Advances in Collaborative Neurodynamic Optimization

Jun Wang

Department of Computer Science & School of Data Science
City University of Hong Kong
Kowloon Tong, Kowloon, Hong Kong
jwang.cs@cityu.edu.hk

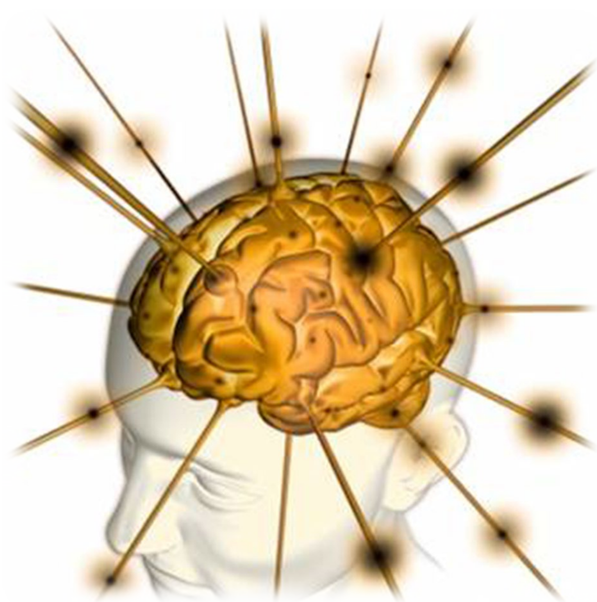
TU-Wien, Vienna, Austria; June 10, 2024

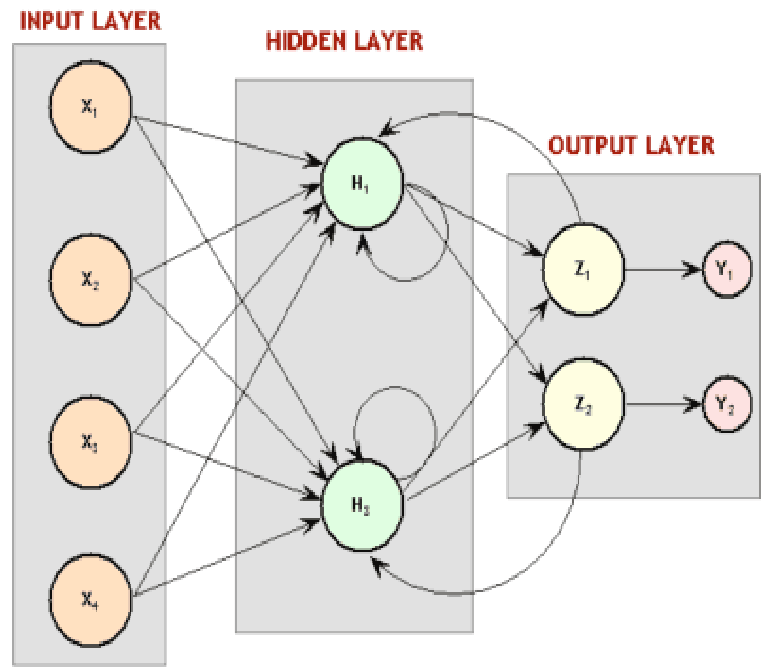
Optimization

- Optimization is omnipresent in nature and society.
- Optimization is an important tool for problem-solving.
- Optimization problems arise in numerous applications such as data processing, machine learning, robotics and control, etc.

Neurodynamic Optimization

As brain-like nonlinear dynamic systems, recurrent neural networks can serve as parallel computational models for optimization (aka. neurodynamic optimization).



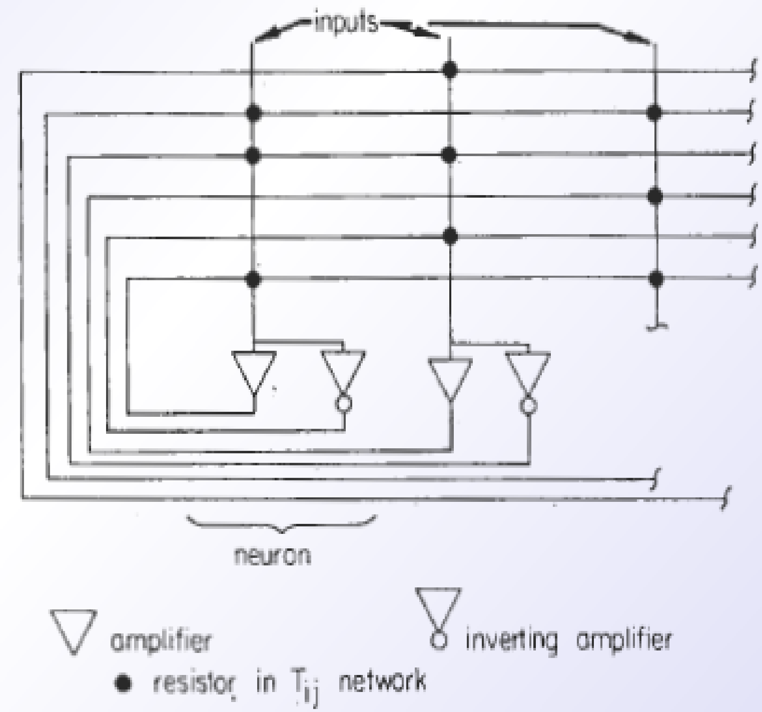


Pioneering Works

1. J. J. Hopfield and D. W. Tank, *Biological Cybernetics*, vol. 52, pp. 141–152, 1985.
2. J. J. Hopfield and D. W. Tank, *Science*, vol. 233. no. 4764, pp. 625–633,1986.
3. D.W. Tank and J. J. Hopfield, *IEEE Trans. Circuits and Systems*, vol. 33, pp. 533–541, 1986.

$$C_i \frac{du_i}{dt} = -\frac{u_i}{R_i} + \sum_j W_{ij} v_j + I_i$$

$$v_i = g_i(u_i)$$



Problem Formulation

Consider the following constrained optimization problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) \le 0 \\ & h(x) = 0 \\ & l \le x \le u \end{aligned} \qquad (1)$$

where $x \in \Re^n$, $f : \Re^n \to \Re$, $g(x) = [g_i(x), \ldots, g_m(x)]^T$, $h(x) = [h_i(x), \ldots, h_q(x)]^T$, l is a lower bound and u is an upper bound where $-\infty < l \le u < +\infty$. Let $\Omega = [l, u] = \prod_{i=1}^{n} [l_i, u_i]$. $f(x)$, $g(x)$ and $h(x)$ are assumed to be twice differentiable. If $f(x)$ or $g(x)$ is nonconvex or $h(x)$ is nonaffine, then (1) is a global optimization problem.

Solvable Problems

- Linear programming
- Convex optimization; e.g., convex quadratic programming
- Nonsmooth optimization
- Generalized convex optimization
- Distributed optimization
- Global optimization with nonconvex functions
- Multi-objective optimization
- Mixed-integer and combinatorial optimization

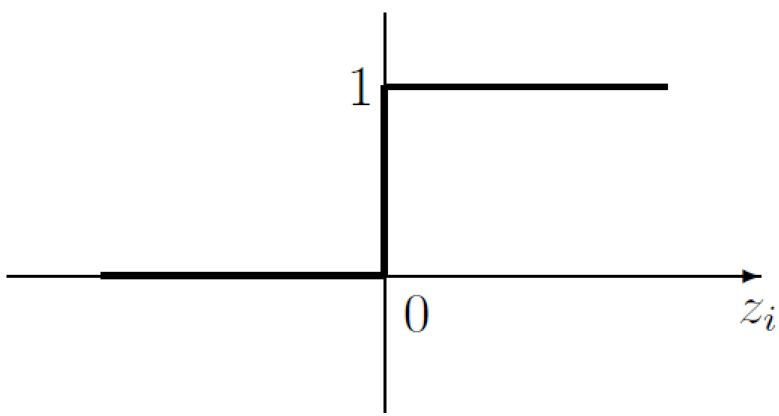
Design Principles

- Smooth penalty function methods
- Lagrangian methods
- Duality methods
- Projection methods
- Nonsmooth penalty function methods
- Multi-agent systems theory and swarm intelligence methods

Simplest RNN for Linear Programming

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{s.t.} & Ax \le b. \end{array}$$

$$\epsilon \frac{dx}{dt} = -\sigma A^T g_{[0,1]}(Ax - b) - c$$



Convergence Condition

- Any equilibrium point is globally stable and an optimal solution to the linear program if

$$\sigma > \frac{\|c\|_2}{\sqrt{\lambda_{\min}(AA^T)}}$$

Q. Liu and J. Wang, "Finite-time convergent recurrent neural network with a hard-limiting activation function for constrained optimization with piecewise-linear objective functions," *IEEE Transactions on Neural Networks*, vol. 22, no. 4, pp. 601-613, 2011.

Nonlinear Programming

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & c(x) \le 0, \quad x \ge 0 \end{array}$$

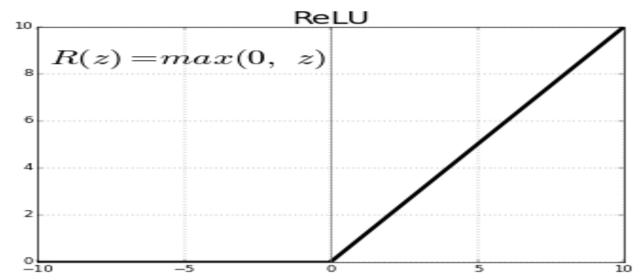
$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Q x + c^T x \\ \text{subject to} & Ax \le b, \quad x \ge 0 \end{array}$$

Projection Networks

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x + (x - \alpha(\nabla f(x) + \nabla c(x)y))^+ \\ -y + (y + \alpha c(x))^+ \end{pmatrix}$$

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x + (x - (Qx + c) + A^T y)^+ \\ -y + (y - Ax + b)^+ \end{pmatrix}$$

$$[x_i]^+ = \max\{0, x_i\}$$



Y. Xia and J. Wang, "A recurrent neural network for nonlinear convex optimization subject to nonlinear inequality constraints," *IEEE Transactions on Circuits and Systems - Part I: Regular Papers*, vol. 51, no. 7, pp. 1385-1394, 2004.

One-layer Neural Net for QP

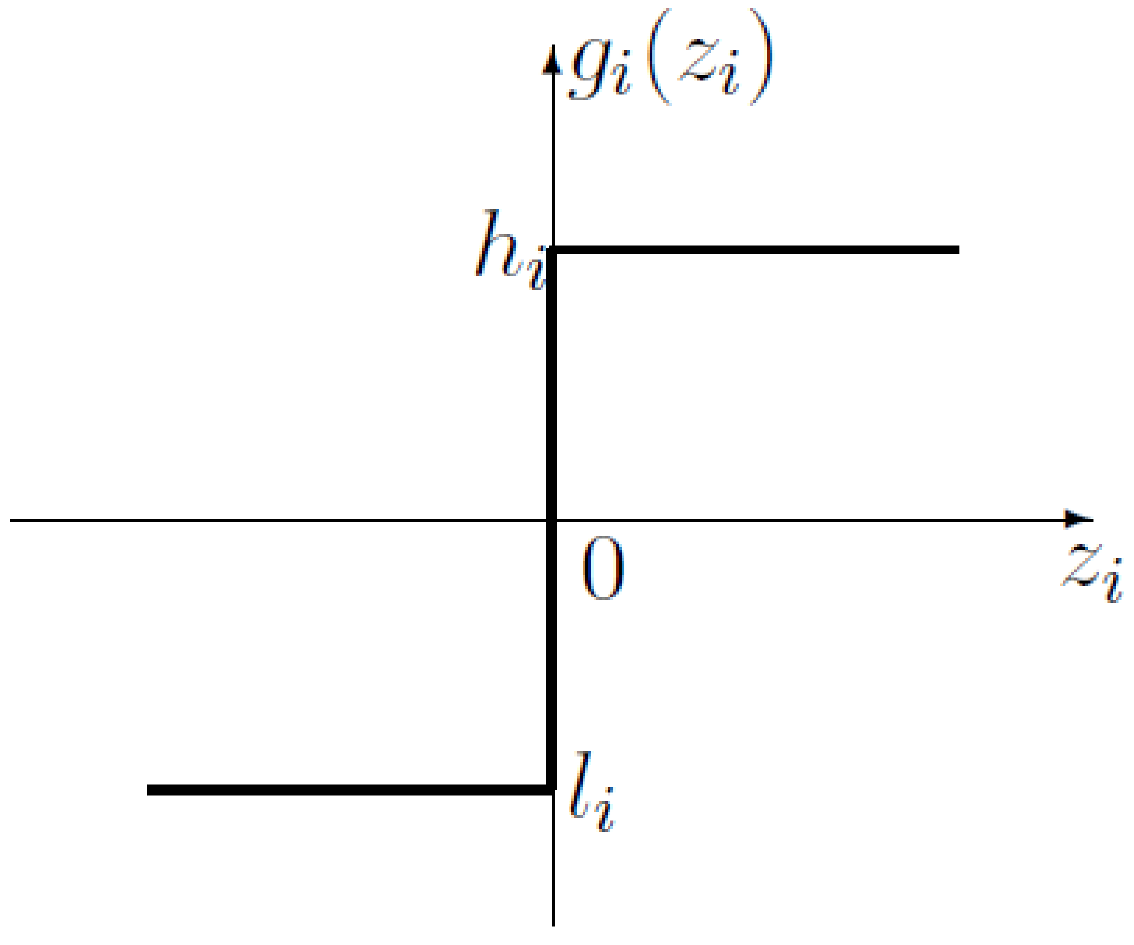
A one-layer recurrent neural net was developed[a]:

$$\epsilon \frac{dz}{dt} = -(I-P)z - [(I-P)Q + \alpha P]g(z) + q,$$
$$x = ((I-P)Q + \alpha P)^{-1}(-(I-P)z + s),$$

where ϵ is a positive scaling constant, $\alpha > 0$ is a parameter, $s = -q + Pq + \alpha A^T(AA^T)^{-1}b$, and $g(\cdot)$ is a vector-valued activation function. $P = A^T(AA^T)^{-1}A$,

[a] Q. Liu, and J. Wang, "A one-layer recurrent neural network with a discontinuous hard-limiting activation function for quadratic programming," *IEEE Transactions on Neural Networks*, vol. 19, no. 4, pp. 558-570, 2008.

Discontinuous Activation Function



Convergence Condition

Assume that the objective function $f(x)$ is strictly convex on the set $S = \{x \in \mathbb{R}^n : Ax = b\}$. If

$$\alpha > \lambda_{\max}(Q^2)\lambda_{\max}(Q^{-1})/4,$$

then the state vector $z(t)$ of the neural network is globally convergent to an equilibrium point and the output vector $x(t)$ is globally convergent to an optimal solution of QP.

Collaborative Neurodynamic Optimization

- For many complex optimization problems, a single neurodynamic optimization model cannot accomplish the tasks.
- More than one neurodynamic optimization models are needed.
- *Collaboration* among the models is essential for the success.

Distributed Optimization

- In many applications, the objective functions are additive:

$$
\begin{aligned}
&\min\ f(x) = \sum_{i=1}^{m} f_i(x) \\
&\text{s.t. } A_i x = b_i \\
&\qquad g_i(x) \le 0, \quad i \in \{1, 2, \ldots, m\}
\end{aligned}
$$

- For examples, data fusion in sensor networks and coordinated operations in swarm robots.
- In such applications, distributed optimization is necessary or desirable.

Collaborative Neurodynamics

A population of coupled neurodynamic models:

$$\begin{cases} \frac{dx_i}{dt} \in 2\Big[-P_i x_i + q_i - (I - P_i)(\partial f_i(x_i) \\ \qquad + (\partial g_i(x_i))^T (z_i + g_i(x_i))^+ \\ \qquad + \sum_{j=1, j\neq i}^{m} a_{ij}(x_i + w_i - x_j - w_j)) \Big] \\ \frac{dz_i}{dt} = -z_i + (z_i + g_i(x_i))^+ \\ \frac{dw_i}{dt} = x_i \end{cases}$$

Global Convergence

- It is proven that the collaborative neurodynamic system is globally convergent or output consensus to optimal solutions if the underlying graph is undirected and connected.
- In its dual formulation for resource allocation, the hidden state vectors *y* or *z* needs to reach a consensus.

Q. Liu, S. Yang, and J. Wang, "A collective neurodynamic approach to distributed constrained optimization," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 28, no. 8, pp. 1747-1758, 2017.

Global Optimization

- Due to nonconvexity, global optimization is much more challenging.
- In recent decades, population-based evolutionary and swarm intelligence algorithms emerged as prevailing methods for global optimization with many success stories in benchmark studies.
- Nevertheless, the meta-heuristic and stochastic natures of the algorithms may not ensure solution consistency or repeatability.

Pros and Cons

- Both neurodynamic optimization and evolutionary optimization approaches have their merits and limitations.
- Neurodynamic approaches are good at constrained and precise local searches with proven convergence, but prone to being trapped at local minima.
- In contrast, evolutionary optimization methods are good at global searches, but weak at constraint handling and guaranteed optimality.

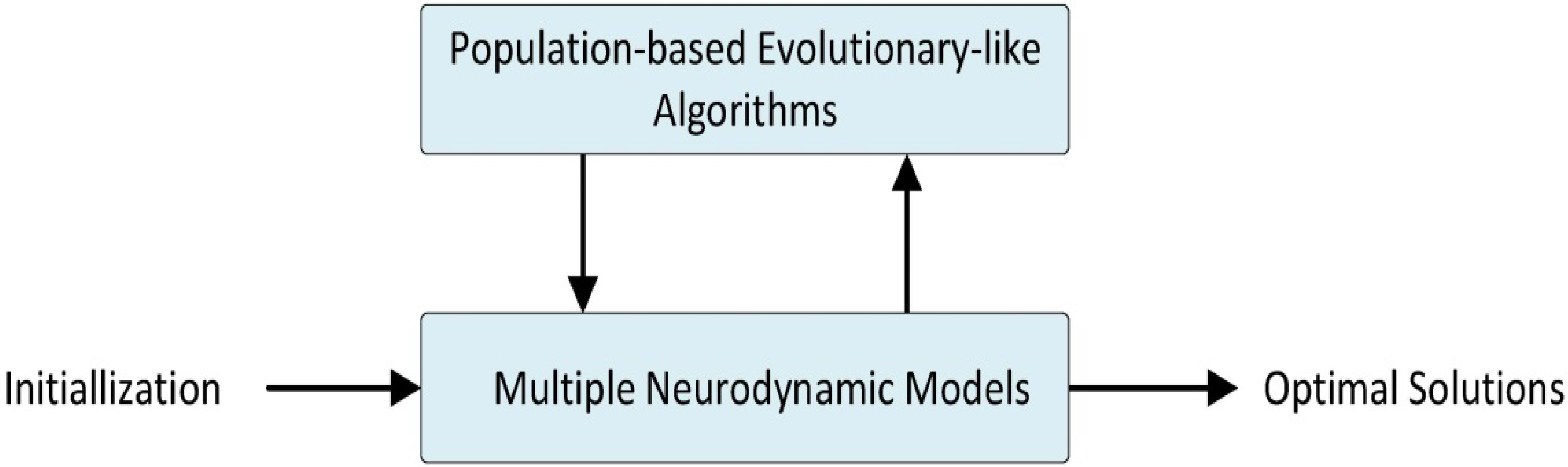
Collaborative Neurodynamic Optimization

- Given the pros and cons of the two types of computationally intelligent optimization approaches, it is natural to integrate them into ones toward hybrid intelligence.
- Collaborative neurodynamic optimization is a hybrid intelligence framework to integrate neurodynamic optimization and swarm intelligence methods.

Collaborative Neurodynamic Optimization (cont'd)

- Multiple projection neural networks are employed for scattered searches.
- A meta-heuristic rule (e.g., PSO) is used to reposition the scattered neurodynamic searches upon their convergence.
- Local searching and global repositioning are carried out alternately until no more reduction of the objective function value could be made.

Collaborative Neurodynamic Optimization (cont'd)



Z. Yan, J. Fan, and J. Wang, "A collective neurodynamic approach to constrained global optimization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 5, pp. 1206-1215, 2017.

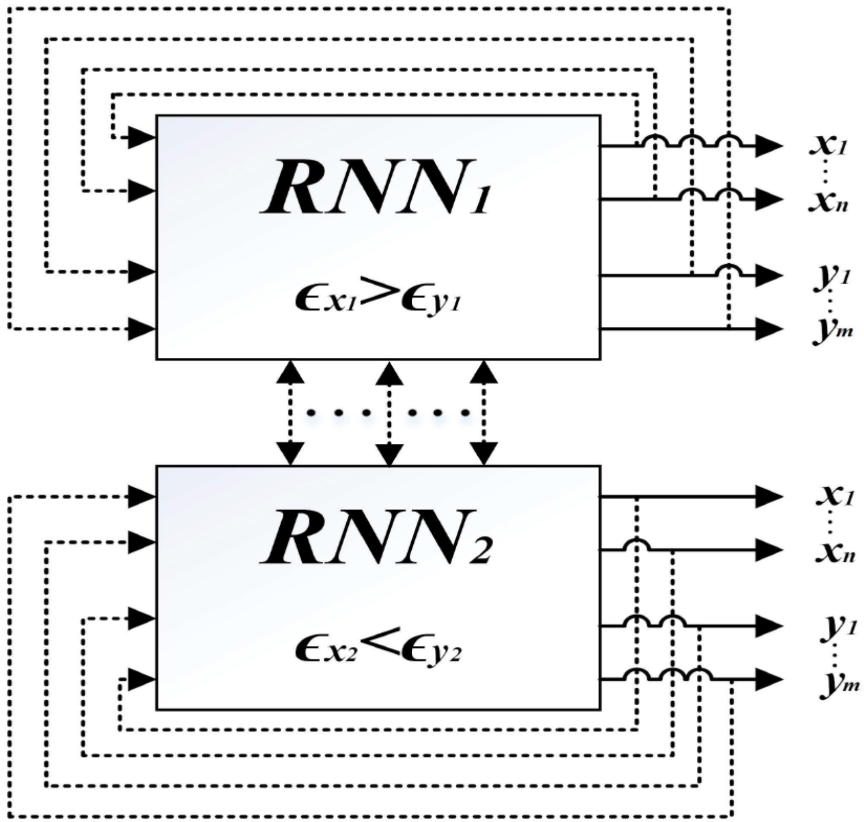
Desirable Properties

- In principle, the collaborative neurodynamic optimization approach is able to find global optimal solutions for any nonconvex objective functions and feasible regions, provided that there are sufficient number of neurodynamic optimization models or sufficient time for search.
- In theory, it is proven that it is globally convergent with probability one (almost sure convergence).

Two-timescale Duplex Neurodynamic Systems

- For a class of special nonconvex functions called biconvex functions, collaborative neurodynamic approaches may be customized.
- The two-timescale Duplex Neurodynamic Systems employs two recurrent neural networks, resulting in minimum spatial complexity.
- They operate in two timescales to enhance diversity.

Two-Timescale Duplex Architecture



Che and J. Wang, "A two-timescale duplex neurodynamic approach to biconvex optimization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 8, pp. 2503-2514, 2019.

Mixed-integer Optimization

$$\min_{\boldsymbol{x},\boldsymbol{y}} \quad f(\boldsymbol{x},\boldsymbol{y})$$

$$\text{s.t.} \quad (\boldsymbol{x}^T,\boldsymbol{y}^T)^T \in \Omega \subseteq \Re^{m+n}$$

$$\boldsymbol{y} \in \{-1,1\}^n \quad \text{or} \quad \boldsymbol{y} \in \{0,1\}^n$$

Constraint Reformulations

- By introducing an instrumental vector *z*, the binary or bipolar constraints cab be converted as one of the following equality/inequality constraints:

TABLE I: Alternative formulations of $\{0,1\}^n$ and $\{-1,1\}^n$.

$\{0,1\}^n$	$(2\boldsymbol{y}-\boldsymbol{e})^T(2\boldsymbol{z}-\boldsymbol{e})=n,\ \Vert 2\boldsymbol{z}-\boldsymbol{e}\Vert_2^2 \le n,\ \boldsymbol{y}\in[0,1]^n$ [12] $\boldsymbol{y}\circ(\boldsymbol{z}-\boldsymbol{e})=\boldsymbol{0}$ and $(\boldsymbol{y}-\boldsymbol{e})\circ\boldsymbol{z}=\boldsymbol{0}$ $(2\boldsymbol{y}-\boldsymbol{e})\circ(2\boldsymbol{z}-\boldsymbol{e})-\boldsymbol{e}=\boldsymbol{0},\ \boldsymbol{y}\in[0,1]^n$, and $\boldsymbol{z}\in[0,1]^n$ $(\boldsymbol{e}-\boldsymbol{z})\circ\boldsymbol{y}\le\boldsymbol{0},\ (\boldsymbol{e}-\boldsymbol{y})\circ\boldsymbol{z}\le\boldsymbol{0},\ \boldsymbol{y}\in[0,1]^n$, and $\boldsymbol{z}\in[0,1]^n$
$\{-1,1\}^n$	$\boldsymbol{y}^T\boldsymbol{z}=n,\ \Vert\boldsymbol{z}\Vert_2^2\le n,\ y\in[-1,1]^n$ [12] $(\boldsymbol{y}+\boldsymbol{e})\circ(\boldsymbol{z}-\boldsymbol{e})=\boldsymbol{0}$ and $(\boldsymbol{y}-\boldsymbol{e})\circ(\boldsymbol{z}+\boldsymbol{e})=\boldsymbol{0}$ $\boldsymbol{y}\circ\boldsymbol{z}-\boldsymbol{e}=\boldsymbol{0},\ \boldsymbol{y}\in[-1,1]^n$, and $\boldsymbol{z}\in[-1,1]^n$ $(\boldsymbol{e}-\boldsymbol{z})\circ(\boldsymbol{e}+\boldsymbol{y})\le\boldsymbol{0},\ (\boldsymbol{e}-\boldsymbol{y})\circ(\boldsymbol{e}+\boldsymbol{z})\le\boldsymbol{0},\ \boldsymbol{y}\in[-1,1]^n$, and $\boldsymbol{z}\in[-1,1]^n$

H. Che and **J. Wang**, "A two-timescale duplex neurodynamic approach to mixed-integer optimization," *IEEE Trans. Neural Networks and Learning Systems*, vol. 32, pp. 36-48, 2021.

