^*, \boldsymbol{y}^*)$ and $f(\boldsymbol{x}^*, \boldsymbol{y}^*)$

Details of the real coupled transportation-power network

The real coupled transportation-power network made up of two DNs and a transportation network. The topologies of the transportation network and the two distribution networks are illustrated in Figs. 1 and 2, respectively.

DN 1 comprises 76 buses, configured with 10 PVs and one CB, with its CS being located at bus 26. DN 2 comprises 119 buses, configured with four PVs and two CBs, with its CS at bus 31. The load curves of the two DNs are depicted in Fig. 3. The substation transformers with an OLTC are installed at the gateway bus of two DNs, respectively, and a tap ratio range is 1 ± 5 × 2%. The maximum capacity of the CBs in two DNs is 5 × 0.2 Mvar. The maximum power output of PVs in DN 1 and DN 2 is 0.8MW and 0.55 MW, respectively. The power factor of PVs is set as 0.90. The forecast output power curves of PVs in two DNs are illustrated in Fig. 4.

Fig. 1. Topology of the practical transportation network.

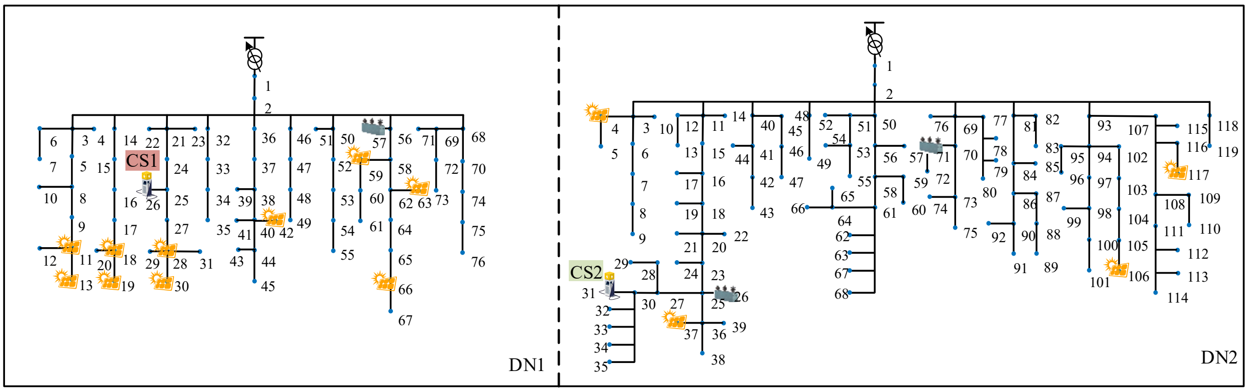
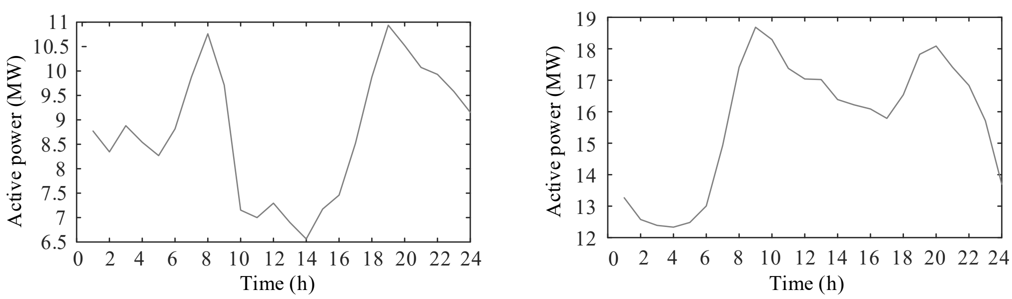
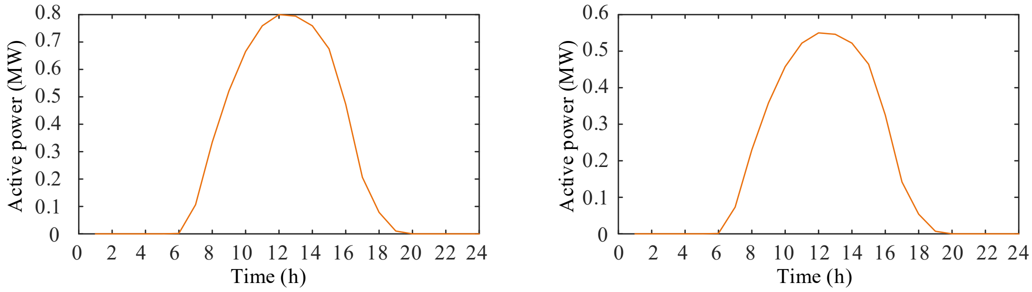


Fig. 2. Topology of two DNs.



(a) DN1 (b) DN2

Fig. 3. Daily load curves of two DNs.



(a) DN1 (b) DN2

Fig. 4. Forecast output power curves of PVs in two DNs.