Two-Timescale Neurodynamics

$$\epsilon_{\tilde{x}} \frac{d\tilde{\boldsymbol{x}}}{dt} = P(\nabla_{\tilde{x}} f(\tilde{\boldsymbol{x}}, \boldsymbol{y}), \tilde{\boldsymbol{x}}),$$

$$\epsilon_{y} \frac{d\boldsymbol{y}}{dt} = P(\nabla_{y} f(\tilde{\boldsymbol{x}}, \boldsymbol{y}), \boldsymbol{y}),$$

where $\tilde{\boldsymbol{x}} = (\boldsymbol{x}^T, \boldsymbol{z}^T)^T$

Existing CNO Paradigms

- Nonnegative matrix factorization
- Feature selection
- Constrained clustering
- Supervised learning
- Portfolio selection
- Hash-bit selection
- Sparse signal reconstruction
- Vehicle-task assignment
- Model predictive control
- HVAC operation planning and control

Feature Selection

- Feature selection is an essential part of data processing.
- It aims at selecting a subset of the most informative features from all available features.
- Apart from the learning capability of neural networks for feature selection, the optimization capability of recurrent neural networks can also play a vital role.

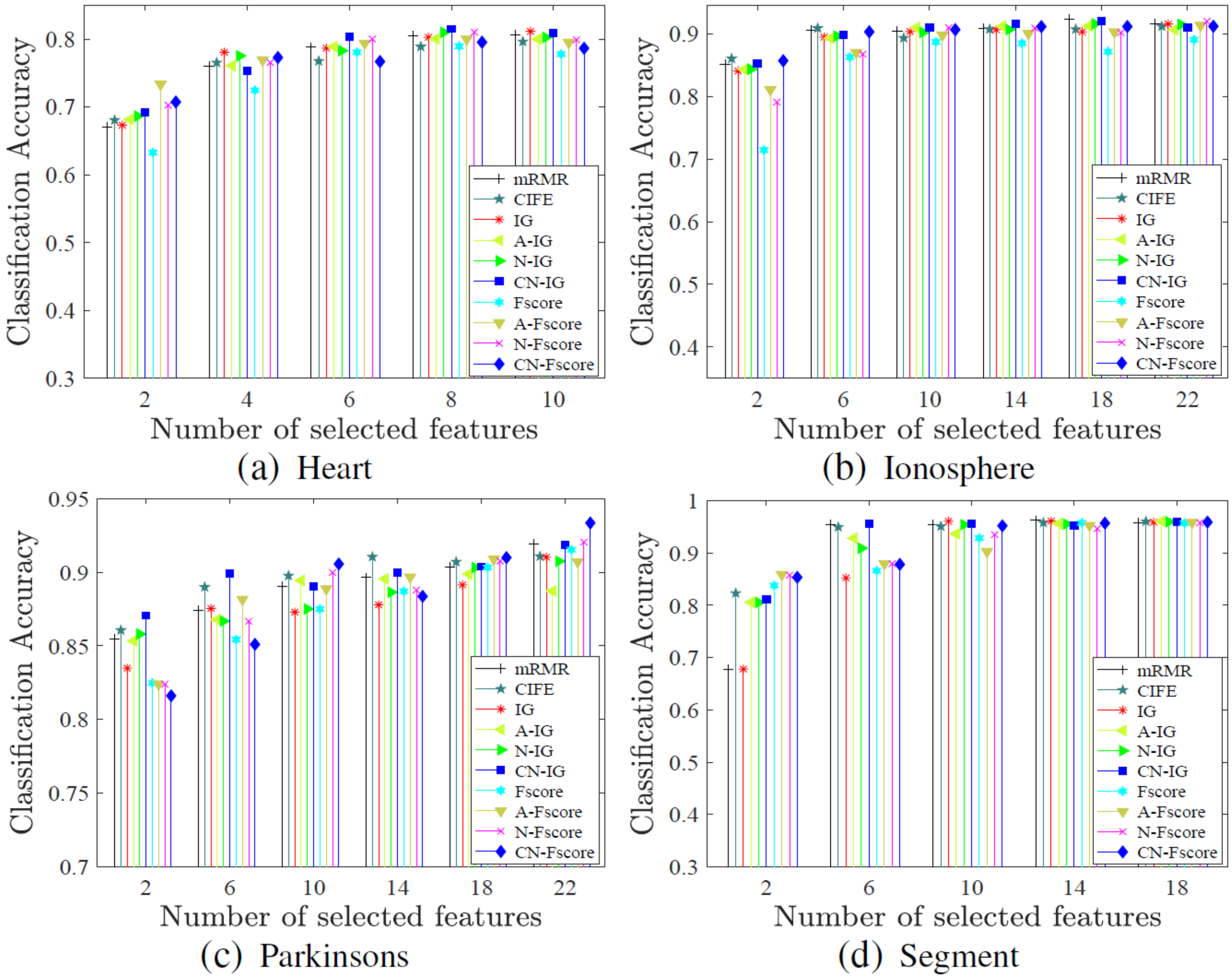
Problem Formulations

$$\min_{w,\gamma} \quad f(w,\gamma) = \frac{\gamma^2}{2} w^T Q w - \gamma \tau^T w,$$

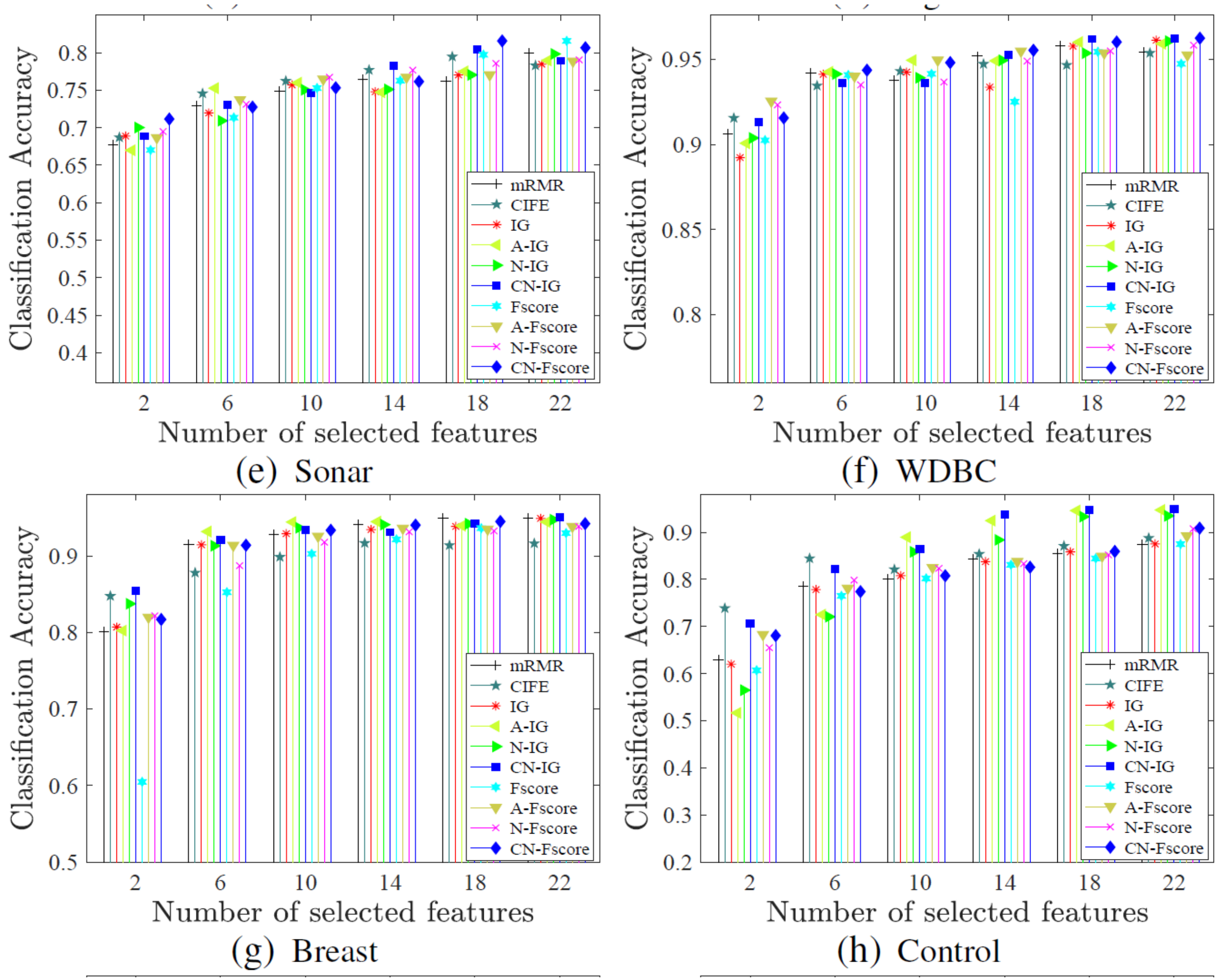
$$\text{s.t.} \quad \sum_{j=1}^{p} w_j = 1,$$

$$w \geq 0, \gamma > 0.$$

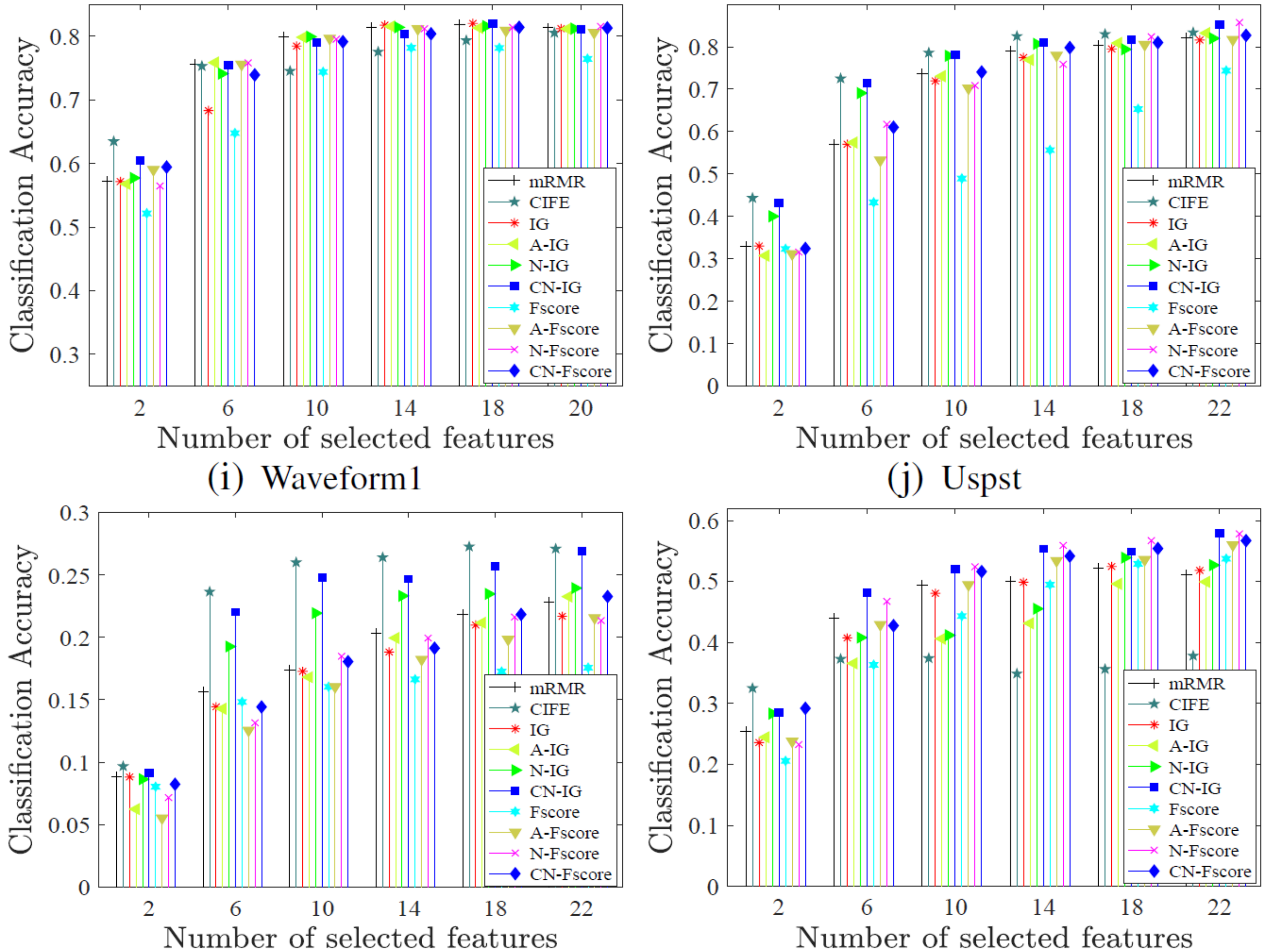
Classification Accuracies



Classification Accuracies



Classification Accuracies



(i) Waveform1

(j) Uspst

(k) Corel_5k

(l) Yale

Classification Accuracies

TABLE IV: Average SVM classification accuracies of the ten feature selection methods on the thirteen benchmark datasets.

Dataset	Method									
	mRMR	CIFE	IG	A-IG	N-IG	CN-IG	Fscore	A-Fscore	N-Fscore	CN-Fscore
Heart	0.8060	0.8006	0.8113	0.8096	0.8110	**0.8135**	0.7394	0.8133	0.8073	0.7991
Ionosphere	0.9178	0.9128	0.9166	0.9170	0.9180	**0.9191**	0.8610	0.9103	0.9109	0.9158
Parkinsons	0.8693	0.8734	0.8644	**0.8790**	0.8644	0.8675	0.8622	0.8772	0.8636	0.8661
Segment	0.8770	0.9066	0.8459	0.8798	0.8778	**0.9081**	0.8814	0.8843	0.8862	0.8919
Sonar	0.7551	**0.7844**	0.7471	0.7579	0.7471	0.7627	0.7726	0.7740	0.7742	0.7763
WDBC	0.9500	0.9513	0.9447	0.9563	0.9498	0.9548	0.9356	0.9604	0.9575	**0.9606**
Breast	0.9430	0.9225	0.9366	0.9403	0.9395	0.9442	0.8624	0.9382	0.9383	**0.9481**
Control	0.8281	0.8876	0.8241	0.8555	0.8433	**0.8982**	0.7920	0.8408	0.8451	0.8390
Waveform1	0.8085	0.7896	0.7973	0.8030	0.7995	0.8032	0.7076	**0.8064**	**0.8064**	0.7946
Uspst	0.7083	0.7704	0.6971	0.7040	0.7200	**0.7739**	0.5768	0.6848	0.7118	0.6945
Corel_5k	0.2100	**0.3034**	0.1971	0.1915	0.1995	0.2618	0.1705	0.1768	0.1955	0.2171
Yale	0.4498	0.3387	0.4172	0.3712	0.3644	0.4431	0.3744	**0.4568**	0.4490	0.4530
Brain	0.7705	0.7306	0.7661	0.6817	0.7859	**0.7863**	0.7346	0.6489	0.6961	0.6950
Average	0.7610	0.7671	0.7512	0.7498	0.7554	**0.7797**	0.7131	0.7517	0.7571	0.7578

Y. Wang, J. Wang, and N. R. Pal, "Supervised feature selection via collaborative neurodynamic optimization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 5, pp. 6878-6892, 2024.

Classification Accuracies

TABLE V: Average 1-NN classification accuracies of the ten feature selection methods on the thirteen benchmark datasets.

Dataset	Method									
	mRMR	CIFE	IG	A-IG	N-IG	CN-IG	Fscore	A-Fscore	N-Fscore	CN-Fscore
Heart	0.7129	0.7111	0.7313	0.7269	0.7313	**0.7357**	0.7306	0.7350	0.7306	0.7277
Ionosphere	0.8780	0.8760	0.8619	0.8678	0.8679	0.8715	0.8495	0.8503	0.8515	**0.8821**
Parkinsons	0.9246	0.9288	0.8971	0.8990	0.9007	**0.9325**	0.8907	0.8993	0.9078	0.9016
Segment	0.9160	**0.9415**	0.9013	0.9330	0.9366	0.9361	0.9205	0.9206	0.9318	0.9339
Sonar	0.7680	0.7670	0.7622	0.7642	0.7647	**0.7707**	0.7447	0.7447	0.7532	0.7657
WDBC	0.9389	0.9381	0.9357	0.9402	0.9403	0.9374	0.9352	0.9390	0.9357	**0.9416**
Breast	0.8593	0.8412	0.8651	0.8731	0.8803	**0.8817**	0.8465	0.8607	0.8469	0.8615
Control	0.7902	0.7609	0.7864	0.8215	0.8054	**0.8557**	0.7867	0.7870	0.7882	0.7961
Waveform1	0.7257	0.7129	0.7107	0.7216	0.7216	0.7364	0.7054	0.7349	0.7236	**0.7368**
Uspst	0.6519	**0.7330**	0.6490	0.6491	0.7090	0.7137	0.5002	0.6513	0.6317	0.6520
Corel_5k	0.1708	0.2053	0.1649	0.1690	0.2098	**0.2107**	0.1441	0.1463	0.1553	0.1483
Yale	0.5019	0.3998	0.5004	0.4616	0.5143	**0.5537**	0.4601	0.4744	0.5449	0.5017
Brain	**0.6832**	0.6288	0.6749	0.6208	0.6346	0.6760	0.6150	0.6674	0.6615	0.6753
Average	0.7324	0.7265	0.7262	0.7268	0.7397	**0.7548**	0.7023	0.7239	0.7279	0.7326

Y. Wang, J. Wang, and N. R. Pal, "Supervised feature selection via collaborative neurodynamic optimization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 5, pp. 6878-6892, 2024.

Classification Accuracies

TABLE VI: Average RF classification accuracies of the ten feature selection methods on the thirteen benchmark datasets.

Dataset	Method									
	mRMR	CIFE	IG	A-IG	N-IG	CN-IG	Fscore	A-Fscore	N-Fscore	CN-Fscore
Heart	0.7788	0.7675	0.7698	0.7624	0.7720	**0.7748**	0.7529	0.7867	0.7884	0.7703
Ionosphere	0.9101	0.9065	0.9044	0.9048	0.9049	**0.9134**	0.8456	0.8880	0.8880	0.9036
Parkinsons	0.8752	0.8861	0.8697	0.8706	0.8834	**0.8914**	0.8768	0.8770	0.8818	0.8822
Segment	0.9121	0.9377	0.8997	0.9366	0.9362	**0.9379**	0.9267	0.9269	0.9285	0.9344
Sonar	0.7170	0.7236	0.7250	0.7252	0.7281	0.7383	0.7389	0.7391	0.7459	**0.7461**
WDBC	0.9361	0.9313	0.9340	0.9343	0.9342	0.9393	0.9351	0.9388	0.9355	**0.9406**
Breast	0.9397	0.9219	0.9344	0.9400	0.9389	**0.9403**	0.8648	0.9356	0.9295	0.9365
Control	0.7761	0.8605	0.7786	0.7980	0.7995	**0.8606**	0.7837	0.8071	0.8018	0.7933
Waveform1	0.7516	0.7512	0.7361	0.7577	**0.7586**	0.7518	0.7072	0.7434	0.7493	0.7457
Uspst	0.6657	0.7185	0.6566	0.6582	0.7068	**0.7163**	0.5224	0.6387	0.6972	0.6813
Corel_5k	0.1538	0.1920	0.1481	0.1482	**0.1937**	0.1934	0.1373	0.1458	0.1578	0.1595
Yale	0.4090	0.3393	0.4152	0.3889	0.4336	0.4874	0.4522	0.4650	0.4705	**0.4946**
Brain	0.6490	0.6501	0.6357	0.6651	0.6559	0.6507	0.6563	0.6408	0.6351	**0.6886**
Average	0.7288	0.7374	0.7236	0.7300	0.7420	**0.7535**	0.7077	0.7333	0.7392	0.7444

Y. Wang, J. Wang, and N. R. Pal, "Supervised feature selection via collaborative neurodynamic optimization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 5, pp. 6878-6892, 2024.

NMF Problem

A representation of patterns as a linear combination of bases:

$$V \approx WH$$

where

V: $n \times m$ matrix. Each column contains n nonnegative values of one of m patterns.

W: ($n \times p$): p columns of W are basis vectors.

H: ($p \times m$): each column of H is a weight vector.

D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788-791, 1999.

Problem Formulations

$$\begin{aligned} \min \quad & f(W, H) \\ \text{s.t.} \quad & W \geq 0, H \geq 0 \end{aligned}$$

Squared Frobenius norm:

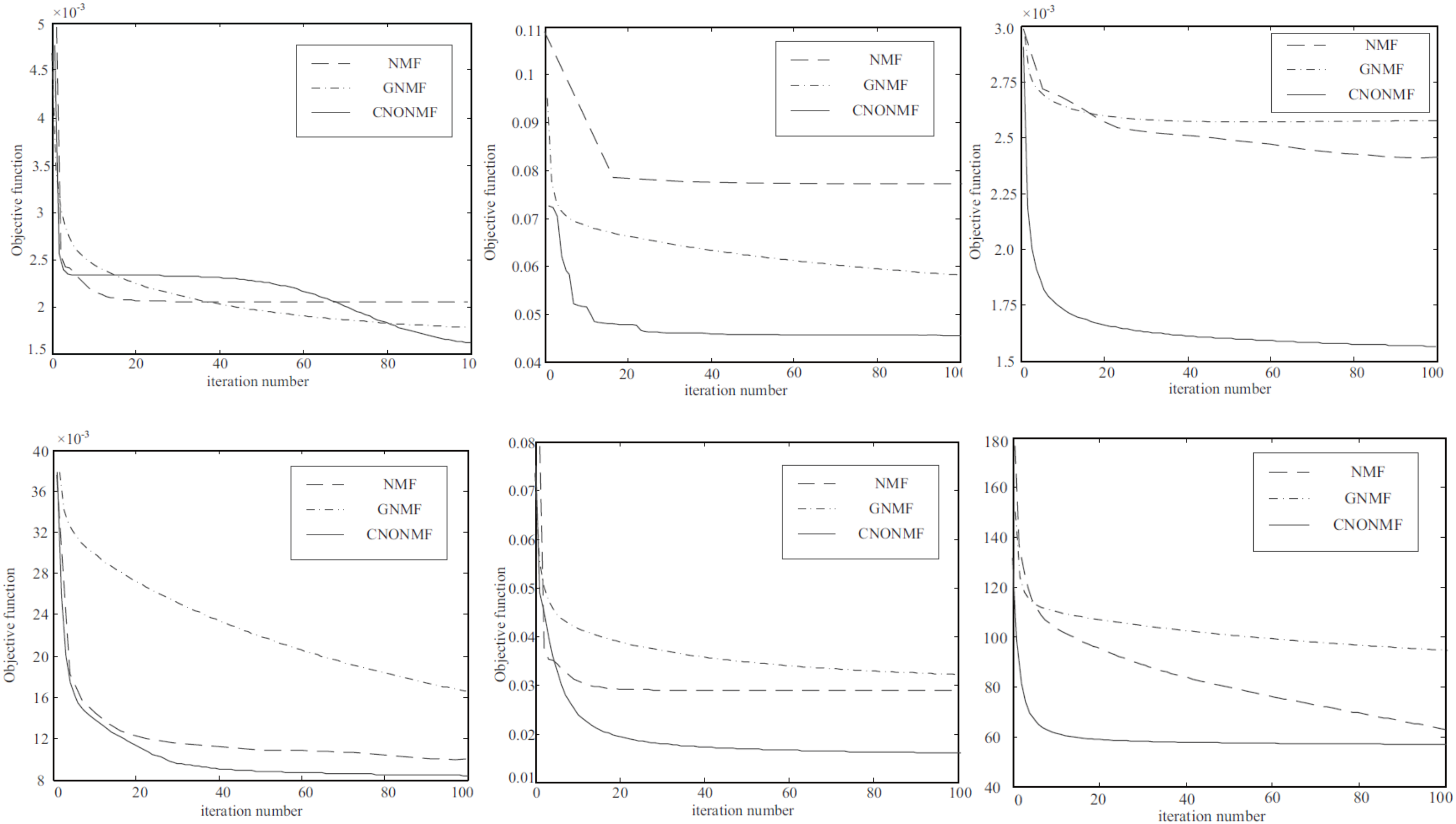
$$f_1(W,H) = \|V - WH\|_2^2 = \sum_{i,j} (V_{ij} - \sum_{k=1}^{r} w_{ik} v_{kj})^2$$

Kullback-Leibler divergence:

$$f_2(W,H) = D(V\|WH) =$$

$$\sum_{i,j} (V_{ij} \log \frac{V_{ij}}{\sum_{k=1}^{r} w_{ik} v_{kj}} V_{ij} - \sum_{k=1}^{r} w_{ik} v_{kj})$$

Convergent Behaviors



Clustering Results (accuracy)

CLUSTERING PERFORMANCE OF THE ACCURACY INDEX [AC (%)]

Dataset	IRIS	Wine	Digit	ORLFace	COIL20	PIE	TDT2
NMF [11]	78.0 ± 5.2	88.7 ± 0.5	97.2 ± 0.0	61.2 ± 0.0	60.5 ± 0.0	63.3 ± 3.7	79.9 ± 11.7
MUNMF [17]	75.5 ± 6.1	88.8 ± 0.8	93.8 ± 0.2	15.6 ± 0.0	13.4 ± 0.0	68.7 ± 2.5	62.6 ± 11.4
APGNMF [20]	88.1 ± 1.0	92.3 ± 0.1	97.2 ± 0.0	39.2 ± 0.0	61.2 ± 0.0	71.4 ± 4.2	87.6 ± 9.2
HALS [18]	88.1 ± 1.0	89.6 ± 0.2	97.2 ± 0.0	65.8 ± 0.0	70.5 ± 0.0	73.9 ± 2.1	81.1 ± 3.7
GNMF [21]	90.7 ± 0.0	93.9 ± 0.0	$\mathbf{99.8 \pm 0.0}$	64.5 ± 0.0	75.3 ± 0.0	78.9 ± 4.5	93.4 ± 2.7
RNNNMF [29]	73.3 ± 4.3	91.0 ± 0.3	92.4 ± 0.0	58.5 ± 0.0	42.0 ± 0.0	66.5 ± 3.1	82.5 ± 4.4
CNONMF [34]	$\mathbf{97.3 \pm 0.0}$	91.5 ± 0.0	97.2 ± 0.0	65.8 ± 0.0	70.5 ± 0.0	75.1 ± 2.2	87.8 ± 1.5
CNOWMNMF	$\mathbf{97.3 \pm 0.0}$	$\mathbf{94.6 \pm 0.0}$	97.9 ± 0.0	70.0 ± 0.0	70.5 ± 0.0	78.5 ± 0.9	88.5 ± 1.3
CNOWMGNMF	90.7 ± 0.0	93.9 ± 0.0	$\mathbf{99.8 \pm 0.0}$	$\mathbf{71.3 \pm 0.0}$	$\mathbf{79.7 \pm 0.0}$	$\mathbf{80.1 \pm 1.4}$	$\mathbf{94.1 \pm 1.2}$

J. Fan and J. Wang, “A collective neurodynamic optimization approach to nonnegative matrix factorization,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, pp. 2344-2356, 2017.

Clustering Results (NMI)

CLUSTERING PERFORMANCE OF THE NORMALIZED MUTUAL INFORMATION INDEX (*NMI* (%))

Dataset	IRIS	Wine	Digit	ORLFace	COIL20	PIE	TDT2
NMF [11]	60.8 ± 6.7	73.7 ± 0.7	89.2 ± 0.0	78.6 ± 0.0	72.5 ± 0.0	80.4 ± 1.1	82.0 ± 9.2
MUNMF [17]	61.5 ± 3.3	70.6 ± 1.1	83.7 ± 0.5	34.8 ± 0.0	11.5 ± 0.0	85.9 ± 1.6	75.2 ± 8.9
APGNMF [20]	74.3 ± 2.6	75.3 ± 0.6	91.2 ± 0.0	62.9 ± 0.0	69.3 ± 0.0	86.5 ± 1.5	86.3 ± 8.4
HALS [18]	74.3 ± 2.0	74.2 ± 1.0	89.2 ± 0.0	79.4 ± 0.0	77.0 ± 0.0	87.4 ± 0.9	82.1 ± 3.1
GNMF [21]	79.3 ± 0.0	80.6 ± 0.0	$\mathbf{98.8 \pm 0.0}$	80.5 ± 0.0	87.5 ± 0.0	89.1 ± 1.6	88.0 ± 5.7
CNONMF	$\mathbf{90.8 \pm 0.1}$	$\mathbf{80.9 \pm 0.0}$	91.7 ± 0.0	81.9 ± 0.0	77.0 ± 0.0	88.8 ± 0.2	86.6 ± 1.1
CNOGNMF	79.3 ± 0.0	80.6 ± 0.0	$\mathbf{98.8 \pm 0.0}$	$\mathbf{83.8 \pm 0.0}$	$\mathbf{91.5 \pm 0.0}$	$\mathbf{89.2 \pm 0.4}$	$\mathbf{91.2 \pm 1.2}$

J. Fan and J. Wang, “A collective neurodynamic optimization approach to nonnegative matrix factorization,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, pp. 2344-2356, 2017.

Sparse Nonnegative Matrix Factorization

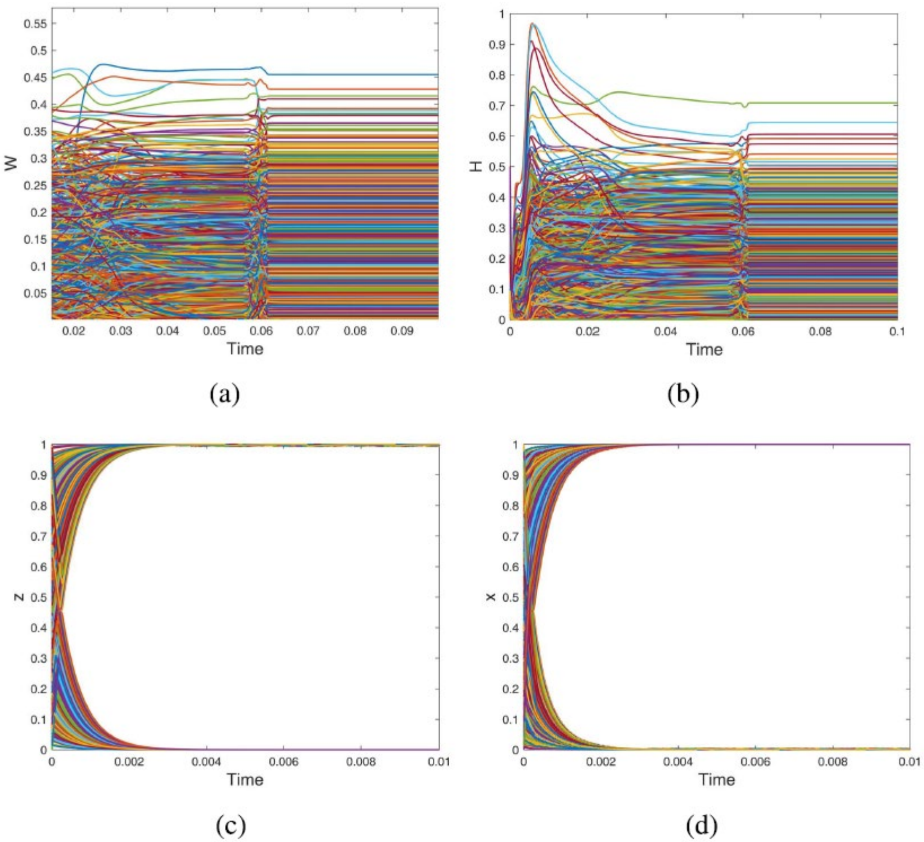
- Factorized matrices with higher sparsity levels show stronger robustness against noises and occupy less storage space.

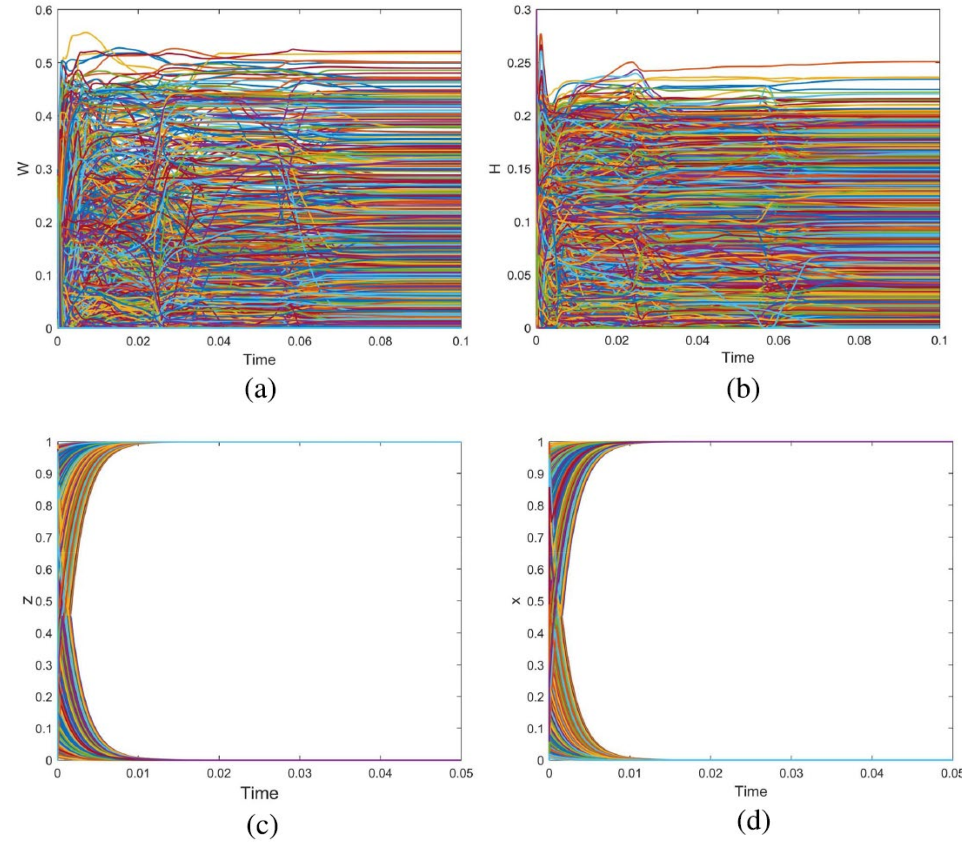
$$\min_{\boldsymbol{W},\boldsymbol{H},\boldsymbol{y}} \|\boldsymbol{V} - \boldsymbol{W}\boldsymbol{H}\|_F^2 + \gamma \sum_{i=1}^{m} \sum_{j=1}^{r} y_{ij}$$

$$\text{s.t. } 0 \le w_{ij} \le M y_{ij}, \quad i = 1, \ldots, m, \ j = 1, \ldots, r$$

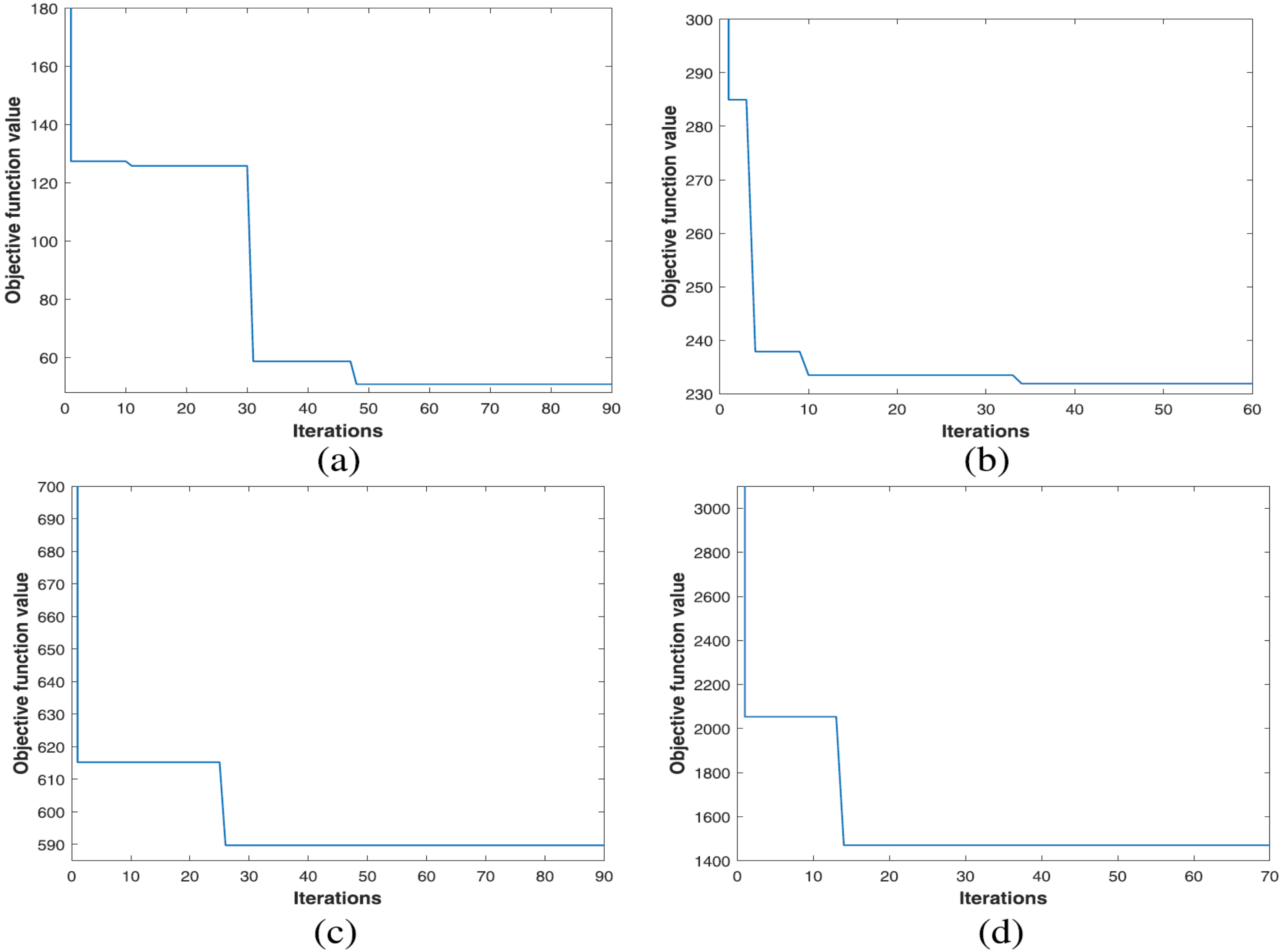
$$\boldsymbol{y} \in \{0, 1\}^{\mathrm{mr}}, \quad h_{pq} \ge 0, \ p = 1, \ldots, r, \ q = 1, \ldots, n$$

Neuronal Convergence





Objective Minimization



Evaluation Measures

$$\mathrm{RRE} = \frac{\| V - WH \|_F^2}{\| V \|_F^2}, \quad \mathrm{SEM} = \sum_{j=1}^{r} \frac{\| \boldsymbol{w}_j \|_0}{\mathrm{mr}}$$

$$\mathrm{GS} = \frac{1}{2}\left(\exp\left(-\frac{\mathrm{RRE}^2}{\sigma_{\mathrm{RRE}}^2}\right) + \exp\left(-\frac{\mathrm{SEM}^2}{\sigma_{\mathrm{SEM}}^2}\right)\right).$$

Comparative Results (Yale)

Performance measure	RRE	SEM	GS
NMFSC [17]	0.0572	0.4746	0.5943
NMFSMU [69]	0.9999	0.4527	0.1218
GA-MIOP [70]	Infeasible	Infeasible	Infeasible
rsNNLS [19]	0.2023	0.4983	0.4668
NeNMF [53]	0.0658	0.4479	0.6103
NMFS [71]	0.9991	0.4722	0.1077
MOSNMF [21]	0.0923	0.3096	0.7293
Tanh-NMF [23]	0.0427	0.4516	0.6159
DTPNN-SNMF [22]	0.0625	0.2751	0.7831
Method herein	0.2068	0.0859	**0.8473**

H. Che, J. Wang, and A. Cichocki, "Bicriteria sparse nonnegative matrix factorization via two-timescale duplex neurodynamic optimization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 8, 2023.

Comparative Results (ORL)

Performance measure	RRE	SEM	GS
NMFSC [17]	0.0166	0.7949	0.5054
NMFSMU [69]	0.9999	0.4116	0.1557
GA-MIOP [70]	Infeasible	Infeasible	Infeasible
rsNNLS [19]	0.0748	0.4983	0.5710
NeNMF [53]	0.0172	0.4430	0.6279
NMFS [71]	0.9990	0.4973	0.0911
MOSNMF [21]	0.0198	0.1019	0.9641
Tanh-NMF [23]	0.0050	0.4589	0.6166
DTPNN-SNMF [22]	0.0105	0.3064	0.7611
Method herein	0.0191	0.0768	**0.9788**

H. Che, J. Wang, and A. Cichocki, "Bicriteria sparse nonnegative matrix factorization via two-timescale duplex neurodynamic optimization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 8, 2023.

Sparse Bayesian Regression

- Regression in a sparse Bayesian learning framework is usually formulated as a global optimization problem with a nonconvex objective function.
- Due to the nonconvexity, the solution quality and consistency depend heavily on the initial values of the optimization solver.

Problem Formulation

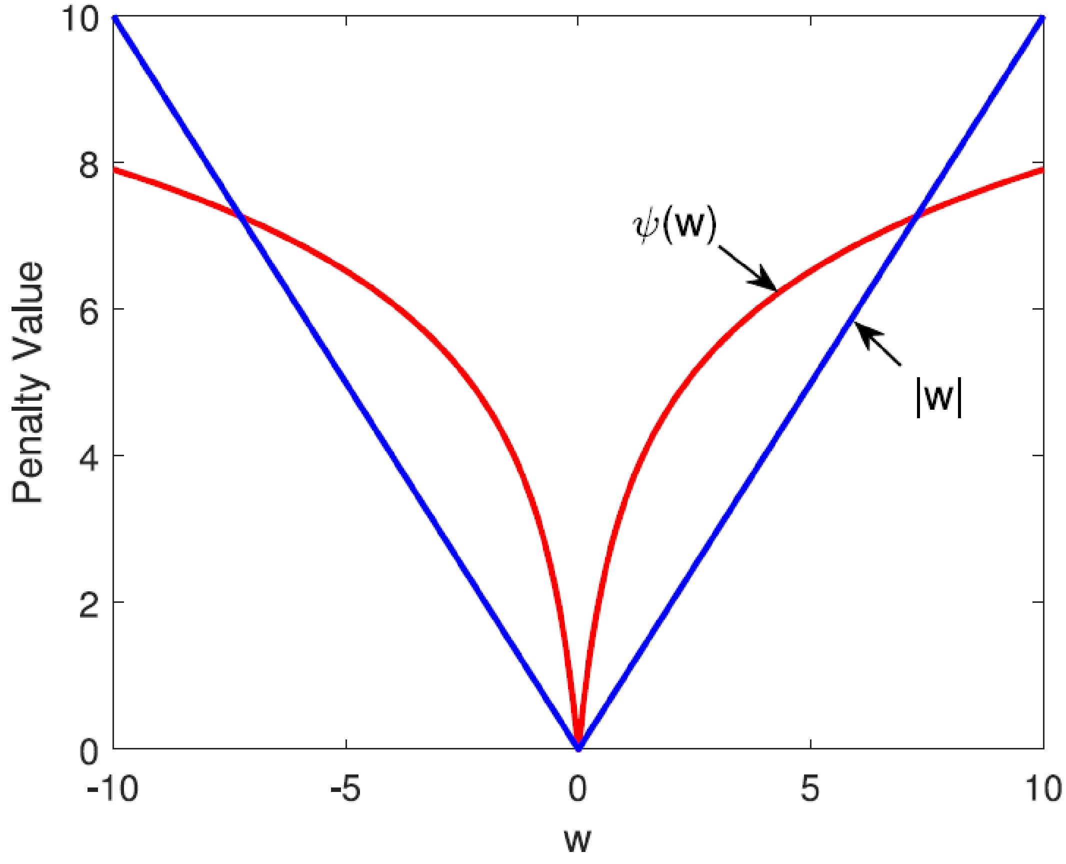
$$\min \; L(w, \sigma, \gamma)$$
$$\text{s.t.} \;\; \sigma \geq \varepsilon, \;\; \gamma_i \geq \varepsilon, \qquad i = 1, \ldots, n.$$

$$L(w, \sigma, \gamma) = \frac{\|y - \Phi w\|^2}{\sigma} + \sum_{i=1}^{n} f(w_i, \sigma, \gamma_i) + (m-2) \ln \sigma$$

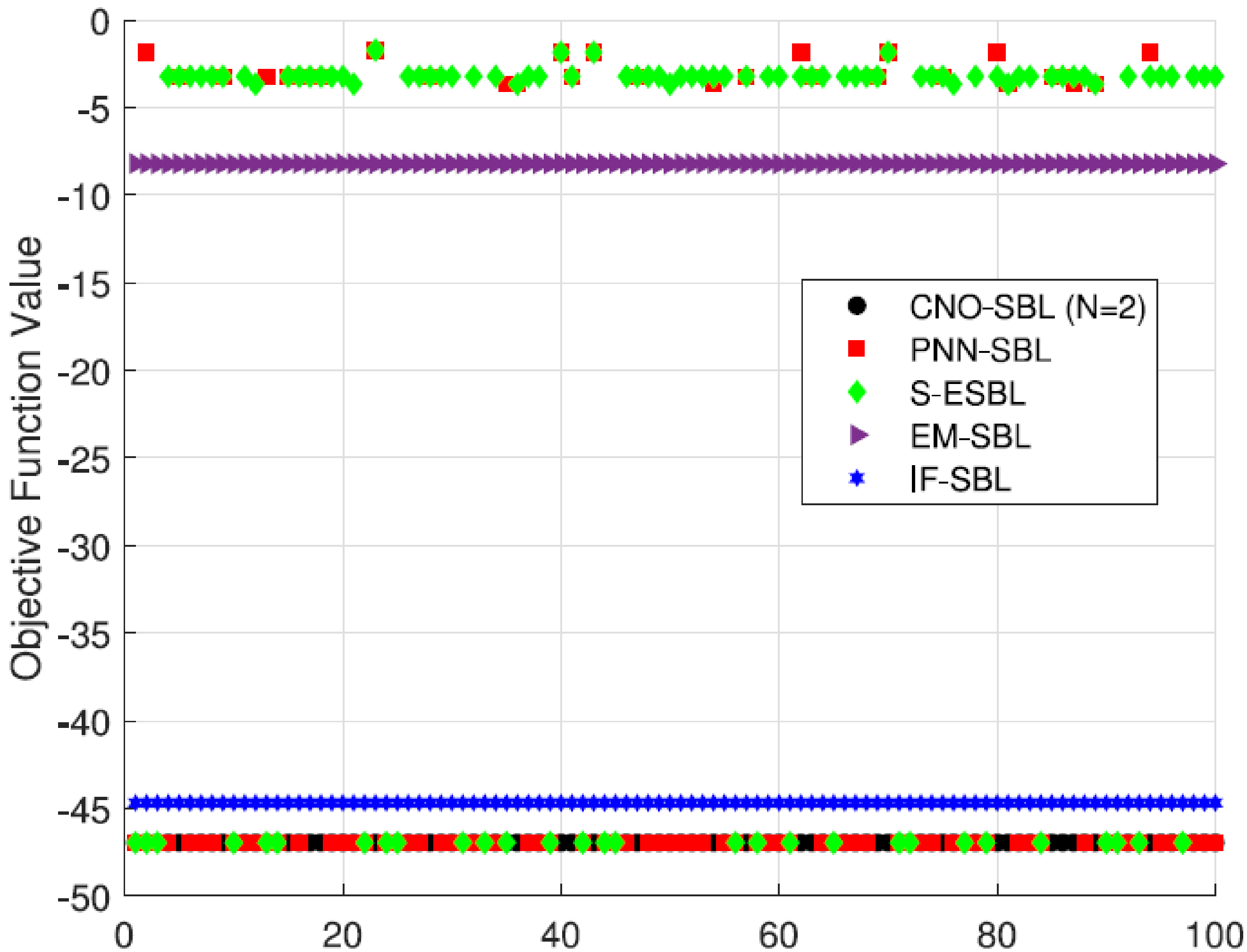
$$f(w_i, \sigma, \gamma_i) := (w_i^2 / \gamma_i) + \ln(1 + [s_0 \gamma_i / \sigma]).$$

W. Zhou, H. Zhang, and J. Wang, "Sparse Bayesian learning based on collaborative neurodynamic optimization," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, 2022.

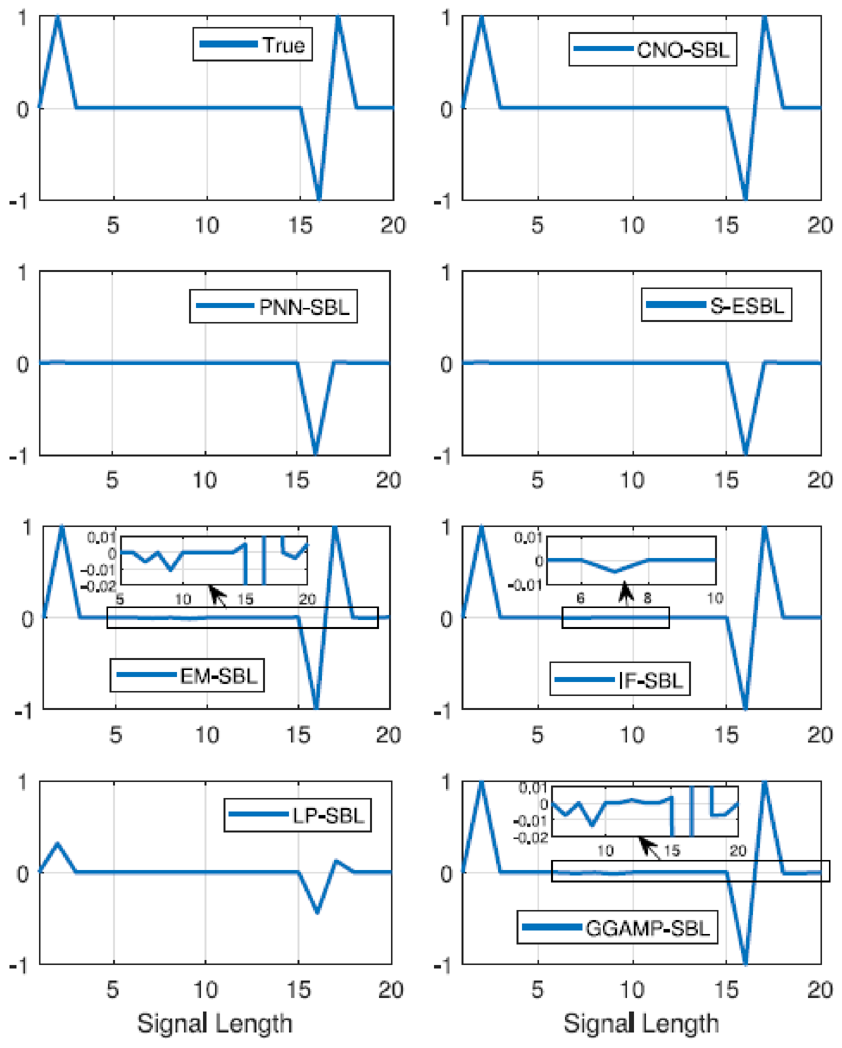
Sparsity-inducing Effect



Objective Function Values



Sparse Signal Reconstruction



Sparse Signal Reconstruction

Signal Types	Methods	nRMSE (mean+std)	SR	$\|\widehat{w}\|_0$
± 1 Spike	PNN-SBL	0.2675 ± 0.4141	71%	4.03
	S-ESBL [26]	0.5965 ± 0.3820	29%	5.64
	EM-SBL [9]	0.0111 ± 0.0000	0%	9.00
	IF-SBL [23]	0.0062 ± 0.0000	0%	4.00
	LP-SBL [15]	0.6286 ± 0.2727	0%	2.44
	GGAMP-SBL [24]	0.0127 ± 0.0000	0%	9.00
	CNO-SBL ($N = 2$)	$\mathbf{0.0057 \pm 0.0000}$	**100%**	**3.00**
Gaussian	PNN-SBL	0.2109 ± 0.3833	71%	3.76
	S-ESBL [26]	0.2451 ± 0.4125	68%	3.92
	EM-SBL [9]	0.0273 ± 0.0000	0%	11.00
	IF-SBL [23]	0.1065 ± 0.2717	84%	3.39
	LP-SBL [15]	0.5807 ± 0.2429	2%	1.80
	GGAMP-SBL [24]	0.0207 ± 0.0000	0%	10.00
	CNO-SBL ($N = 2$)	$\mathbf{0.0014 \pm 0.0000}$	**100%**	**3.00**
ST	PNN-SBL	0.4571 ± 0.5325	58%	3.52
	S-ESBL [26]	0.4680 ± 0.5328	57%	3.63
	EM-SBL [9]	0.0185 ± 0.0000	0%	9.00
	IF-SBL [23]	0.0515 ± 0.2168	96%	3.14
	LP-SBL [15]	0.5279 ± 0.2957	0%	2.95
	GGAMP-SBL [24]	0.0204 ± 0.0000	0%	11.00
	CNO-SBL ($N = 2$)	$\mathbf{0.0074 \pm 0.0000}$	**100%**	**3.00**

PDE Identification

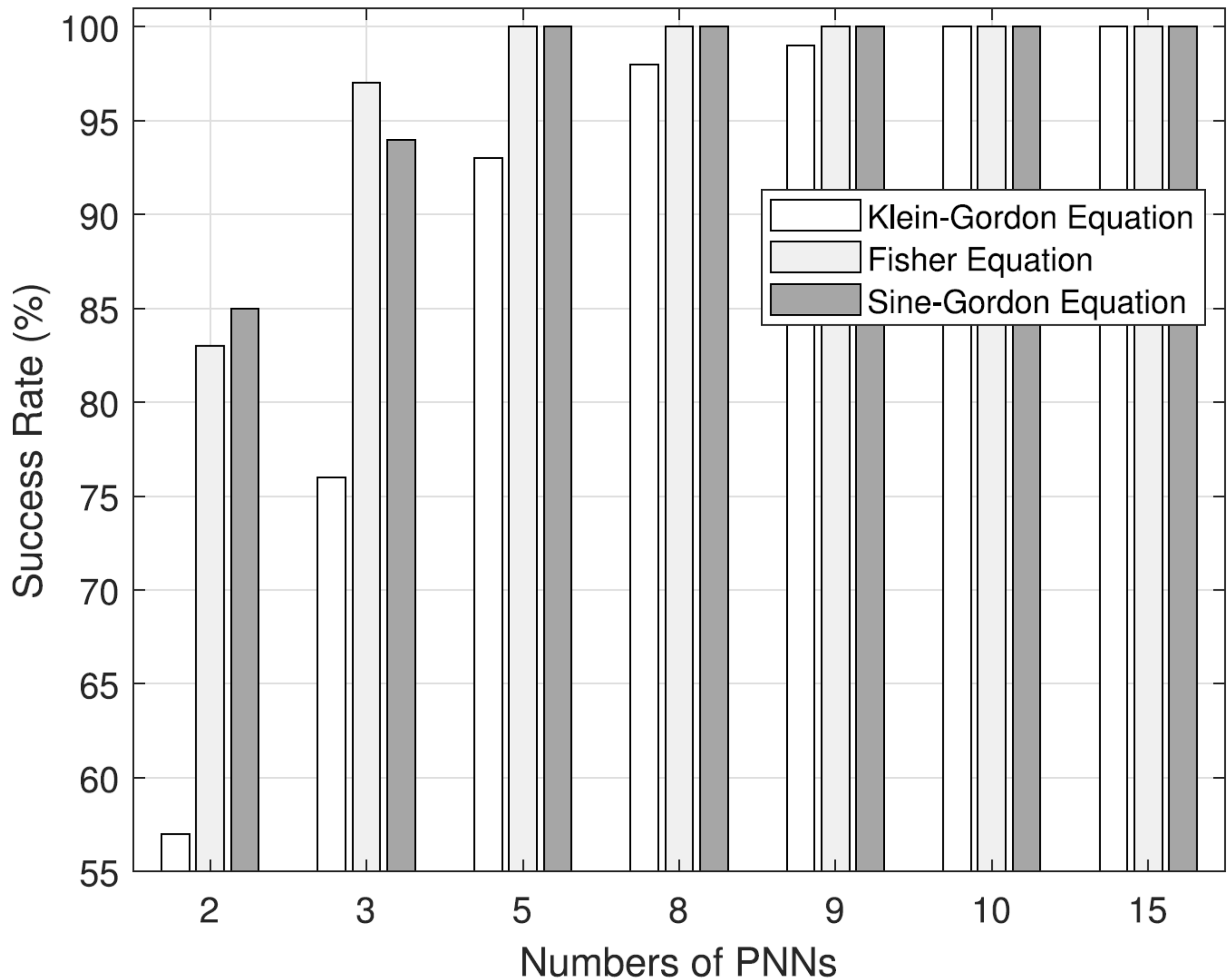
THREE TYPICAL PDES AND IDENTIFICATION RESULTS USING CNO-SBL

Name	PDE to be identified	Identified PDE
Klein-Gordon equation	$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} - u - u^3$	$\frac{\partial^2 u}{\partial t^2} = 0.9995 \frac{\partial^2 u}{\partial x^2} - 1.0003u - 0.9965u^3$
Fisher's equation	$\frac{\partial u}{\partial t} = u - u^2 + 0.1 \frac{\partial^2 u}{\partial x^2}$	$\frac{\partial u}{\partial t} = 1.0009u - 1.0012u^2 + 0.0998 \frac{\partial^2 u}{\partial x^2}$
Sine-Gordon equation	$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} - \sin(u)$	$\frac{\partial^2 u}{\partial t^2} = 0.9990 \frac{\partial^2 u}{\partial x^2} - 0.9987 \sin(u)$

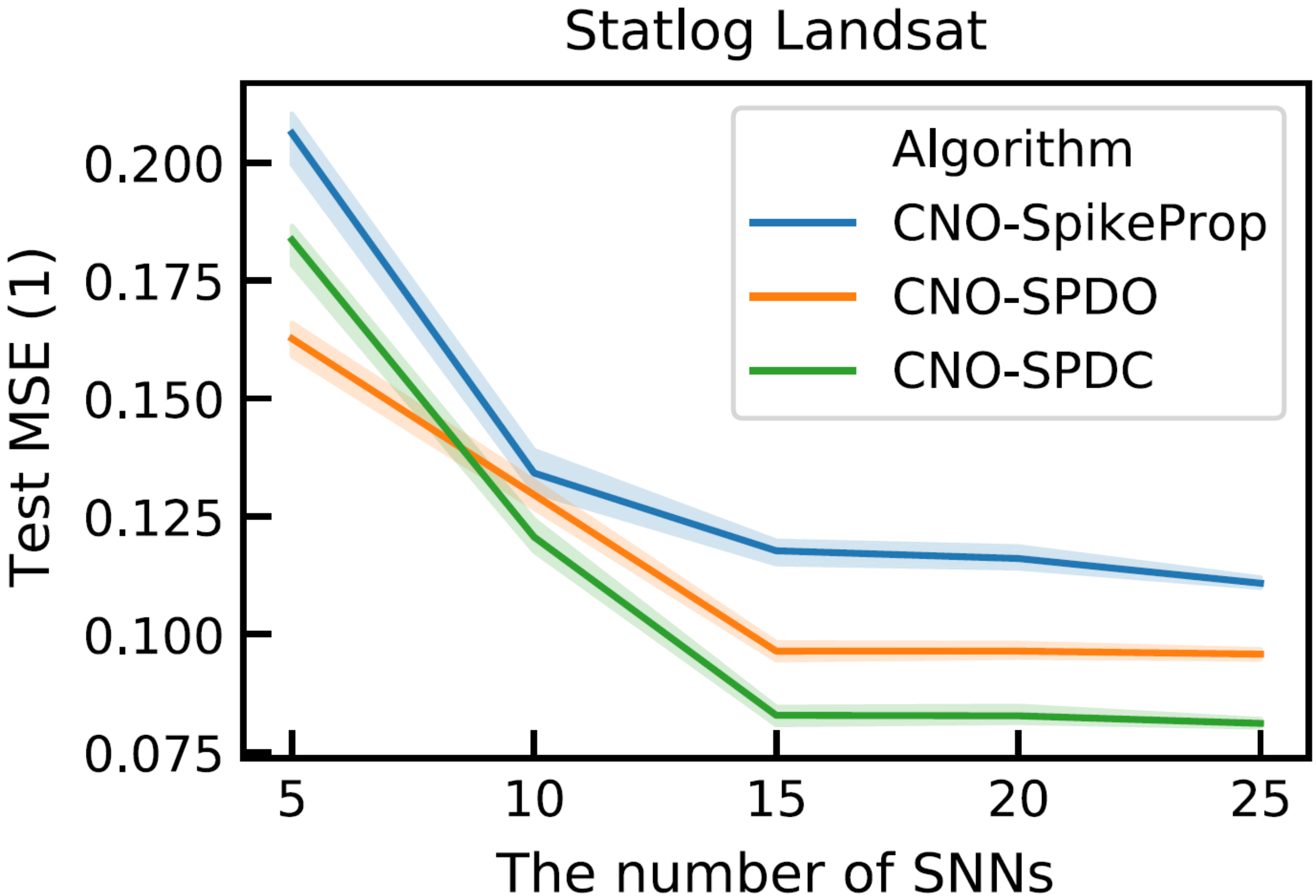
PDE Identification Results

PDE	Methods	nRMSE (mean+std)	SR	$\|\widehat{w}\|_0$
Klein-Gordon equation	PNN-SBL	1.3599 ± 0.6038	6%	5.32
	S-ESBL [26]	1.3493 ± 0.5020	0%	9.31
	EM-SBL [9]	0.0135 ± 0.0000	0%	6.00
	IF-SBL [23]	1.4998 ± 0.5794	0%	3.60
	LP-SBL [15]	0.0023 ± 0.0073	96%	3.09
	GGAMP-SBL [24]	0.0310 ± 0.1239	0%	4.00
	CNO-SBL (N=10)	$\mathbf{0.0020 \pm 0.0000}$	**100%**	**3.00**
Fisher's equation	PNN-SBL	0.7428 ± 0.5670	23%	3.21
	S-ESBL [26]	1.0713 ± 0.2375	0%	9.12
	EM-SBL [9]	0.0236 ± 0.0000	0%	7.00
	IF-SBL [23]	1.1470 ± 0.3709	0%	5.81
	LP-SBL [15]	0.6924 ± 0.3553	21%	**3.00**
	GGAMP-SBL [24]	0.6898 ± 0.3644	0%	4.79
	CNO-SBL (N=5)	$\mathbf{0.0010 \pm 0.0000}$	**100%**	**3.00**
Sine-Gordon equation	PNN-SBL	0.8953 ± 0.3671	10%	2.91
	S-ESBL [26]	0.8989 ± 0.4050	12%	2.97
	EM-SBL [9]	0.0034 ± 0.0000	**100%**	**2.00**
	IF-SBL [23]	0.8456 ± 0.3854	8%	3.23
	LP-SBL [15]	0.5741 ± 0.4369	1%	2.61
	GGAMP-SBL [24]	0.3709 ± 0.4956	3%	3.35
	CNO-SBL (N=5)	$\mathbf{0.0011 \pm 0.0000}$	**100%**	**2.00**

Success Rate vs. Population Size



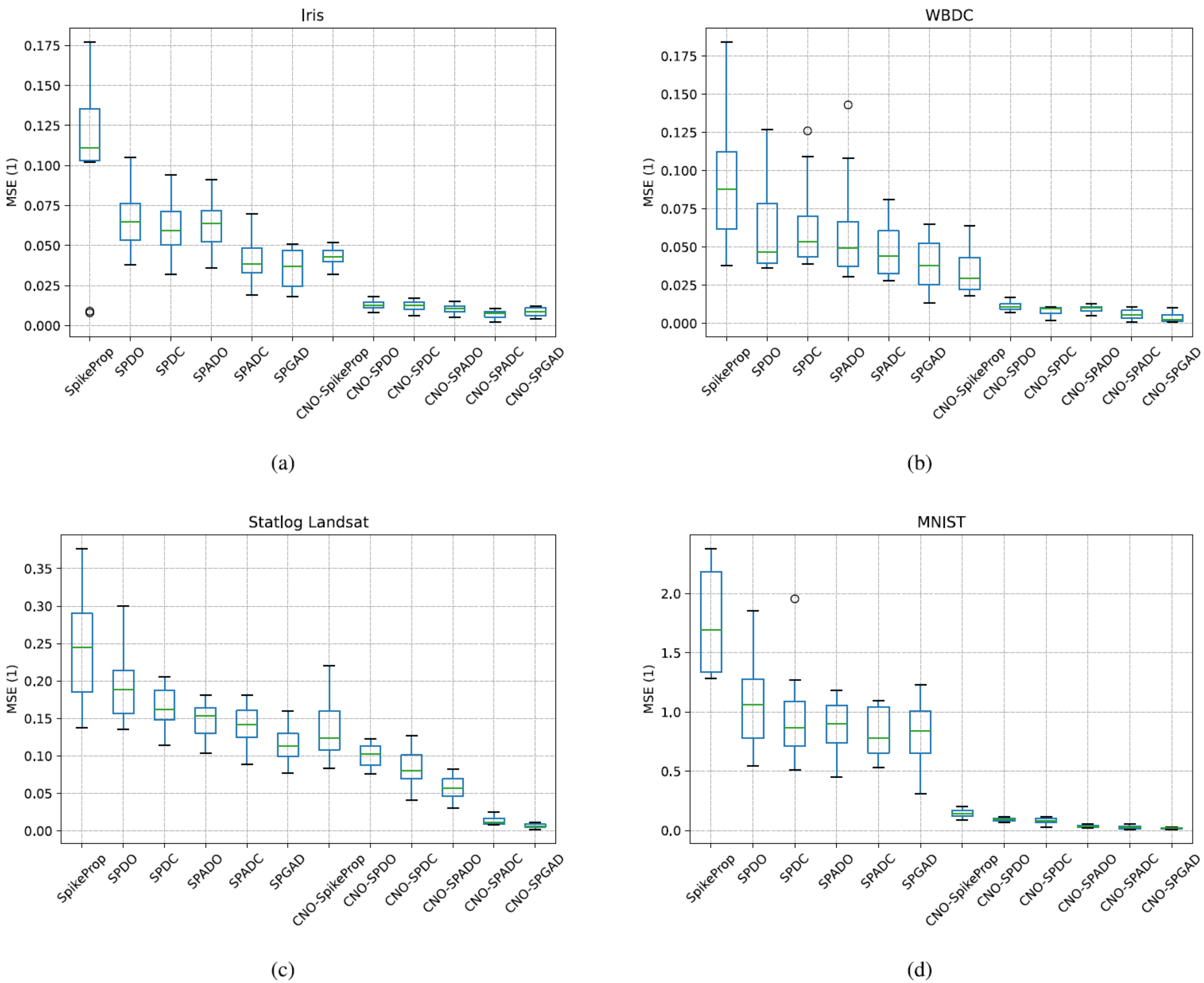
Spiking NN Supervised Learning



Average Training Errors

Algorithm \ Dataset	Iris	WBCD	Statlog Landsat	MNIST
SpikeProp [11]	0.102	0.057	0.205	1.562
SpikePropRT [36]	0.080	0.012	3.921	-
SpikePropAD [18]	0.010	0.421	3.203	-
SPSL1/2 [34]	0.023	0.035	-	-
MC-SEFRON [37]	0.033	0.017	-	0.086
TMM [38]	0.018	0.028	-	-
SPDO	0.058	0.040	0.192	0.966
SPDC	0.049	0.043	0.152	0.848
P-SPDO	0.108	0.069	0.234	1.487
P-SPDC	0.094	0.066	0.226	1.375
SPADO	0.053	0.035	0.138	0.843
SPADC	0.034	0.032	0.123	0.621
SPGAD	0.032	0.025	0.112	0.635
CNO-SpikeProp	0.042	0.023	0.119	0.123
CNO-SPDO	0.012	0.009	0.081	0.082
CNO-SPDC	0.013	0.004	0.098	0.079
CNO-SPADO	0.010	0.006	0.064	0.030
CNO-SPADC	**0.007**	0.003	0.018	0.038
CNO-SPGAD	0.009	**0.002**	**0.008**	**0.024**

Monte Carlo Results



Training and Testing Accuracy

Classification Performances of the 16 SNN Algorithms in Terms of Micro Averaged Accuracy With Training/Test Samples in the Four Benchmark Data Sets

Algorithm \ Dataset	Iris	WBCD	Statlog Landsat	MNIST
SpikeProp [11]	97.2±0.15 /96.7±0.13	97.3±0.17/94.2±0.20	85.8±0.16/80.3±0.18	-
MC-SEFRON [37]	98.4/97.1	98.4/97.4	-	93.64/92.3
TMM [38]	97.5±0.8/97.2±1.0	97.4±0.3/97.2±0.5	-	-
SPDO	98.5±0.10/97.8±0.12	98.2±0.09/96.6±0.13	88.4±0.08/85.3±0.06	94.5±0.11/92.3±0.10
SPDC	98.7±0.08/98.0±0.06	98.6±0.12/97.1±0.14	88.9±0.08/85.9±0.07	95.2±0.09/93.6±0.06
P-SPDO	96.8±0.14/94.3±0.16	96.4±0.15/95.2±0.14	87.7±0.19/84.3±0.16	92.8±0.13/90.7±0.11
P-SPDC	97.0±0.09/95.8±0.13	96.8±0.14/95.4±0.12	87.2±0.10/84.8±0.09	91.9±0.08/90.2±0.08
SPADO	99.2± 0.08/98.3±0.10	99.6±0.07/98.2±0.06	89.3±0.06/88.4±0.05	96.6±0.06/95.7±0.04
SPADC	99.5±0.05/98.3±0.03	99.8±0.06/98.5±0.08	91.2±0.06/89.1±0.04	97.7±0.07/96.3±0.05
SPGAD	99.8±0.03/99.3±0.02	99.8±0.05/99.1±0.02	93.7±0.04/92.5±0.03	98.8±0.05/97.8±0.04
CNO-SpikeProp	98.4±0.03/96.6±0.02	99.1±0.01/98.4±0.02	88.2±0.03/86.3±0.03	95.4±0.02/93.3±0.01
CNO-SPDO	99.8±0.02/98.7±0.01	99.6±0.02/98.4±0.02	95.6±0.01/94.3±0.02	98.7±0.02/98.2±0.01
CNO-SPDC	99.8±0.01/99.2±0.02	99.6±0.01/98.7±0.02	95.8±0.03/95.1±0.02	98.9±0.01/98.5±0.01
CNO-SPADO	**100±0.00/100±0.00**	**100±0.00/100±0.00**	98.6±0.02/97.8±0.03	99.2±0.01/98.4±0.01
CNO-SPADC	**100±0.00/100±0.00**	**100±0.00/100±0.00**	98.4±0.02/97.5±0.01	98.9±0.01/98.2±0.01
CNO-SPGAD	**100±0.00/100±0.00**	**100±0.00/100±0.00**	**98.8±0.01/98.2±0.00**	**99.4±0.02/98.9±0.01**

J. Zhao, J. Yang, J. Wang, and W. Wu, "Spiking neural network regularization with fixed and adaptive drop-keep probabilities," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, pp. 4096-4109, 2022.

Classification Performance

AVERAGE PRECISION, RECALL, AND F1-SCORE OF THE 14 SNN ALGORITHMS WITH TEST SAMPLES IN THE FOUR BENCHMARK DATA SETS

Algorithm \ Measure	Precision	Recall	F1-score
SpikeProp [11]	0.75	0.77	0.73
SPDO	0.82	0.78	0.78
SPDC	0.84	0.81	0.80
P-SPDO	0.75	0.69	0.70
P-SPDC	0.79	0.76	0.75
SPADO	0.85	0.86	0.83
SPADC	0.86	0.84	0.83
SPGAD	0.86	0.86	0.85
CNO-SpikeProp	0.79	0.80	0.76
CNO-SPDO	0.89	0.82	0.84
CNO-SPDC	0.91	0.84	0.86
CNO-SPADO	0.86	0.89	0.88
CNO-SPADC	0.87	**0.90**	0.88
CNO-SPGAD	**0.93**	0.89	**0.91**

Portfolio Optimization

- As a cornerstone of modern finance involved with the work of at least six Nobel Laureates, portfolio selection or optimization is of great interest for financial investments from both academic and economic points of view.
- A major breakthrough of modern portfolio theory was highlighted by Nobel Laureate H. M. Markowitz in his mean-variance framework.

Mean-Variance Framework

A classical mean-variance portfolio optimization problem can be formulated as a bi-objective optimization problem [1]:

$$\begin{aligned} \min_{y} \; & y^T V y, -\mu^T y \\ \text{s.t. } & e^T y = 1, \\ & 0 \le y \le e, \end{aligned} \tag{1}$$

where $y = (y_1, y_2, \dots, y_n)^T$ is the proportion of wealth invested to the stocks, V is the covariance matrix, $\mu = (\mu_1, \mu_2, \dots, \mu_n)^T$ is the vector of mean returns of n stocks, e is the vector of ones. The terms $\mu^T y$ and $y^T V y$ measure the expected return and risk of a portfolio, respectively.

Mean-CVaR Formulation

$$\min_{y,\alpha,\rho} \ \alpha + \frac{1}{N(1-\theta)} \sum_{j=1}^{N} \rho_j$$

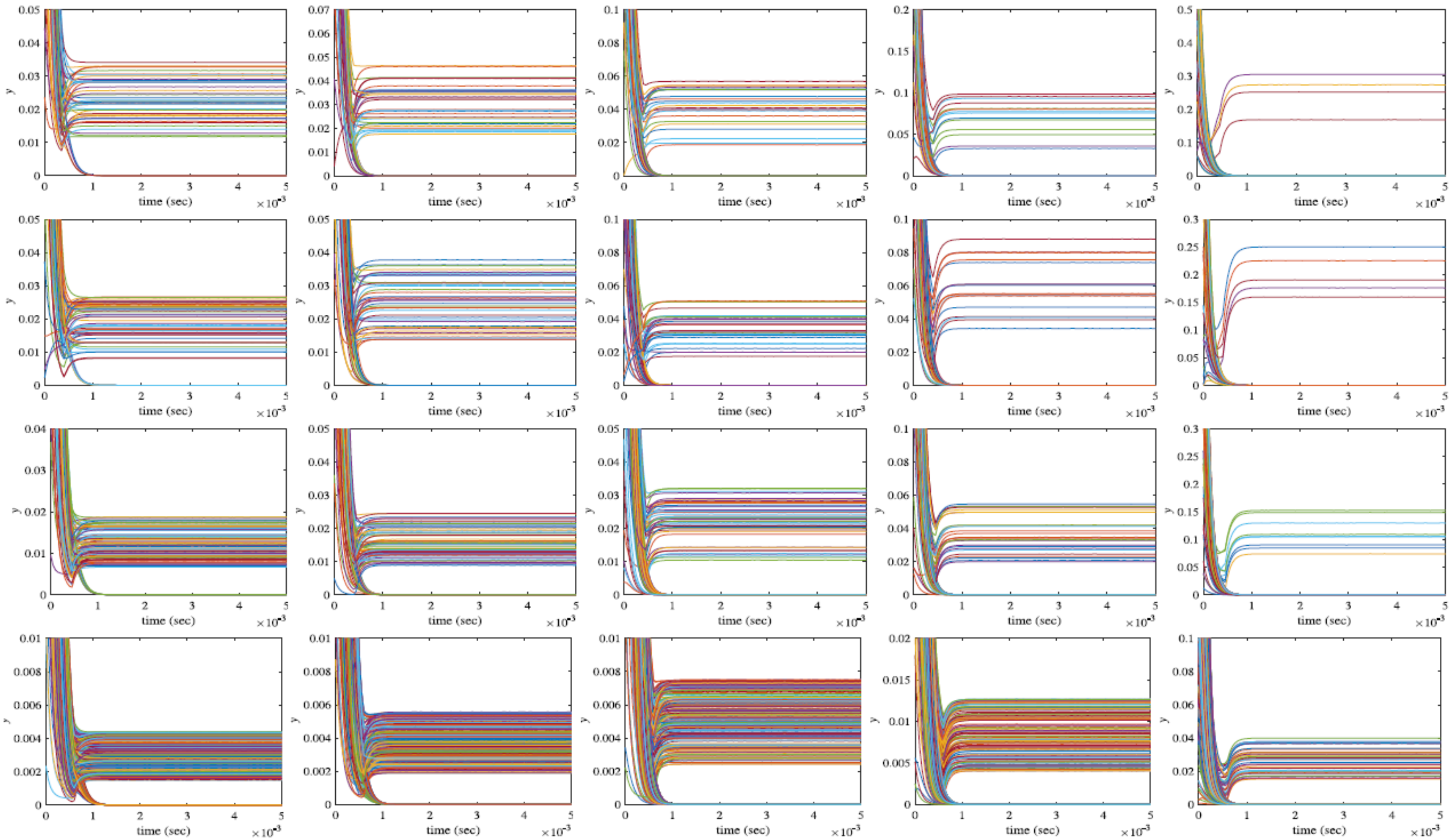
$$\begin{aligned} \min_{y,\alpha,\rho} \ & -\mu^T y, \\ \text{s.t.} \ & \rho_j \geq y^T \xi_j - \alpha, \ \ j = 1, \ldots, N \\ & e^T y = 1, \\ & 0 \leq y \leq e. \end{aligned}$$

Cardinality-constrained Mean-CVaR Formulation

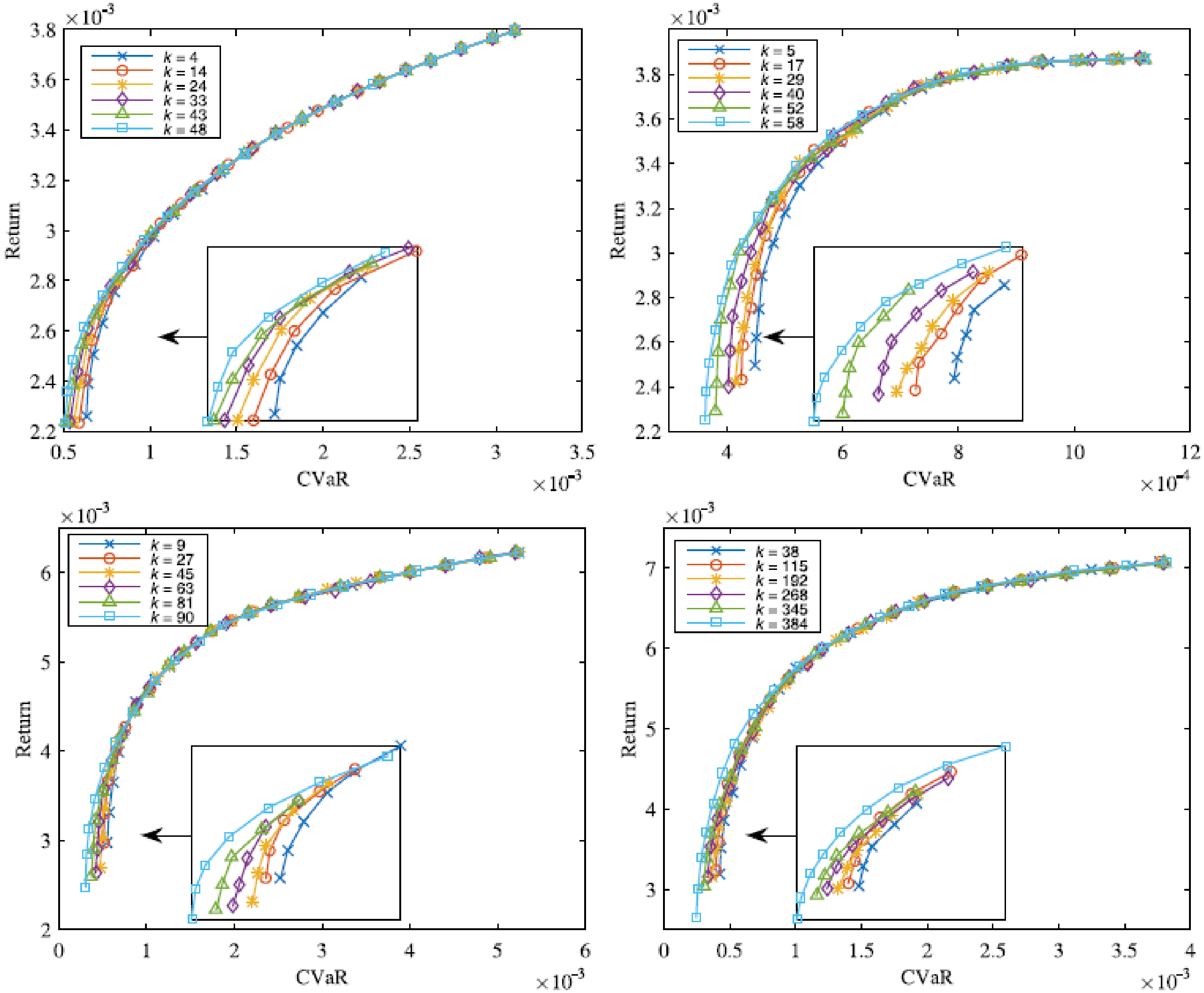
$$\min_{y,z,\alpha,\rho} \ \alpha + \frac{1}{N(1-\theta)} \sum_{j=1}^{N} \rho_j$$

$$\begin{aligned}
\min_{y,z,\alpha,\rho} \quad & -\mu^T y \\
\text{s.t.} \quad & \xi_j^T y - \alpha \le \rho_j, \ \ j = 1, \ldots, N \\
& e^T y = 1, \\
& e^T z \le k, \\
& 0 \le y \le e \circ z, \\
& z \in \{0,1\}^n,
\end{aligned}$$

Neuronal Convergence



Pareto Fronts



Comparative Results

Statistics of experimental results of six cardinality-constrained bi-objective portfolio optimization methods CNO-CBPS, MINLP, CPLEX, NSGAII, PBIEDE, and HMA, where $\lambda = 0.5$, best results are boldfaced, and 'NA' denotes the baseline could not obtain a solution within the time limit.

Datasets	n	k	MINLP solver		CPLEX		NSGAII		PBIEDE		HMA		CNO-CBPS	
			time (s)	f_λ	time (s)	f_λ	time (s)	f_λ	time (s)	f_λ	time (s)	f_λ	time (s)	f_λ
HDAX	48	15	**2.46**	**0.5811**	4.31	**0.5811**	5.36	0.5921	11.30	0.5910	23.96	0.5822	8.98	**0.5811**
		10	4.99	**0.5932**	4.75	**0.5932**	**3.78**	0.6011	4.51	0.5986	21.70	0.5956	6.37	**0.5932**
		5	21.88	**0.6186**	5.38	**0.6186**	**2.19**	0.6230	3.21	0.6201	19.93	0.6234	3.71	**0.6186**
FTSE	58	15	9.22	**0.3754**	**6.26**	**0.3754**	8.77	0.3825	16.26	0.3802	30.13	0.3779	17.02	0.3770
		10	14.14	**0.3816**	7.63	**0.3816**	**5.26**	0.3930	10.36	0.3897	27.79	0.3898	8.69	**0.3816**
		5	89.46	**0.3962**	9.71	**0.3962**	**4.10**	0.4098	8.49	0.3991	26.99	0.3966	6.53	**0.3962**
HSI	90	15	82.33	**1.6842**	**7.80**	**1.6842**	9.39	1.7052	20.61	1.6944	45.05	1.6861	14.13	1.6856
		10	189.19	**1.7120**	11.12	**1.7120**	**8.42**	1.7749	14.12	1.7321	39.87	1.7163	12.01	**1.7120**
		5	310.11	**1.7486**	27.83	**1.7486**	**5.53**	1.7853	8.99	1.7657	34.92	1.7646	7.81	**1.7486**
SP500	384	15	3600	NA	2408	**1.2502**	**439.87**	1.3021	779.4	1.2841	1113.56	1.2751	623.80	**1.2502**
		10	3600	NA	3600	NA	**248.41**	1.3278	361.32	1.3030	1077.66	1.2945	458.65	**1.2896**
		5	3600	NA	3600	NA	**148.72**	1.3348	210.14	1.3228	932.17	1.3158	245.36	**1.3107**

M.F. Leung and **J. Wang**, "Cardinality-constrained portfolio selection based on collaborative neurodynamic optimization," *Neural Networks,* vol. 145, pp. 68-79, 2022.

Annualized Returns

Ranges of annualized returns of four resulting portfolios with various values of cardinality bound k.

	Datasets	Market index	Solution sparsity					
			0%	~10%	~30%	~50%	~70%	~90%
In-sample 2000–2011	HDAX	5.66%	10.18–15.42%	8.70–15.42%	8.18–15.42%	7.39–15.42%	7.11–15.42%	6.97–15.42%
	FTSE	5.29%	11.22–15.90%	10.16–15.90%	9.01–15.90%	7.95–15.90%	7.35–15.90%	6.99–15.90%
	HSI	6.77%	13.68–21.02%	12.56–21.02%	12.15–21.02%	11.80–21.02%	11.36–21.02%	10.09–21.02%
	SP500	5.31%	11.98–17.88%	10.28–17.88%	10.02–17.88%	9.39–17.88%	8.68–17.88%	8.25–17.88%
Out-of-sample 2012–2017	HDAX	6.64%	10.32–14.03%	9.91–14.37%	9.37–14.32%	8.61–14.59%	7.50–14.87%	6.60–14.11%
	FTSE	4.43%	11.51–15.04%	11.17–14.72%	10.61–15.02%	9.93–14.93%	9.72–14.88%	9.33–15.40%
	HSI	6.18%	12.05–18.95%	11.71–18.98%	11.47–18.90%	11.30–18.83%	11.16–18.73%	10.28–18.92%
	SP500	5.93%	13.57–26.84%	13.38–27.41%	12.47–26.20%	12.08–26.54%	11.30–26.92%	11.25–25.32%

M.F. Leung and **J. Wang**, "Cardinality-constrained portfolio selection based on collaborative neurodynamic optimization," *Neural Networks,* vol. 145, pp. 68-79, 2022.

Problem Reformulation

Maximizing conditional Sharpe ratio subject to self-financing and cardinality constraints.

$$\max_{y} \ \frac{\mu^T y - r_f}{\mathrm{CVaR}_\theta(y)}$$
$$\text{s.t. } e^T y = 1,$$
$$\|y\|_0 \leq k,$$
$$y \geq 0.$$

M. Leung, **J. Wang**, and H. Che, "Cardinality-constrained portfolio selection via two-timescale duplex neurodynamic optimization," *Neural Networks*, vol. 153, pp. 399-410, 2022.

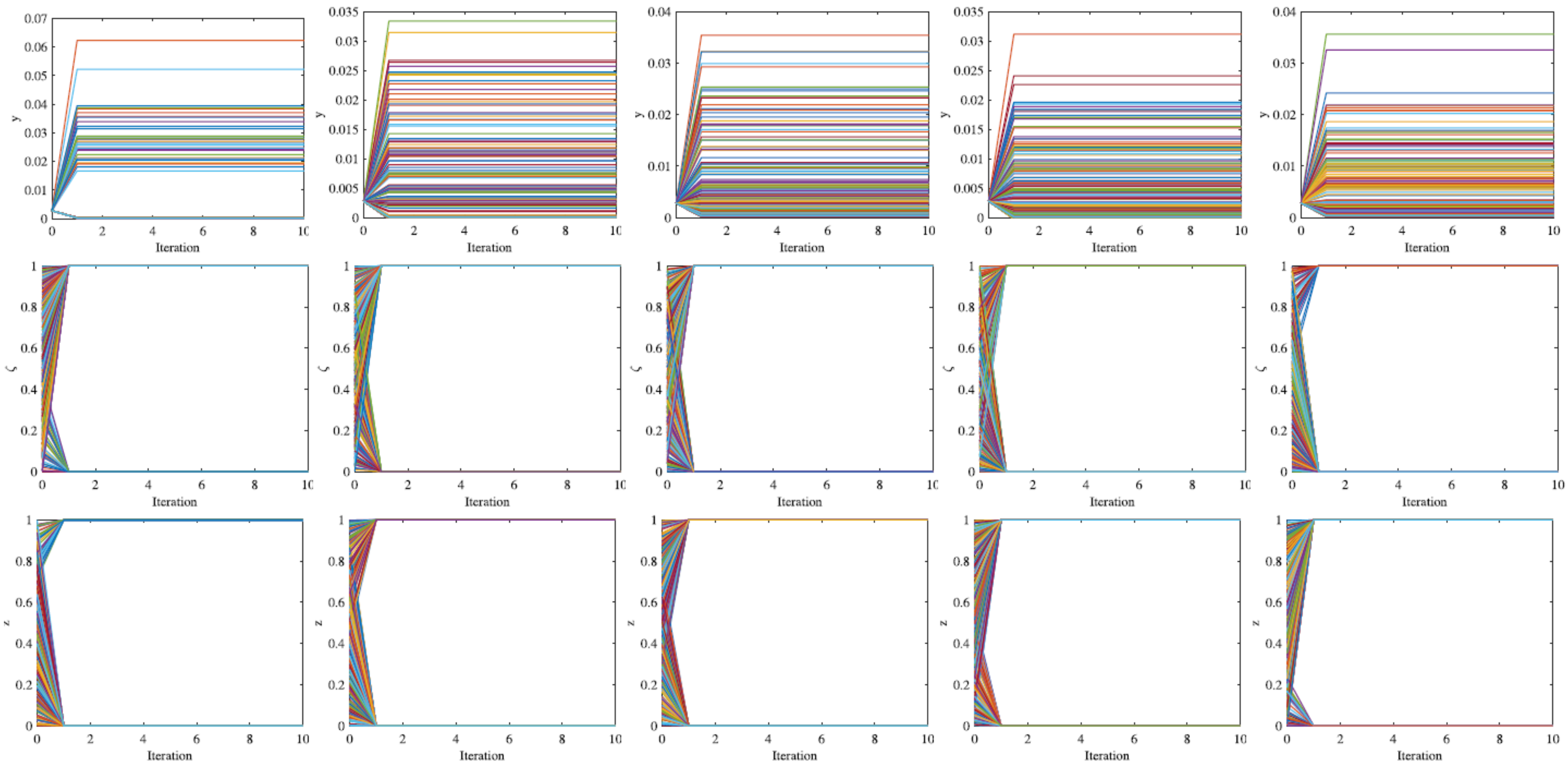
Problem Reformulation (cont'd)

$$\min_{\gamma,\rho,\sigma,y,z} \frac{\gamma^2}{2}\left(\rho + \frac{1}{N(1-\theta)}\sum_{J=1}^{N}\sigma_J\right)^2 - \gamma(\mu^T y - r_f)$$

$$\begin{aligned} \text{s.t. } & \sigma_i \geq -\xi_i^T y - \rho, \ \ \sigma_i \geq 0, \ \ i = 1, 2, \ldots, N; \\ & e^T y = 1; \\ & e^T z \leq k; \\ & 0 \leq y \leq z; \\ & z \in \{0, 1\}^n; \end{aligned}$$

where $z \in \{0, 1\}^n$, $e^T z \leq k$ is the cardinality constraint.

Convergence Behaviors

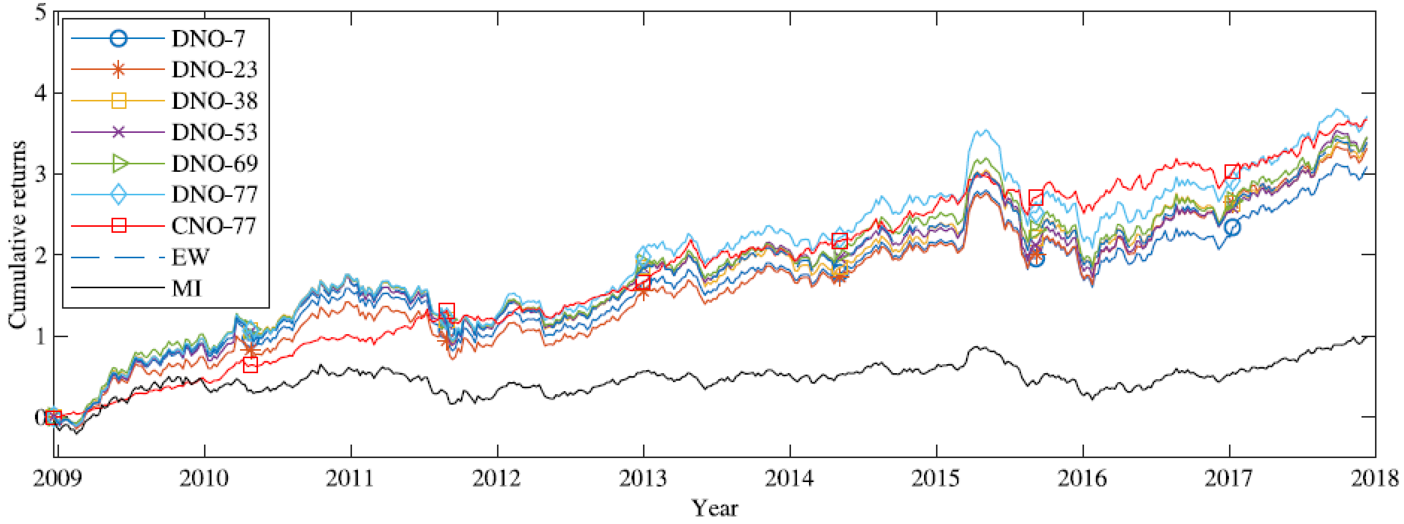


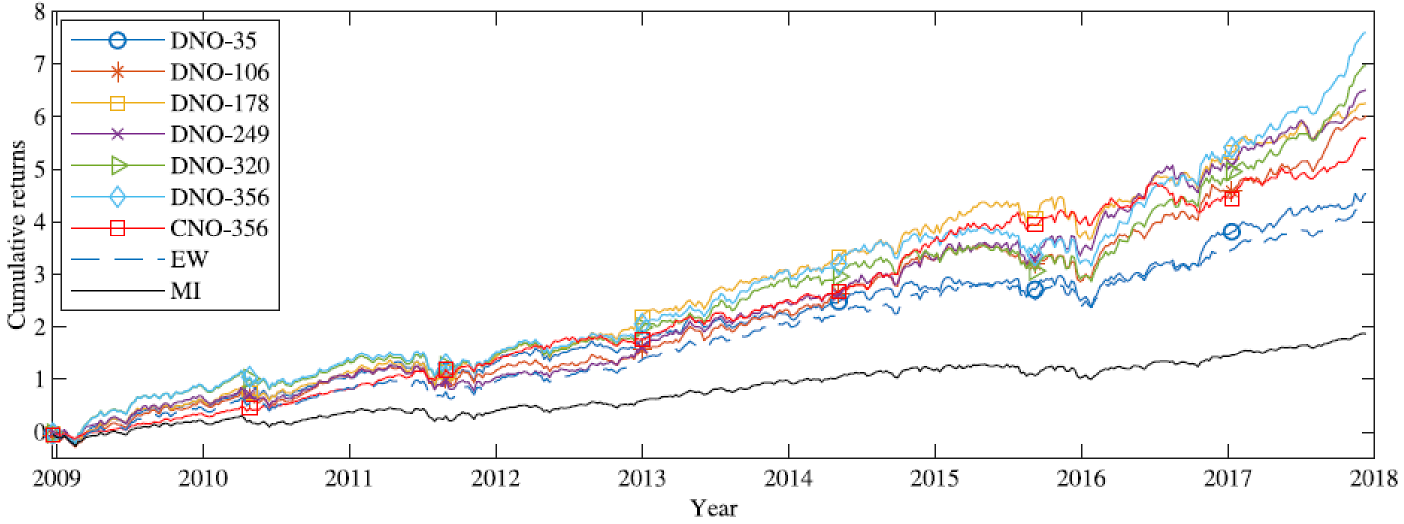
Sharpe Ratio, Conditional Sharpe Ratio, and Annualized Returns

Resulting annualized Sharpe ratio, conditional Sharpe ratio and returns based on half-and-half partitioned datasets.

Dataset	*n*	*k*	Sharpe ratio				Conditional Sharpe ratio				Annualized return (%)			
			DNO	CNO	EW	MI	DNO	CNO	EW	MI	DNO	CNO	EW	MI
HDAX	49	4	0.5486	0.5231	0.6873	0.6726	0.2475	0.2349	0.2734	0.3447	10.4758	7.9459	12.9857	12.2463
		14	0.5975	0.5689			0.2863	0.2487			11.2641	11.1892		
		24	0.6687	0.6357			0.2963	0.2532			12.7997	12.1576		
		34	0.7271	0.6848			0.3401	0.2688			13.8972	13.4195		
		44	0.7574	0.7119			0.3658	0.2899			14.5806	13.6594		
		49	0.7582	0.7565			0.3836	0.3129			15.1259	13.8340		
FTSE	56	5	0.4844	0.4349	0.8303	0.4364	0.2053	0.1936	0.3659	0.2374	7.7156	7.9548	12.9811	5.9705
		16	0.6338	0.6187			0.2963	0.2311			9.8699	8.2565		
		28	0.6600	0.6302			0.3051	0.2412			10.2688	8.4036		
		39	0.6755	0.6554			0.3159	0.2631			10.7203	9.1148		
		50	0.7214	0.7201			0.3562	0.2824			11.1810	12.4485		
		56	0.8756	0.8542			0.3965	0.3554			15.8881	15.6841		
HSCI	77	7	0.9315	0.8645	0.9791	0.4782	0.4353	0.4086	0.3459	0.1710	16.9461	16.9073	17.8610	7.9394
		23	0.9411	0.8790			0.4420	0.4185			17.6416	17.1799		
		38	0.9551	0.8941			0.4487	0.4214			17.8137	17.4809		
		53	0.9640	0.9074			0.4521	0.4267			18.0446	17.7833		
		69	1.9881	0.9186			0.4566	0.4332			18.0600	18.0823		
		77	1.0147	0.9369			0.4635	0.4471			18.7928	18.6832		
SP500	356	35	1.1314	1.1145	1.1599	0.8255	0.5019	0.4147	0.4120	0.2801	20.9448	19.7990	20.3834	12.4253
		106	1.1834	1.1440			0.5633	0.4417			24.1746	20.3223		
		178	1.2083	1.1748			0.5852	0.5189			24.6381	20.9244		
		249	1.2141	1.1904			0.6197	0.5435			25.1157	21.5870		
		320	1.2187	1.2126			0.6362	0.5679			25.9424	22.3755		
		356	1.2231	1.2175			0.6539	0.6488			26.9975	23.2873		

Cumulative Returns





Distributed Portfolio Selection

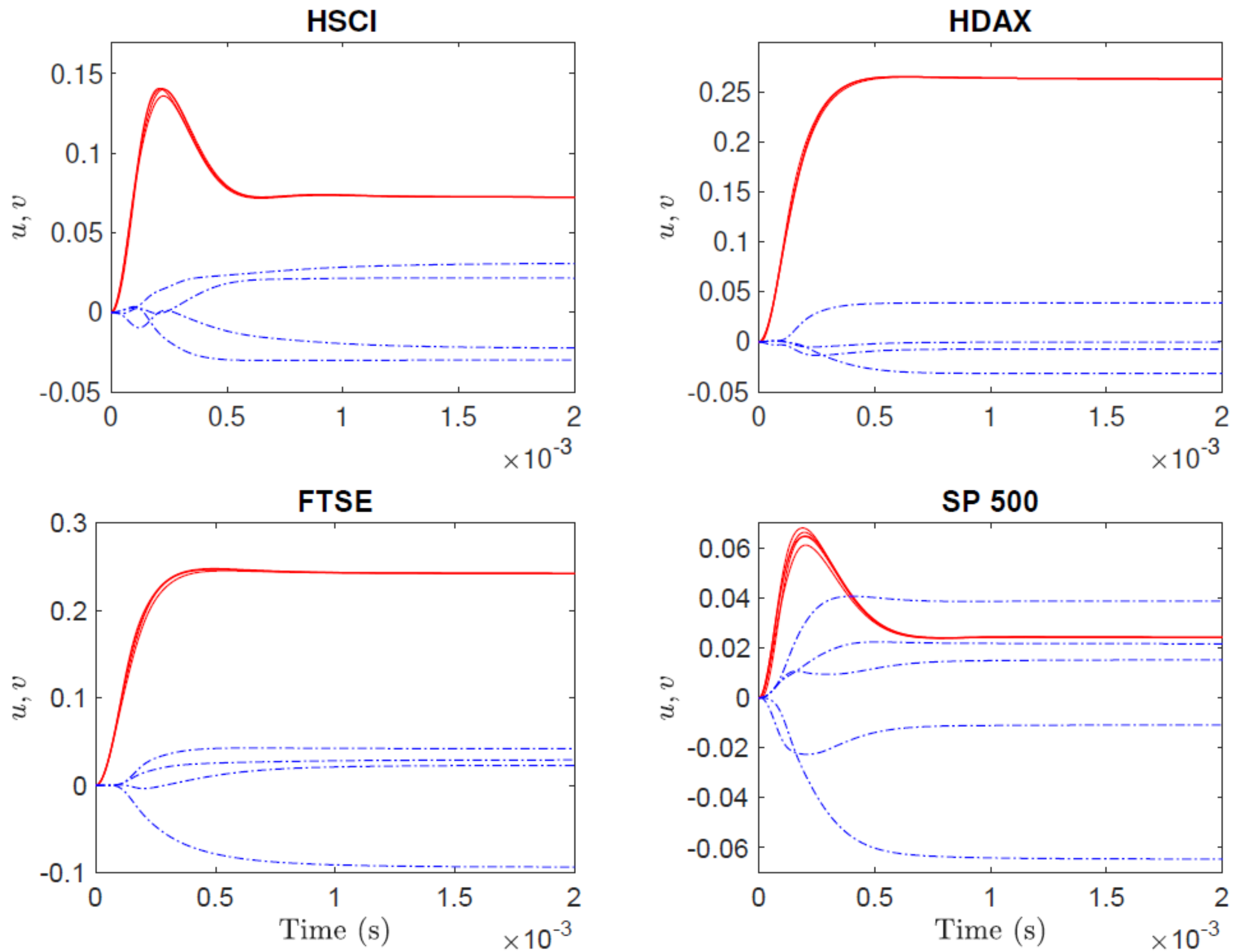
- As a paradigm of decentralized decision making in the finance industry, decentralized portfolio optimization is advantageous in terms of preferential and geographical diversification.
- In distributed portfolio optimization, more specialized investment decisions could be made by leveraging the specific expertise of fund managers toward higher returns and lower risks.

Coupled Projection Neural Networks

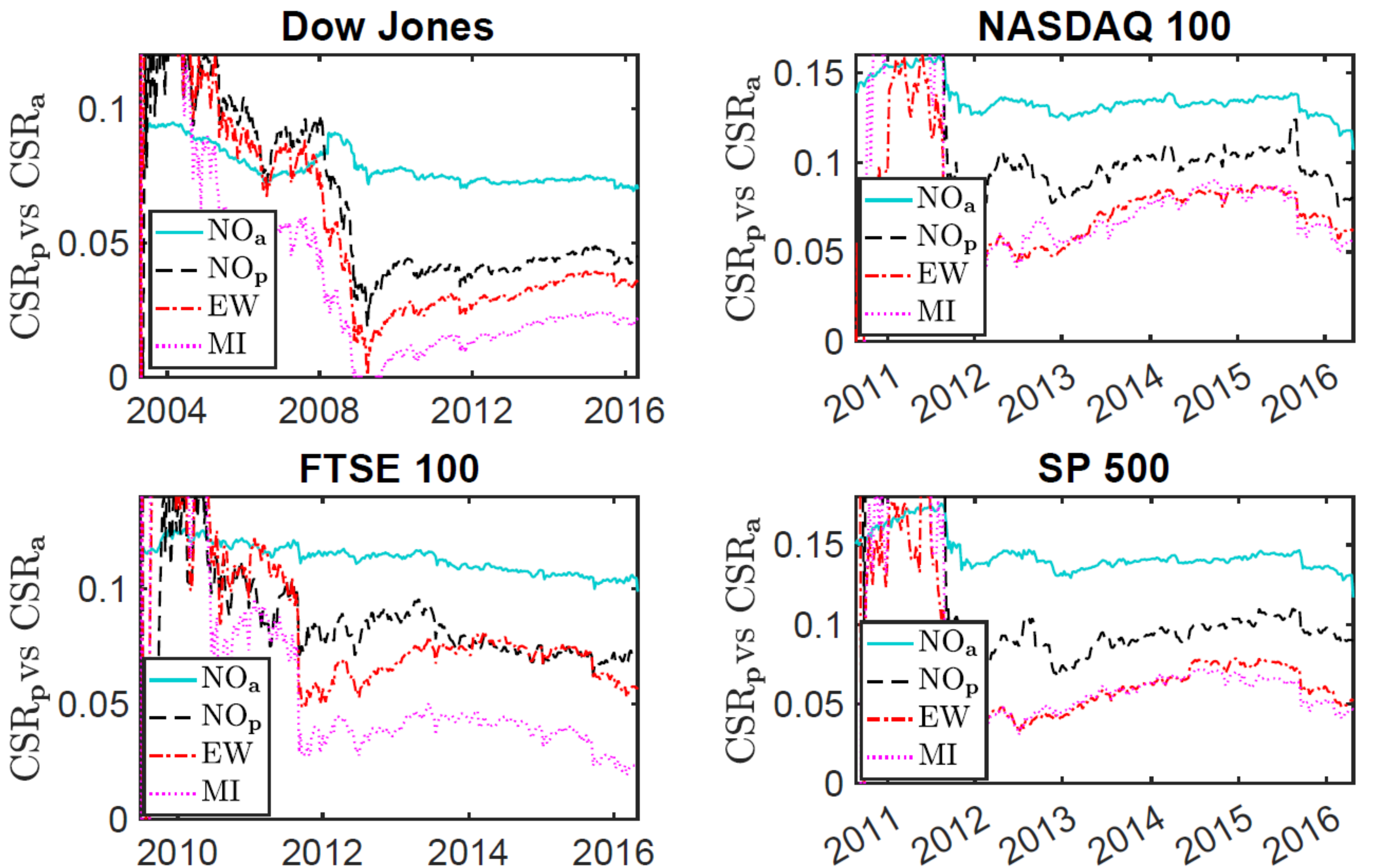
$$\epsilon \frac{dx_i}{dt} = -x_i + g(x_i - \nabla_{x_i} f_i(x_i, \rho) - e_{n_i} u_i),$$
$$\epsilon \frac{d\rho}{dt} = -\rho + g(\rho - \nabla_{\rho} f_i(x_i, \rho)),$$
$$\epsilon \frac{du_i}{dt} = e_{n_i}^T x_i - b_i - \sum_{j \in \mathcal{N}_i} (u_i - u_j + v_i - v_j),$$
$$\epsilon \frac{dv_i}{dt} = \sum_{j \in \mathcal{N}_i} (u_i - u_j),$$

J. Wang and X. Gan, "Neurodynamics-driven portfolio optimization with targeted performance criteria," *Neural Networks,* vol. 157, pp. 404-421, 2023.

Convergence/Consensus Behaviors



Ex-ante vs. Ex-post Performances



Performance Comparisons

Sortino Ratio (SoR_p)		
NO	EW	MI
1.7914	1.3315	1.4357
4.6473	1.5292	1.6859
3.2970	2.4649	2.0759
2.4611	**3.5893**	3.3459
1.5529	**1.5910**	0.7165
1.2181	0.9863	0.9612
1.3493	1.2127	0.6182
2.0778	1.8232	1.2355

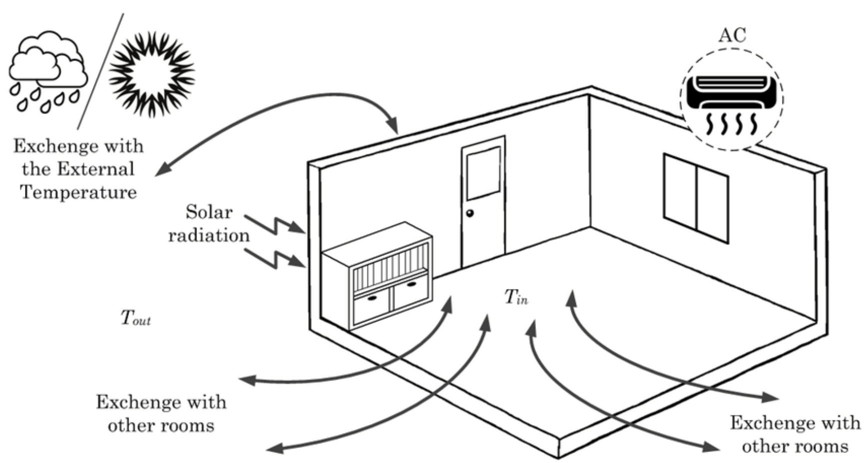
Cndtnl. Sharpe Ratio (CSR_p)		
NO	EW	MI
0.3732	0.4164	**0.4629**
0.9923	0.3291	0.3287
0.7442	0.5963	0.6525
0.3446	0.9860	**1.0737**
0.4685	0.3459	0.1710
0.3898	0.2734	0.3447
0.3975	0.3659	0.2374
0.6653	0.4120	0.2801

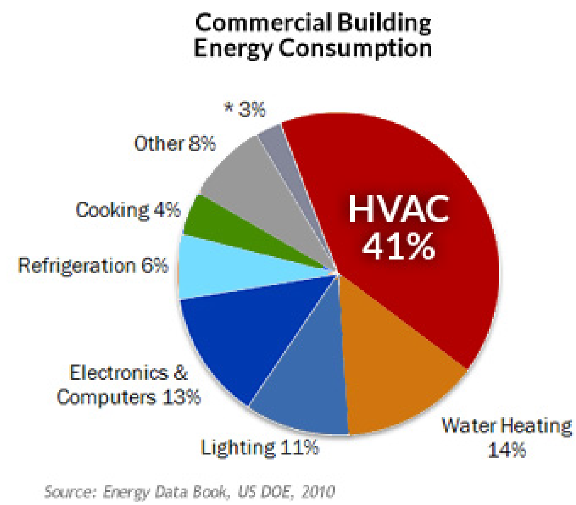
Performance Comparisons (cont'd)

BCST-Library	Dataset	NO	EW	MI	M-V	L-SSD	LR-ASSD	RMZ-SSD	PK-SSD	CZeSD
	DJIA	**0.7568**	0.6100	0.3704	0.5635	0.4596	0.5297	0.3261	0.6473	0.4895
Sharpe	NASDAQ100	**1.2844**	1.1131	0.9191	1.0528	1.3800	1.1686	1.1613	1.2323	1.0079
Ratio	FTSE100	1.1646	0.9152	0.3810	1.2832	1.2771	0.7767	**1.4363**	1.1980	0.7482
(SR_p)	S&P500	**1.3219**	0.8541	0.7672	0.8440	0.7966	0.8633	1.0451	0.8548	0.9550
	DJIA	**1.0516**	0.9200	0.7247	0.8453	0.6730	0.8070	0.4571	0.9415	0.7247
Sortino	NASDAQ100	1.9902	1.6369	1.4832	1.5542	**2.1006**	1.7874	1.7283	1.9192	1.4832
Ratio	FTSE100	1.7634	1.3124	1.0723	1.9765	1.9277	1.1118	**2.2624**	1.8296	1.0723
(SoR_p)	S&P500	**2.0476**	1.2543	1.4024	1.2017	1.1165	1.3083	1.5475	1.2917	1.4024
	DJIA	**2.4784**	1.9987	1.1252	1.7624	1.4112	1.7530	0.9501	1.9386	1.5490
Omega	NASDAQ100	4.6692	4.0148	3.0966	3.4222	**4.7902**	3.9412	3.9563	4.3455	3.4389
Ratio	FTSE100	3.6241	2.8986	1.1084	4.3032	4.4438	2.4450	**5.1034**	4.2804	2.3341
(OR_p)	S&P500	**4.9801**	3.0322	2.6243	2.6638	2.5327	2.8671	3.3727	2.7897	3.3734
Conditional	DJIA	**0.3264**	0.2601	0.1578	0.2440	0.1896	0.2211	0.1325	0.2673	0.2044
Sharpe	NASDAQ100	0.5763	0.4462	0.4037	0.4252	**0.5835**	0.4781	0.4560	0.5274	0.3982
Ratio	FTSE100	0.5157	0.4114	0.1662	0.5563	0.5582	0.2949	**0.6185**	0.5325	0.3346
(CSR_p)	S&P500	**0.6440**	0.3738	0.3334	0.3403	0.3134	0.3790	0.4247	0.3610	0.3835
Entropic	DJIA	0.2190	**0.2191**	0.1787	0.1703	0.1477	0.1965	0.0845	0.1953	0.1787
Sharpe	NASDAQ100	**0.6177**	0.3384	0.3319	0.2232	0.2635	0.2346	0.3357	0.3948	0.3319
Ratio	FTSE100	**0.5187**	0.2976	0.2370	0.2581	0.2065	0.1901	0.3937	0.2796	0.2370
(ESR_p)	S&P500	**0.6071**	0.2237	0.2806	0.1353	0.1600	0.1905	0.2167	0.2309	0.2806

HVAC operation optimization

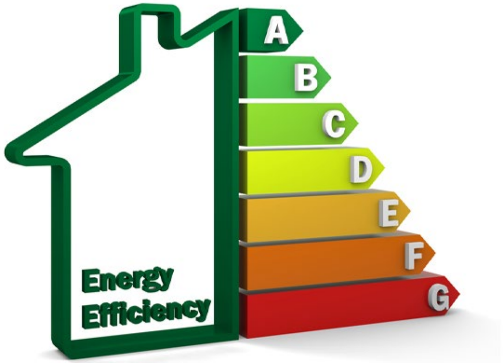
- Heating, ventilation, and air conditioning (HVAC) systems
  - Vital facilities for regulating temperature and humidity in the ambient environments of buildings
  - To meet thermal comfort and air quality requirements
- HVAC systems consume a substantial amount (~40%) of energy in commercial buildings





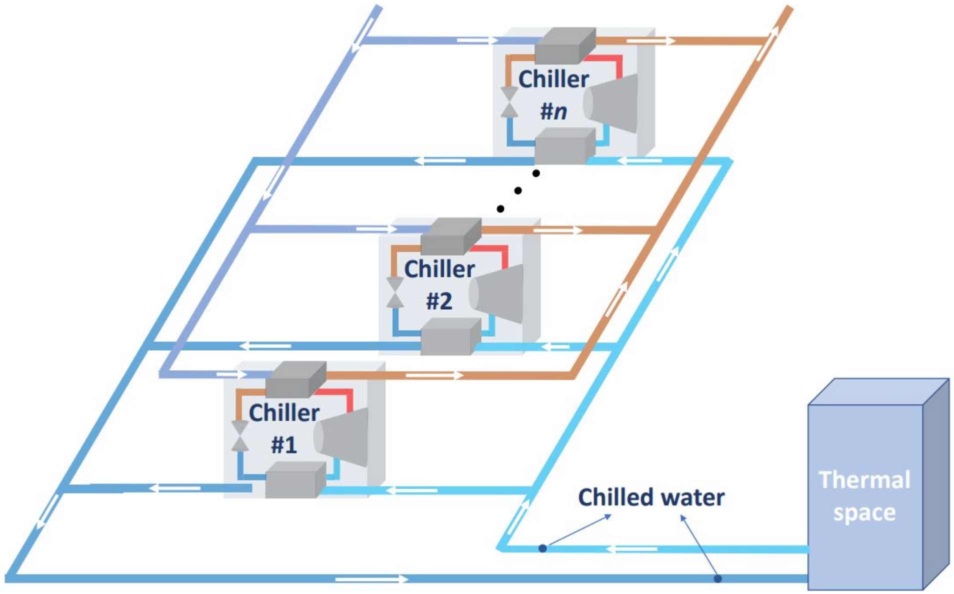
HVAC operation optimization (cont'd)

- In the global urbanization process, HVAC systems will take up an increasing portion of energy consumption



- It is crucial to optimize HVAC operations
  - Increase energy efficiency
  - Reduce carbon dioxide emissions

Optimal chiller loading



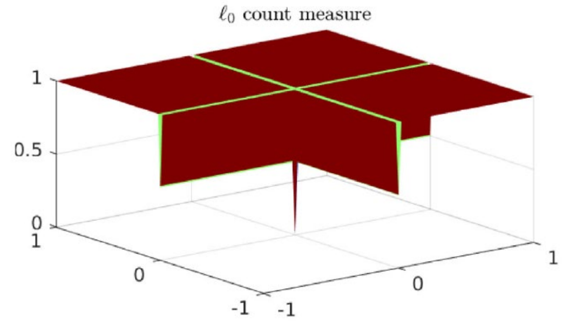
- Chillers: > **60%** of energy consumption
- Optimize chiller outputs
- To meet demands
- With minimal energy consumption

Optimal chiller loading (cont'd)

A **power consumption** function of chillers: for chiller i (i = 1, 2, . . . , n),

$$P_i(PLR_i) = a_i PLR_i^3 + b_i PLR_i^2 + c_i PLR_i + d_i$$

Cardinality constraint: to confine the number of chillers with ON status $\|PLR\|_0 \leq k$



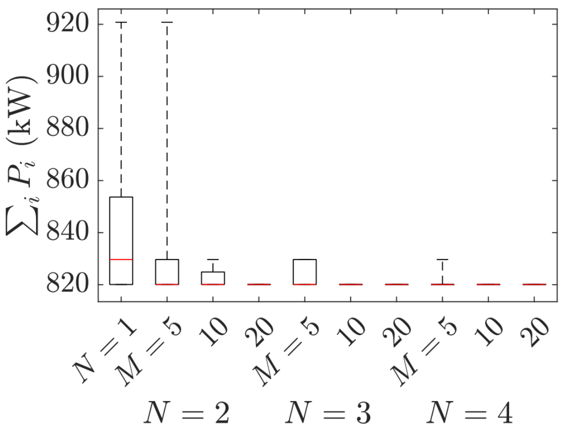
$$\sum_{i=1}^{n} y_i \leq k, \in \{0, 1\}$$

$\min\limits_{PLR,y} \sum_{i=1}^{n} P_i(PLR_i)$	**Power consumption**
subject to $\sum_{i=1}^{n} \overline{P}_i PLR_i - P_D = 0,$	**Supply-demand constraint**
$\underline{PLR}_i y_i \leq PLR_i \leq \overline{PLR}_i y_i,$	**Capacity constraints**
$\sum_{i=1}^{n} y_i \leq k,\ y_i(y_i - 1) = 0 \quad \forall i = 1, \ldots, n.$	**Cardinality constraint** / **Integer constraints**

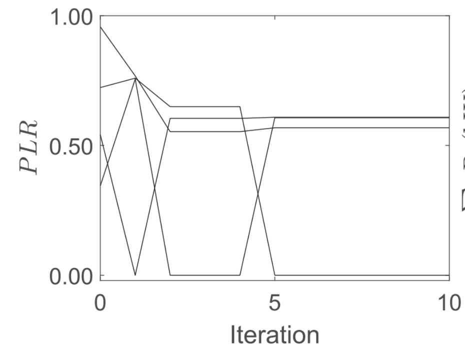
Optimal chiller loading (cont'd)

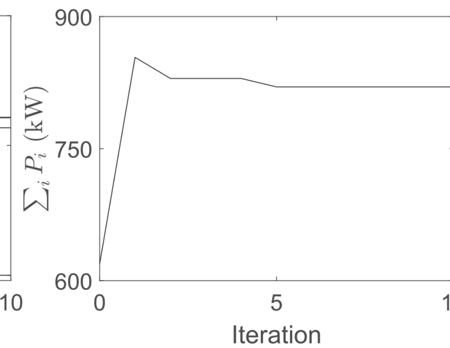
Hyperparameter Selection

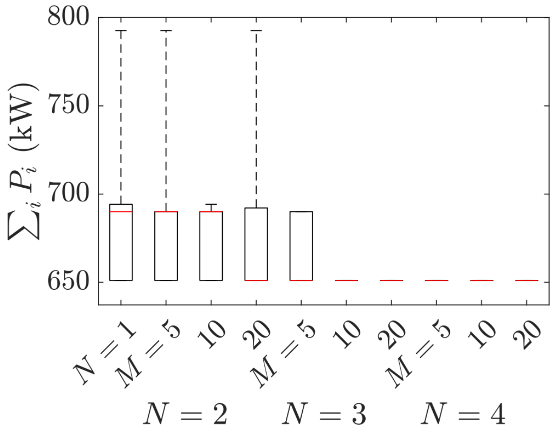
Convergence Behaviors



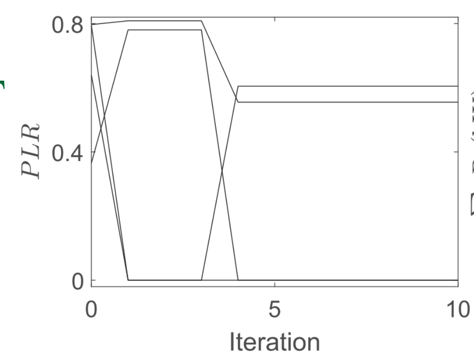
$P_D = 1450$ RT
$k = 3$

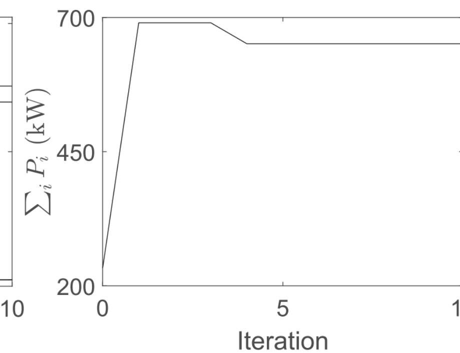






$P_D = 1160$ RT
$k = 2$





Optimal chiller loading (cont'd)

A four-chiller system (a hotel)

up to 23.85% of savings

Method	P_D=2610 RT best	worst	mean	SD	P_D=2320 RT best	worst	mean	SD	P_D=2030 RT best	worst	mean	SD
GAMS [11]	**1857.30**	-	-	-	**1455.66**	-	-	-	**1178.14**	-	-	-
GA [10]	1862.18	-	-	-	1457.23	-	-	-	1183.80	-	-	-
LGM [10]	**1857.30**	1864.17	-	-	**1455.66**	1461.05	-	-	**1178.14**	1182.50	-	-
PSO [13]	**1857.30**	1857.45	1857.43	0.04	**1455.66**	1522.42	1462.34	20.03	**1178.14**	**1178.14**	**1178.14**	**0.00**
DE [14]	**1857.30**	1858.57	1857.43	0.40	**1455.66**	**1455.66**	**1455.66**	**0.00**	**1178.14**	**1178.14**	**1178.14**	**0.00**
DCSA [16]	**1857.30**	1857.40	1857.32	0.02	1455.67	1458.48	1455.81	0.53	**1178.14**	1199.50	1181.07	4.80
CNO-CL	**1857.30**	**1857.30**	**1857.30**	**0.00**	**1455.66**	**1455.66**	**1455.66**	**0.00**	**1178.14**	**1178.14**	**1178.14**	**0.00**

Method	P_D=1740 RT best	worst	mean	SD	P_D=1450 RT best	worst	mean	SD	P_D=1160 RT best	worst	mean	SD
GAMS [11]	**998.53**	-	-	-	**820.07**	-	-	-	**651.07**	-	-	-
GA [10]	1001.62	-	-	-	907.72	-	-	-	856.30	-	-	-
LGM [10]	**998.53**	1002.22	-	-	904.62	907.97	-	-	849.99	853.13	-	-
PSO [13]	**998.53**	1013.43	1005.36	5.71	**820.07**	847.53	826.52	10.88	**651.07**	691.19	667.12	19.65
DE [14]	**998.53**	1009.20	1000.21	3.66	**820.07**	821.28	820.19	0.38	**651.07**	655.63	651.53	1.44
DCSA [16]	1008.24*	1074.55*	1038.13*	25.72	825.72*	897.06*	838.05*	17.43	652.16*	794.25*	713.17*	44.02
CNO-CL	**998.53**	**998.53**	**998.53**	**0.00**	**820.07**	**820.07**	**820.07**	**0.00**	**651.07**	**651.07**	**651.07**	**0.00**

A 20-chiller system → up to 21.36% of savings

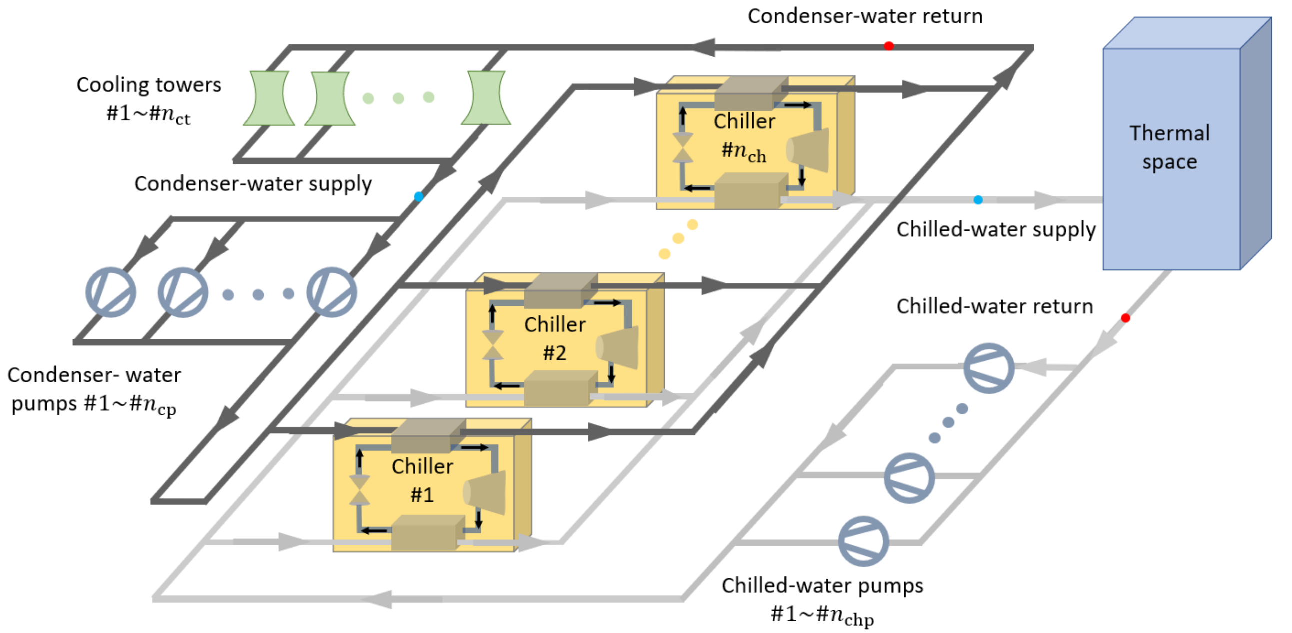
Method	P_D=13050 RT best / worst	mean	SD	average time	P_D=11600 RT best / worst	mean	SD	average time	P_D=10150 RT best / worst	mean	SD	average time
GA [12]	9293.78 / 9401.31	9326.04	19.29	7.50	7293.64 / 7361.76	7322.59	15.50	7.51	5910.94 / 6024.68	5947.46	22.22	7.40
PSO [12]	9307.41 / 10650.73	9734.65	351.31	4.68	7942.34 / 10992.43	9149.29	658.79	4.76	6269.72 / 7563.12	6799.53	220.73	4.74
DE [14]	10188.39 / 11267.59	10746.82	219.46	7.09	7917.54 / 9502.17	8921.39	307.28	7.11	6956.12 / 8048.28	7458.87	306.82	7.05
IFA [15]	9286.71 / 9287.49	9286.98	0.17	134.02	7278.37 / 7278.58	7278.45	0.04	129.78	5890.74 / 5965.45	5896.54	19.51	129.47
CNO-CL	**9286.49 / 9286.49**	**9286.49**	**0.00**	**0.25**	**7278.32 / 7278.32**	**7278.32**	**0.00**	**4.31**	**5890.69 / 5890.69**	**5890.69**	**0.00**	**4.36**

Method	P_D=8700 RT best / worst	mean	SD	average time	P_D=7250 RT best / worst	mean	SD	average time	P_D=5800 RT best / worst	mean	SD	average time
GA [12]	4975.30 / 5079.74	5019.05	23.24	7.39	4157.11 / 4481.54	4258.88	58.81	7.69	3322.83 / 3563.15	3437.40	57.32	7.65
PSO [12]	5080.86 / 5697.55	5407.83	126.59	**4.87**	4125.79 / 4699.08	4348.79	117.64	**4.79**	3245.64 / 3705.91	3508.28	105.54	**4.78**
DE [14]	5679.43 / 6747.46	6285.46	256.47	7.03	4423.34 / 5773.08	5164.85	272.14	7.02	3634.22 / 4624.62	4015.83	189.90	7.00
IFA [15]	4942.76 / 5005.12	4970.55	17.55	129.04	4103.48 / 4327.59	4181.83	51.55	130.60	3267.88 / 3503.05	3351.42	48.33	130.11
CNO-CL	**4942.64 / 4942.64**	**4942.64**	**0.00**	21.46	**4074.55 / 4080.84**	**4076.00**	**2.66**	62.02	**3225.91 / 3233.19**	**3226.63**	**2.20**	52.56

Z. Chen, J. Wang, and Q.L. Han, "Optimal chiller loading based on collaborative neurodynamic optimization," *IEEE Transactions on Industrial Informatics*, vol. 19, pp. 3057-3067, 2023.

CityU

Chiller Plant Operation Planning



- Determine the output of each device

e.g., on/off status, mass flow rates of water and air….

Challenges:

- More devices: complex equations
- More coupling constraints for the conservation of energy

Chiller Plant Operation Planning (cont'd)

Power consumption function: Chillers

$$P_{\mathrm{ch}}^{(i)}\left(T_{\mathrm{chws}}, T_{\mathrm{cws}}, Q_{\mathrm{ch}}^{(i)}\right) = \overline{P}_{\mathrm{ch}}^{(i)} \Phi_c^{(i)} \Phi_e^{(i)} \Phi_{\mathrm{ep}}^{(i)}$$

$$\Phi_c^{(i)}(T_{\mathrm{chws}}, T_{\mathrm{cws}}) = a_1^{(i)} + a_2^{(i)} T_{\mathrm{chws}} + a_3^{(i)} T_{\mathrm{chws}}^2 + a_4^{(i)} T_{\mathrm{cws}} + a_5^{(i)} T_{\mathrm{cws}}^2 + a_6^{(i)} T_{\mathrm{chws}} T_{\mathrm{cws}}$$

$$\Phi_e^{(i)}(T_{\mathrm{chws}}, T_{\mathrm{cws}}) = b_1^{(i)} + b_2^{(i)} T_{\mathrm{chws}} + b_3^{(i)} T_{\mathrm{chws}}^2 + b_4^{(i)} T_{\mathrm{cws}} + b_5^{(i)} T_{\mathrm{cws}}^2 + b_6^{(i)} T_{\mathrm{chws}} T_{\mathrm{cws}}$$

$$\Phi_{\mathrm{ep}}^{(i)}\left(\Phi_c^{(i)}, Q_{\mathrm{ch}}^{(i)}\right) = d_1^{(i)} H\left(Q_{\mathrm{ch}}^{(i)}\right) + d_2^{(i)} \frac{Q_{\mathrm{ch}}^{(i)}}{\Phi_c^{(i)} \overline{Q}_{\mathrm{ch}}^{(i)}} + d_3^{(i)} \left(\frac{Q_{\mathrm{ch}}^{(i)}}{\Phi_c^{(i)} \overline{Q}_{\mathrm{ch}}^{(i)}}\right)^2$$

Power consumption function: Pumps

$$P_{\mathrm{chp}}^{(j)}\left(r_{\mathrm{chpw}}^{(j)}\right) = \frac{k_{\mathrm{ch0}}^{(j)} \left(\frac{r_{\mathrm{chpw}}^{(j)}}{\overline{r}_{\mathrm{chpw}}^{(j)}}\right)^3}{\eta_{\mathrm{chp}}^{(j)} \left(1 - \exp\left(-k_{\mathrm{chp}}^{(j)} \frac{r_{\mathrm{chpw}}^{(j)}}{\overline{r}_{\mathrm{chpw}}^{(j)}}\right)\right)}$$

$$\eta_{\mathrm{chp}}^{(j)}\left(r_{\mathrm{chpw}}^{(j)}\right) = k_{\mathrm{ch1}}^{(j)} + k_{\mathrm{ch2}}^{(j)} \frac{r_{\mathrm{chpw}}^{(j)}}{\overline{r}_{\mathrm{chpw}}^{(j)}} + k_{\mathrm{ch3}}^{(j)} \left(\frac{r_{\mathrm{chpw}}^{(j)}}{\overline{r}_{\mathrm{chpw}}^{(j)}}\right)^2 + k_{\mathrm{ch4}}^{(j)} \left(\frac{r_{\mathrm{chpw}}^{(j)}}{\overline{r}_{\mathrm{chpw}}^{(j)}}\right)^3$$

$$P_{\mathrm{cp}}^{(k)}\left(r_{\mathrm{cpw}}^{(k)}\right) = \frac{k_{\mathrm{cp0}}^{(k)} \left(\frac{r_{\mathrm{cpw}}^{(k)}}{\overline{r}_{\mathrm{cpw}}^{(k)}}\right)^3}{\eta_{\mathrm{cp}}^{(k)} \left(1 - \exp\left(-k_{\mathrm{cp}}^{(k)} \frac{r_{\mathrm{cpw}}^{(k)}}{\overline{r}_{\mathrm{cpw}}^{(k)}}\right)\right)}$$

$$\eta_{\mathrm{cp}}^{(k)}\left(r_{\mathrm{chpw}}^{(k)}\right) = k_{\mathrm{cp1}}^{(k)} + k_{\mathrm{cp2}}^{(k)} \frac{r_{\mathrm{cpw}}^{(k)}}{\overline{r}_{\mathrm{chpw}}^{(k)}} + k_{\mathrm{cp3}}^{(k)} \left(\frac{r_{\mathrm{cpw}}^{(k)}}{\overline{r}_{\mathrm{cpw}}^{(k)}}\right)^2 + k_{\mathrm{cp4}}^{(k)} \left(\frac{r_{\mathrm{cpw}}^{(k)}}{\overline{r}_{\mathrm{cpw}}^{(k)}}\right)^3$$

Power consumption function: Cooling Towers

$$P_{\mathrm{ct}}^{(l)}\left(r_a^{(l)}\right) = \overline{P}_{\mathrm{ct}}^{(l)} \left(\frac{r_a^{(l)}}{\overline{r}_a^{(l)}}\right)^3$$

Chiller Plant Operation Planning (cont'd)

Supply–demand constraints

$$C_p \sum_{j=1}^{n_{\mathrm{chp}}} r_{\mathrm{chpw}}^{(j)} (T_{\mathrm{chwr}} - T_{\mathrm{chws}}) = P_D \quad P_D = \sum_{i=1}^{n_{\mathrm{ch}}} Q_{\mathrm{ch}}^{(i)}$$

Constraints for the conservation of energy

$$P_D + \sum_{i=1}^{n_{\mathrm{ch}}} P_{\mathrm{ch}}^{(i)} = C_p \sum_{k=1}^{n_{\mathrm{cp}}} r_{\mathrm{cpw}}^{(k)} (T_{\mathrm{cwr}} - T_{\mathrm{cws}})$$

$$P_D + \sum_{i=1}^{n_{\mathrm{ch}}} P_{\mathrm{ch}}^{(i)} = \varepsilon_a \sum_{l=1}^{n_{\mathrm{ct}}} r_a^{(l)} (h_{\mathrm{sw}} - h_{\mathrm{in}}) \quad \varepsilon_a(T_{\mathrm{cws}}, T_{\mathrm{cwr}}) = \frac{T_{\mathrm{cwr}} - T_{\mathrm{cws}}}{T_{\mathrm{cwr}} - T_{\mathrm{wb}}}$$

$$h_{\mathrm{sw}}(T_{\mathrm{cwr}}) = c_{t0} + c_{t1} T_{\mathrm{cwr}} + c_{t2} T_{\mathrm{cwr}}^2 + c_{t3} T_{\mathrm{cwr}}^3$$

Quadratic equations

$$y_{\mathrm{ch}} \circ (y_{\mathrm{ch}} - e) = 0, \quad y_{\mathrm{chp}} \circ (y_{\mathrm{chp}} - e) = 0$$
$$y_{\mathrm{cp}} \circ (y_{\mathrm{cp}} - e) = 0, \quad y_{\mathrm{ct}} \circ (y_{\mathrm{ct}} - e) = 0.$$

Capacity constraints

$$y_{\mathrm{ch}} \circ \underline{Q}_{\mathrm{ch}} \le Q_{\mathrm{ch}} \le y_{\mathrm{ch}} \circ \overline{Q}_{\mathrm{ch}}$$
$$y_{\mathrm{chp}} \circ \underline{r}_{\mathrm{chpw}} \le r_{\mathrm{chpw}} \le y_{\mathrm{chp}} \circ \overline{r}_{\mathrm{chpw}}$$
$$y_{\mathrm{cp}} \circ \underline{r}_{\mathrm{cpw}} \le r_{\mathrm{cpw}} \le y_{\mathrm{cp}} \circ \overline{r}_{\mathrm{cpw}}$$
$$y_{\mathrm{ct}} \circ \underline{r}_a \le r_a \le y_{\mathrm{ct}} \circ \overline{r}_a$$
$$\underline{T}_{\mathrm{chws}} \le T_{\mathrm{chws}} \le \overline{T}_{\mathrm{chws}}, \quad \underline{T}_{\mathrm{cws}} \le T_{\mathrm{cws}} \le \overline{T}_{\mathrm{cws}}$$
$$\underline{T}_{\mathrm{chwr}} \le T_{\mathrm{chwr}} \le \overline{T}_{\mathrm{chwr}}, \quad \underline{T}_{\mathrm{cwr}} \le T_{\mathrm{cwr}} \le \overline{T}_{\mathrm{cwr}}$$

Cardinality constraints

$$e^{\mathrm{T}} y_{\mathrm{ch}} \le k_{\mathrm{ch}}, \quad e^{\mathrm{T}} y_{\mathrm{chp}} \le k_{\mathrm{chp}},$$
$$e^{\mathrm{T}} y_{\mathrm{cp}} \le k_{\mathrm{cp}}, \quad e^{\mathrm{T}} y_{\mathrm{ct}} \le k_{\mathrm{ct}}$$

Chiller Plant Operation Planning (cont'd)

Problem formulation for chiller plant operation planning

min	**Total power consumption**
s.t.	**Supply–demand constraints**
	Constraints for the conservation of energy
	Capacity constraints
	Cardinality constraints
	Quadratic equations

Chiller Plant Operation Planning (cont'd)

A Chiller Plant With Homogeneous Devices: Save up to 55.69% of power consumption

Method	$P_D = 3033$ kW					$P_D = 6065$ kW				
	$\sum_{i=1}^{8} P_{\text{ch}}^{(i)}$	$\sum_{j=1}^{8} P_{\text{chp}}^{(j)}$	$\sum_{k=1}^{8} P_{\text{cp}}^{(k)}$	$\sum_{l=1}^{8} P_{\text{ct}}^{(l)}$	P_{total}	$\sum_{i=1}^{8} P_{\text{ch}}^{(i)}$	$\sum_{j=1}^{8} P_{\text{chp}}^{(j)}$	$\sum_{k=1}^{8} P_{\text{cp}}^{(k)}$	$\sum_{l=1}^{8} P_{\text{ct}}^{(l)}$	P_{total}
DE [17]	Solution infeasible					Solution infeasible				
PSO-GA [11]	626.58	118.42	88.68	34.90	868.57^{+}	1096.61	38.21	143.42	31.96	1310.19^{+}
GA [8]	790.41	93.62	80.79	36.82	1001.65^{+}	1077.35	143.30	26.42	21.16	1268.23^{+}
DC [9]	468.74	16.65	5.81	20.00	511.21	832.55	28.46	8.42	3.88	873.31
CNO-CPOP	**428.05**	**12.66**	**2.95**	**0.15**	**443.80**	**828.84**	**25.31**	**5.84**	**1.17**	**861.16**

Method	$P_D = 9098$ kW					$P_D = 12131$ kW				
	$\sum_{i=1}^{8} P_{\text{ch}}^{(i)}$	$\sum_{j=1}^{8} P_{\text{chp}}^{(j)}$	$\sum_{k=1}^{8} P_{\text{cp}}^{(k)}$	$\sum_{l=1}^{8} P_{\text{ct}}^{(l)}$	P_{total}	$\sum_{i=1}^{8} P_{\text{ch}}^{(i)}$	$\sum_{j=1}^{8} P_{\text{chp}}^{(j)}$	$\sum_{k=1}^{8} P_{\text{cp}}^{(k)}$	$\sum_{l=1}^{8} P_{\text{ct}}^{(l)}$	P_{total}
DE [17]	Solution infeasible					Solution infeasible				
PSO-GA [11]	1482.01	157.07	50.30	27.44	1716.82^{+}	1861.84	302.16	73.75	40.58	2278.34^{+}
GA [8]	1416.97	105.61	55.22	30.18	1607.98^{+}	1814.10	158.12	57.21	21.77	2051.19^{+}
DC [9]	**1253.41**	56.35	**10.37**	28.09	1348.22	**1701.94**	**105.57**	**21.34**	**9.43**	**1838.29**
CNO-CPOP	1256.82	**53.07**	10.38	**3.95**	**1324.22**	**1701.94**	**105.57**	**21.34**	**9.43**	**1838.29**

A Chiller Plant With Heterogeneous Devices: Save up to 58.30% of power consumption

Method	$P_D = 1312$ kW					$P_D = 2625$ kW				
	$\sum_{i=1}^{4} P_{\text{ch}}^{(i)}$	$\sum_{j=1}^{4} P_{\text{chp}}^{(j)}$	$\sum_{k=1}^{4} P_{\text{cp}}^{(k)}$	$\sum_{l=1}^{4} P_{\text{ct}}^{(l)}$	P_{total}	$\sum_{i=1}^{4} P_{\text{ch}}^{(i)}$	$\sum_{j=1}^{4} P_{\text{chp}}^{(j)}$	$\sum_{k=1}^{4} P_{\text{cp}}^{(k)}$	$\sum_{l=1}^{4} P_{\text{ct}}^{(l)}$	P_{total}
DE [17]	Solution infeasible					Solution infeasible				
PSO-GA [11]	188.53	28.19	41.85	19.28	277.85^{+}	354.52	56.79	29.36	20.00	460.67^{+}
GA [8]	255.32	22.93	20.40	10.21	308.87^{+}	350.38	43.94	21.99	14.17	430.48^{+}
CNO-CPOP	**122.41**	**4.80**	**1.55**	**0.04**	**128.81**	**274.72**	**11.23**	**2.80**	**0.64**	**289.39**

Method	$P_D = 3937$ kW					$P_D = 5250$ kW				
	$\sum_{i=1}^{4} P_{\text{ch}}^{(i)}$	$\sum_{j=1}^{4} P_{\text{chp}}^{(j)}$	$\sum_{k=1}^{4} P_{\text{cp}}^{(k)}$	$\sum_{l=1}^{4} P_{\text{ct}}^{(l)}$	P_{total}	$\sum_{i=1}^{4} P_{\text{ch}}^{(i)}$	$\sum_{j=1}^{4} P_{\text{chp}}^{(j)}$	$\sum_{k=1}^{4} P_{\text{cp}}^{(k)}$	$\sum_{l=1}^{4} P_{\text{ct}}^{(l)}$	P_{total}
DE [17]	Solution infeasible					Solution infeasible				
PSO-GA [11]	474.29	73.15	35.64	25.83	608.92^{+}	691.37	81.34	57.82	29.72	860.25^{+}
GA [8]	491.95	50.31	45.60	16.59	604.44^{+}	671.80	50.98	33.30	21.74	777.82^{+}
CNO-CPOP	**432.23**	**30.26**	**5.01**	**1.15**	**468.64**	**599.78**	**51.08**	**8.31**	**2.76**	**661.92**

Z. Chen, J. Wang, and Q.L. Han, "Chiller plant operation planning via collaborative neurodynamic optimization," *IEEE Transactions on Systems, Man and Cybernetics: Systems*, vol. 53, pp. 4623-4635, 2023.

Distributed Chiller Loading

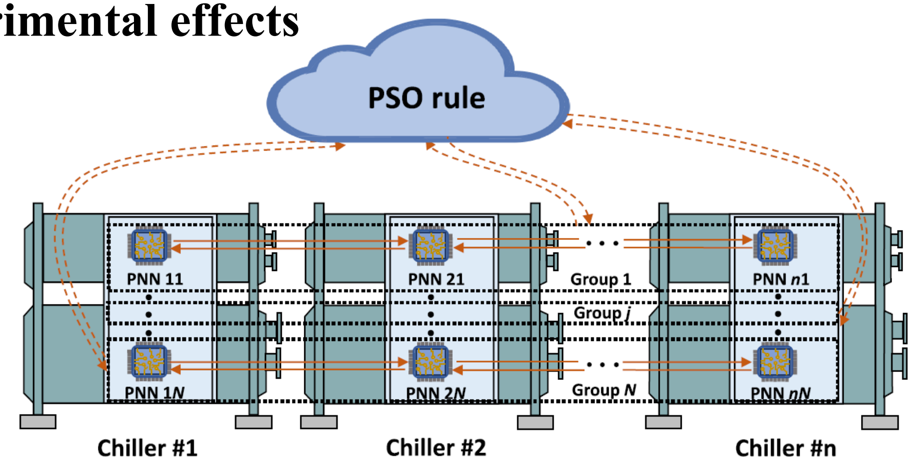
Motivations

- Most existing chiller systems **operate in centralized locations**
- There are **several disadvantages** in centralized chiller systems
  - Centralized information processing **entails high reliability of communication** for effective chiller loading
  - Centrally located chillers with **long circulation pipelines** are **usually inefficient with time delays** and **energy losses** in cooling load dispatching

Distant chilled-water circulation Communication-link failures	→	Chiller loading delays Energy losses	→ **Distributed optimization** / **IoT devices** →	√
Reason		**Detrimental effects**		

IoT-based distributed chillier loading

- Networked IoT devices are installed in chillers for **communication**, **computation**, and **control**
- Could **overcome the limitations** of centralized HVAC systems



Distributed Chiller Loading I (cont'd)

$$\min_{PLR,y} \sum_{i=1}^{n} P_i(PLR_i)$$

s.t.

$$\sum_{i=1}^{n} \overline{P}_i PLR_i - P_D = 0$$

$$\underline{PLR}_i y_i \le PLR_i \le \overline{PLR}_i y_i$$

$$\sum_{i=1}^{n} y_i \le k$$

$$y_i(y_i - 1) = 0, \quad i = 1, \ldots, n$$

$$\sum_{i=1}^{n}(\overline{P}_i PLR_i - P_{D_i}) = 0, \quad \sum_{i=1}^{n}(y_i - k_i) \le 0.$$

$$\sum_{i=1}^{n} P_{D_i} = P_D \quad \sum_{i=1}^{n} k_i = k$$

$$\mathcal{L}(PLR, y, \lambda, \mu, \gamma, \rho) = \sum_{i=1}^{n} P_i(PLR_i) + \lambda^T g(PLR, y) + \mu^T h(PLR, y) + \gamma^T(D(\overline{P})PLR - \tilde{P}_D) + \rho^T(y - \tilde{k})$$

$$L\gamma = 0, \quad L\rho = 0$$

$$\hat{\mathcal{L}}(PLR, y, \lambda, \mu, \gamma, \rho, \nu, \omega) = \mathcal{L}(PLR, y, \lambda, \mu, \gamma, \rho) + \nu^T L\gamma + \omega^T L\rho$$

$$\mathcal{L}_a(PLR, y, \lambda, \mu, \gamma, \rho, \nu, \omega) = \hat{\mathcal{L}}(PLR, y, \lambda, \mu, \gamma, \rho, \nu, \omega) + \frac{\alpha}{2}\sum_{i=1}^{2n} \lambda_i g_i(PLR, y)^2 + \frac{\beta}{2}\sum_{i=1}^{n} h_i(PLR, y)^2$$

$$x = [PLR^T, y^T]^T$$

Equivalent in terms of the optimal solution

$$\min_{x} \max_{\{\lambda,\mu,\gamma,\rho,\nu,\omega\}} \mathcal{L}_a(x, \lambda, \mu, \gamma, \rho, \nu, \omega)$$

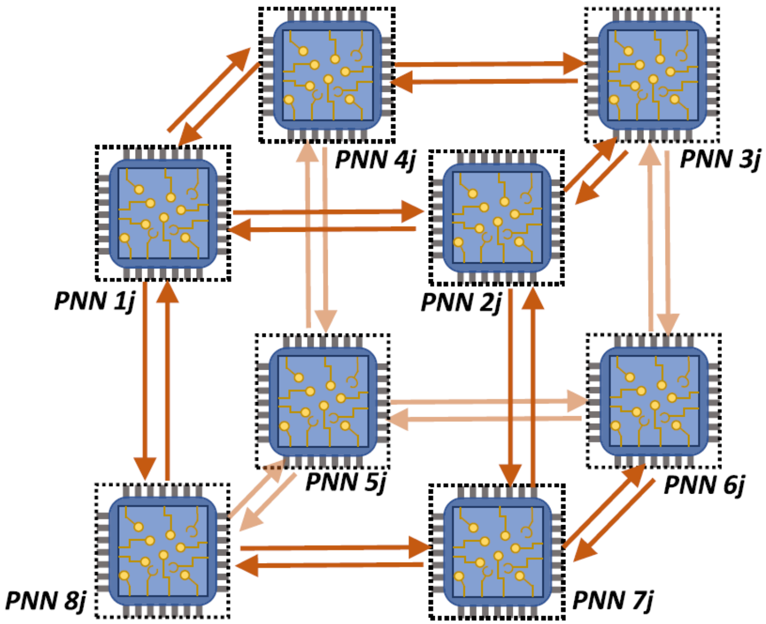
Distributed Chiller Loading I (cont'd)

Coupled projection neural networks

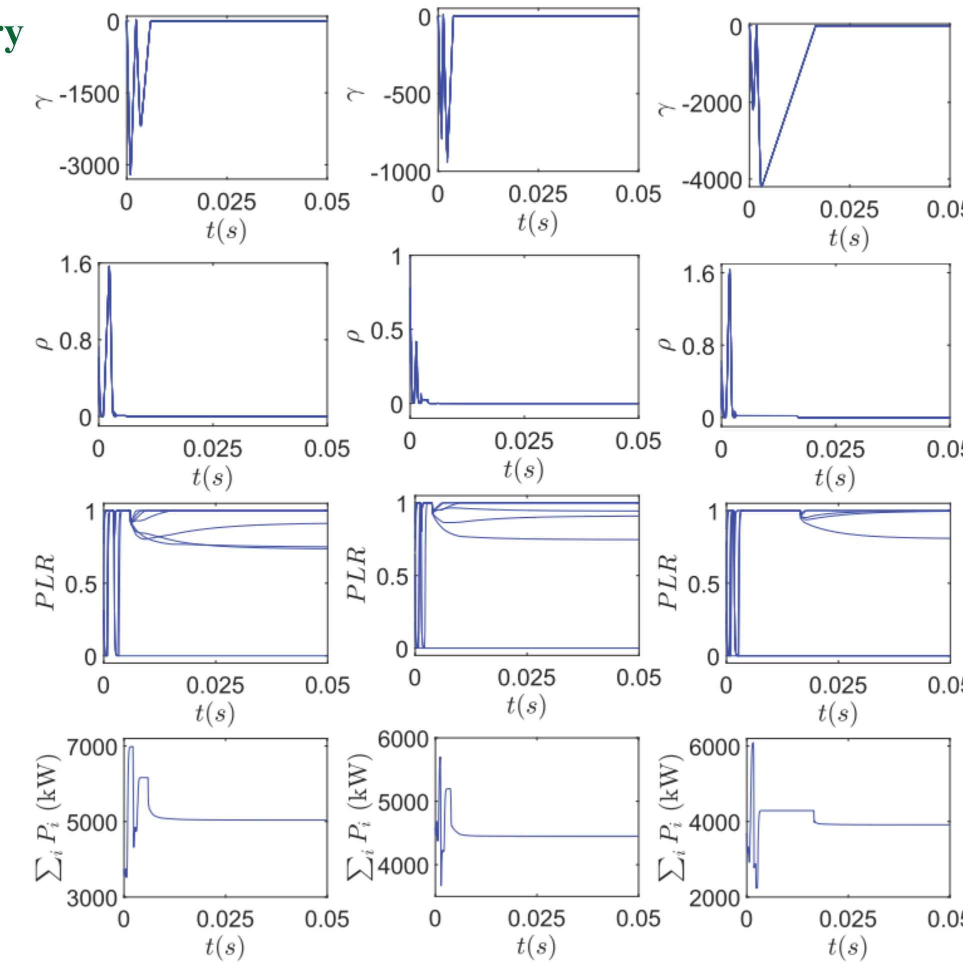
$$
\begin{cases}
\epsilon \dfrac{dx^{(j)}}{dt} = -x^{(j)} + P_\Omega\Bigg(x^{(j)} - \Bigg(\nabla P(x^{(j)}) + \nabla g(x^{(j)})\lambda^{(j)} \\
\qquad + \nabla h(x^{(j)})\mu^{(j)} + \alpha \nabla g(x^{(j)}) D(\lambda^{(j)}) g(x^{(j)}) \\
\qquad + \beta \nabla h(x^{(j)}) h(x^{(j)}) + \begin{bmatrix} D(\overline{P})\gamma^{(j)} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \rho^{(j)} \end{bmatrix} \Bigg)\Bigg) \\
\epsilon \dfrac{d\lambda^{(j)}}{dt} = -\lambda^{(j)} + (\lambda^{(j)} + g(x^{(j)}))^+ \\
\epsilon \dfrac{d\mu^{(j)}}{dt} = h(x^{(j)}) \\
\epsilon \dfrac{d\gamma^{(j)}}{dt} = D(\overline{P})PLR^{(j)} - \tilde{P}_D - L\nu^{(j)} - L\gamma^{(j)} \\
\epsilon \dfrac{d\nu^{(j)}}{dt} = L\gamma^{(j)} \\
\epsilon \dfrac{d\rho^{(j)}}{dt} = -\rho^{(j)} + \left(\rho^{(j)} + \left(y^{(j)} - \tilde{k}\right) - L\rho^{(j)} - L\omega^{(j)}\right)^+ \\
\epsilon \dfrac{d\omega^{(j)}}{dt} = L\rho^{(j)}.
\end{cases}
$$

Distributed Chiller Loading I (cont'd)

An eight-chiller system in a semiconductor factory in Taiwan Hsinchu Science Industrial District



Communication Links among Eight Chillers



Distributed Chiller Loading I (cont'd)

Performance Comparisons

A 20-chiller system

Method	Type	$P_D = 8000$ RT				$P_D = 7000$ RT				$P_D = 6000$ RT			
		best	worst	average	SD	best	worst	average	SD	best	worst	average	SD
GA [5]	Cen.	4753.76	4937.83	4820.09	41.49	4007.62	4306.64	4188.72	41.47	3405.03	3762.39	3625.51	104.95
PSO [5]	Cen.	4734.77	5952.85	5022.53	256.79	3935.20	4736.15	4147.30	200.63	**3216.29**	3976.31	3336.61	166.04
DE [7]	Cen.	4834.23	5544.16	5145.81	157.69	4014.22	4997.32	4425.73	210.09	3273.93	4106.53	3690.33	204.26
IFA [8]	Cen.	4735.71	4735.83	4735.75	0.02	3963.71	4112.02	4087.98	46.04	3296.50	3714.69	3486.36	114.06
CNO-CL [12]	Cen.	**4734.01**	**4734.01**	**4734.01**	**0.00**	**3935.19**	**3935.19**	**3935.19**	**0.00**	**3216.29**	**3216.29**	**3216.29**	**0.00**
CNO-DCL	Dis.	**4734.01**	**4734.01**	**4734.01**	**0.00**	**3935.19**	**3935.19**	**3935.19**	**0.00**	**3216.29**	**3216.29**	**3216.29**	**0.00**

Method	Type	$P_D = 5000$ RT				$P_D = 4000$ RT				$P_D = 3000$ RT			
		best	worst	average	SD	best	worst	average	SD	best	worst	average	SD
GA [5]	Cen.	2766.31	3280.86	2994.23	123.97	2123.48	2837.19	2368.83	155.69	1533.06	2131.79	1751.24	153.42
PSO [5]	Cen.	2587.76	3239.43	2706.77	163.63	**1922.78**	2272.56	1984.96	97.14	**1363.33**	2089.65	1498.63	147.25
DE [7]	Cen.	2587.76	2650.48	2589.64	10.75	1952.11	2762.98	2308.12	192.71	**1363.33**	1515.19	1377.59	20.77
IFA [8]	Cen.	**2557.33**	3578.22	2939.19	185.82	1922.79	2936.13	2364.70	196.63	1408.76	2297.29	1790.33	177.22
CNO-CL [12]	Cen.	**2557.33**	**2557.33**	**2557.33**	**0.00**	**1922.78**	**1922.78**	**1922.78**	**0.00**	**1363.33**	**1363.33**	**1363.33**	**0.00**
CNO-DCL	Dis.	**2557.33**	**2557.33**	**2557.33**	**0.00**	**1922.78**	**1922.78**	**1922.78**	**0.00**	**1363.33**	**1363.33**	**1363.33**	**0.00**

Z. Chen, J. Wang, and Q.L. Han, "A collaborative neurodynamic optimization approach to distributed chiller loading," *IEEE Transactions on Neural Networks and Learning Systems*, in press, 2023.

Event-triggered Projection Neural Network

Z. Xia, Y. Liu, and J. Wang, "An event-triggered collaborative neurodynamic approach to distributed global optimization," *Neural Networks*, vol. 169, pp.181-190, January 2024.

$$
\begin{cases}
\epsilon \dfrac{dx_i(t)}{dt} = -\theta(t)x_i(t) + \theta(t)P_{\Omega_i}\Bigg(x_i(t) - \Bigg(\nabla f_i(x_i(t)) + \nabla g_i'(x_i(t))^T \lambda_i'(t) \\
\qquad + \nabla h_i'(x_i(t))^T \nu_i'(t) + \nabla \bar{g}_i''(x_i(t))^T \zeta_i''(t) \\
\qquad + \nabla \bar{h}_i''(x_i(t))^T \mu_i''(t) \\
\qquad + \alpha \nabla g_i'(x_i(t))^T \mathrm{diag}(\lambda_i'(t)^2) g_i'(x_i(t)) \\
\qquad + \beta \nabla h_i'(x_i(t))^T h_i'(x_i(t)) \Bigg)\Bigg), \\
\epsilon \dfrac{d\lambda_i'(t)}{dt} = \theta(t)\left(-\lambda_i'(t) + (\lambda_i'(t) + g_i'(x_i(t)))^+\right), \\
\epsilon \dfrac{d\nu_i'(t)}{dt} = \theta(t) h_i'(x_i(t)), \\
\epsilon \dfrac{d\zeta_i''(t)}{dt} = \theta(t)\Bigg(-\zeta_i''(t) + \Bigg(\zeta_i''(t) + g_i''(x_i(t)) \\
\qquad - \sum_{j=1}^{N} a_{ij}\left(\zeta_i''(t) + \rho_i''(t) - \zeta_j^e(t) - \rho_j^e(t)\right)\Bigg)^+\Bigg), \\
\epsilon \dfrac{d\rho_i''(t)}{dt} = \theta(t)\sum_{j=1}^{N} a_{ij}\left(\zeta_i''(t) - \zeta_j^e(t)\right), \\
\epsilon \dfrac{d\mu_i''(t)}{dt} = \theta(t)\left(h_i''(x_i(t)) - \sum_{j=1}^{N} a_{ij}\left(\mu_i''(t) + \omega_i''(t) - \mu_j^e(t) - \omega_j^e(t)\right)\right), \\
\epsilon \dfrac{d\omega_i''(t)}{dt} = \theta(t)\sum_{j=1}^{N} a_{ij}\left(\mu_i''(t) - \mu_j^e(t)\right)
\end{cases}
$$

Event-triggered Updating Rules

$$\hat{\zeta}_i^e\left(\tau_{k_i+1}^{(i)}\right) = \hat{\zeta}_i^e\left(\tau_{k_i}^{(i)}\right) - \eta\left(\tau_{k_i+1}^{(i)}\right)\theta\left(\tau_{k_i+1}^{(i)}\right)Q\left(\sigma_i^{\zeta}\left(\tau_{k_i+1}^{(i)}\right)\right)$$

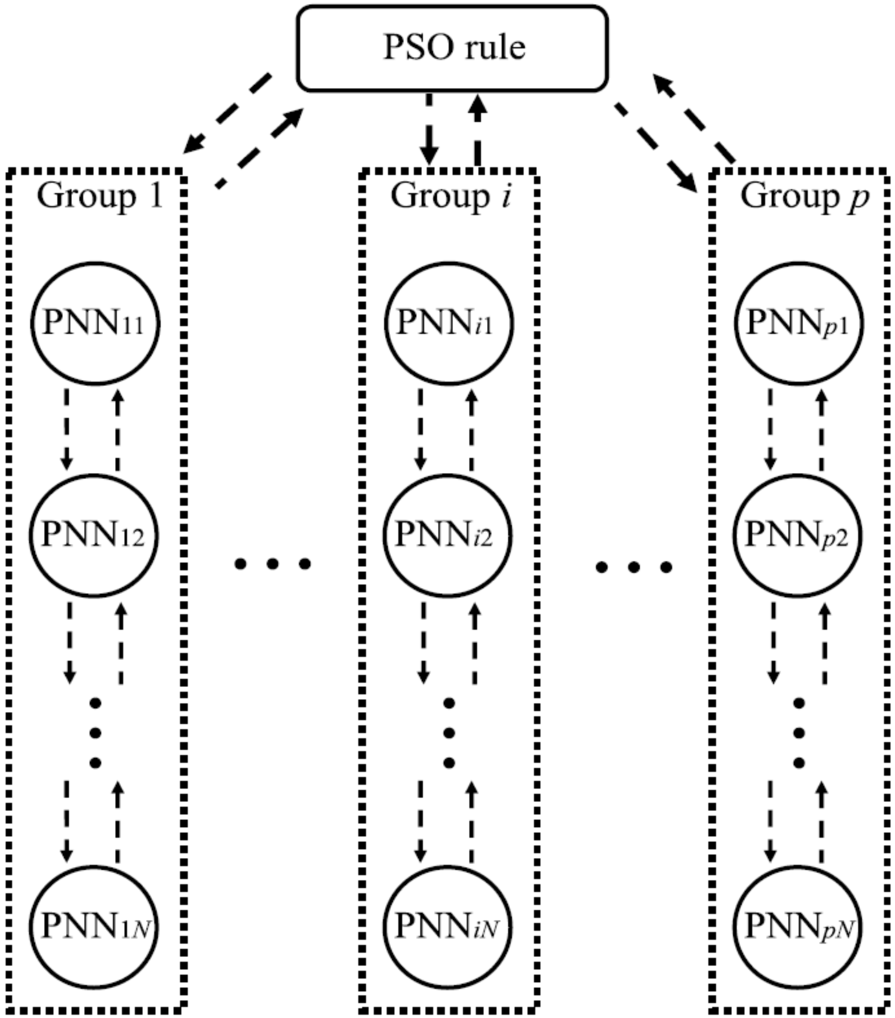
$$\hat{\rho}_i^e\left(\tau_{k_i+1}^{(i)}\right) = \hat{\rho}_i^e\left(\tau_{k_i}^{(i)}\right) - \eta\left(\tau_{k_i+1}^{(i)}\right)\theta\left(\tau_{k_i+1}^{(i)}\right)Q\left(\sigma_i^{\rho}\left(\tau_{k_i+1}^{(i)}\right)\right)$$

$$\hat{\mu}_i^e\left(\tau_{k_i+1}^{(i)}\right) = \hat{\mu}_i^e\left(\tau_{k_i}^{(i)}\right) - \eta\left(\tau_{k_i+1}^{(i)}\right)\theta\left(\tau_{k_i+1}^{(i)}\right)Q\left(\sigma_i^{\mu}\left(\tau_{k_i+1}^{(i)}\right)\right)$$

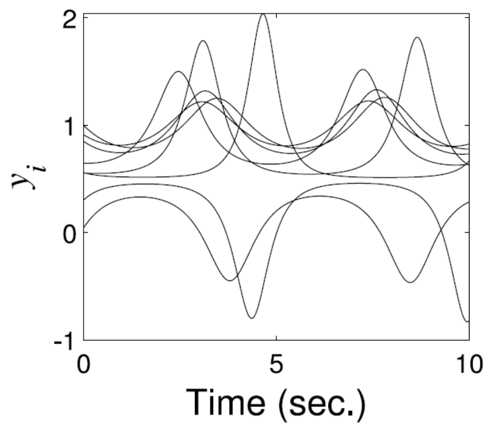
$$\hat{\omega}_i^e\left(\tau_{k_i+1}^{(i)}\right) = \hat{\omega}_i^e\left(\tau_{k_i}^{(i)}\right) - \eta\left(\tau_{k_i+1}^{(i)}\right)\theta\left(\tau_{k_i+1}^{(i)}\right)Q\left(\sigma_i^{\omega}\left(\tau_{k_i+1}^{(i)}\right)\right)$$

Z. Xia, Y. Liu, and **J. Wang**, "An event-triggered collaborative neurodynamic approach to distributed global optimization," *Neural Networks*, vol. 169, pp.181-190, January 2024.

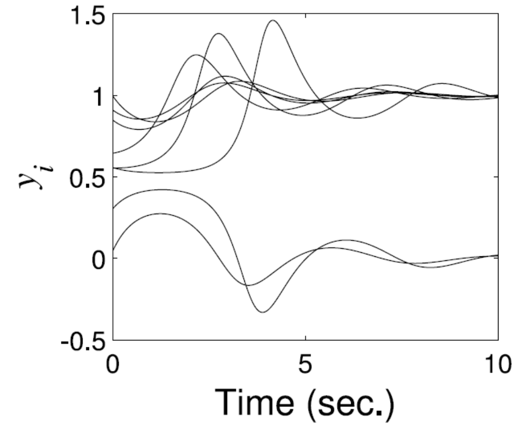
Schematic Diagram



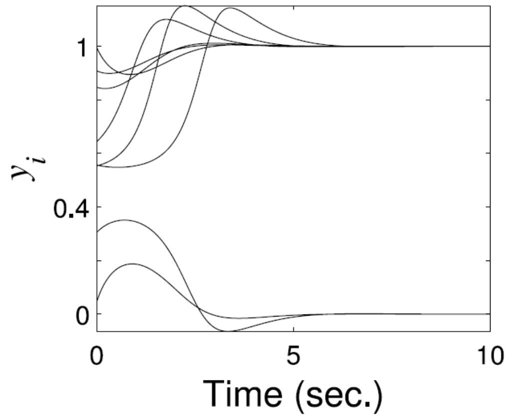
Parametrical Effects



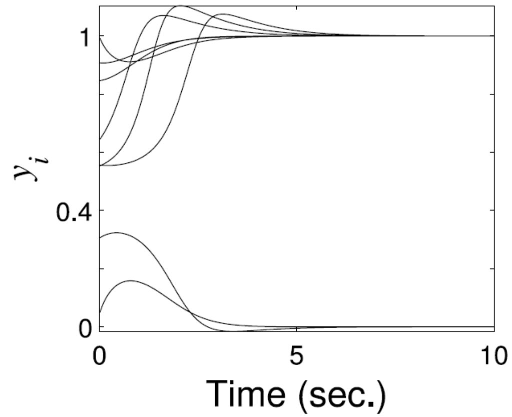
(a) $\beta = 0$.



(b) $\beta = 1$.

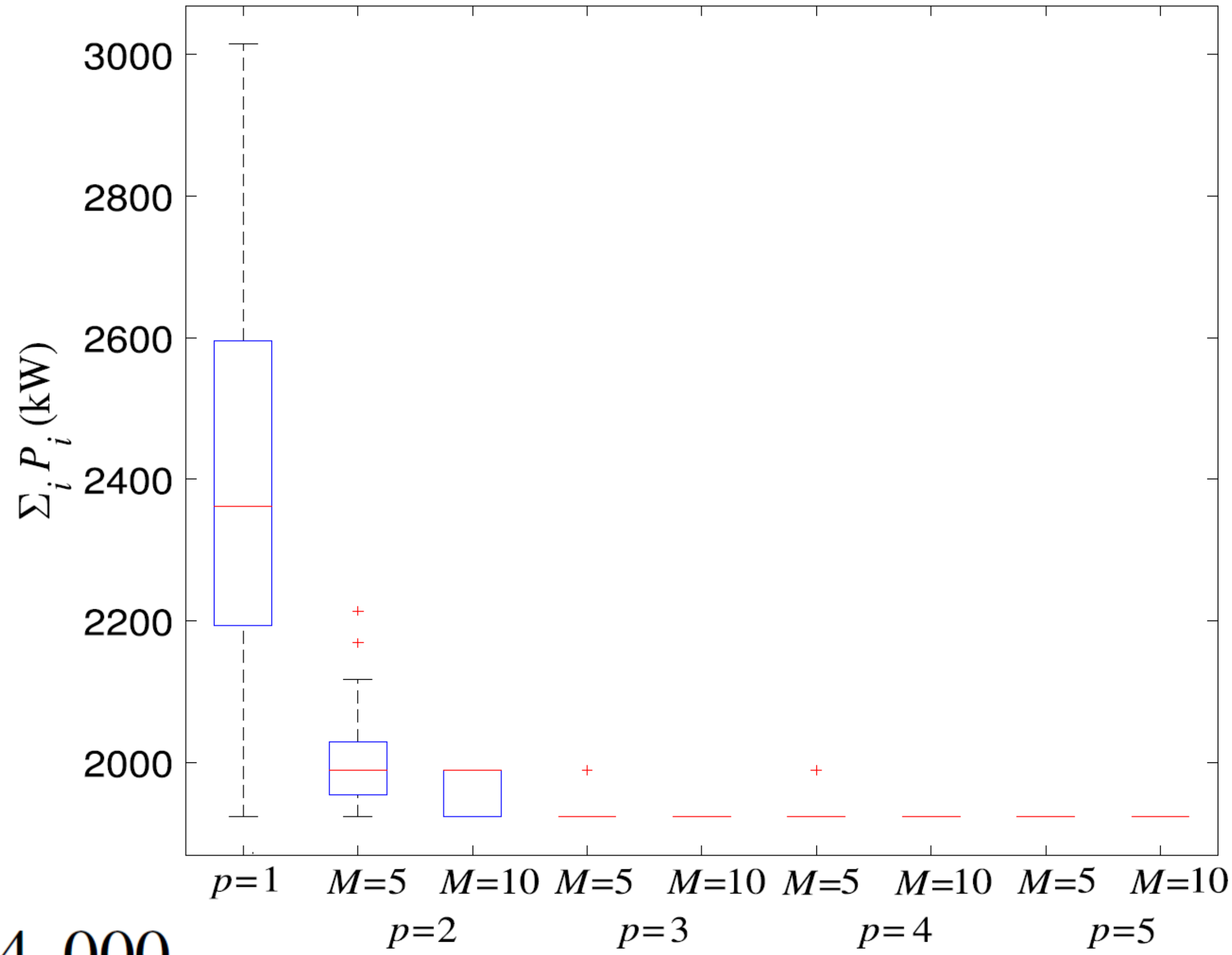


(c) $\beta = 2$.



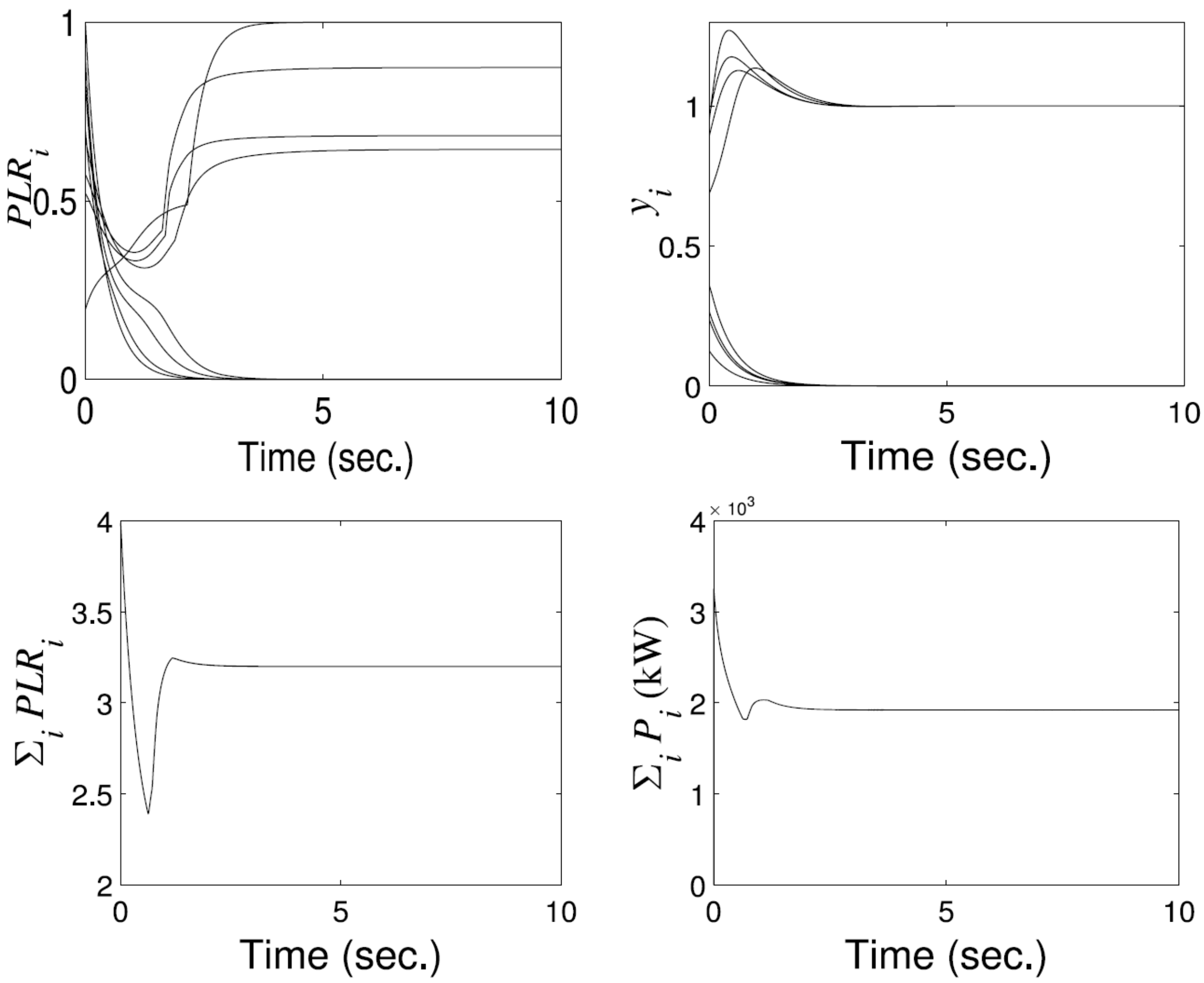
(d) $\beta = 3$.

Parametrical Effects (cont'd)

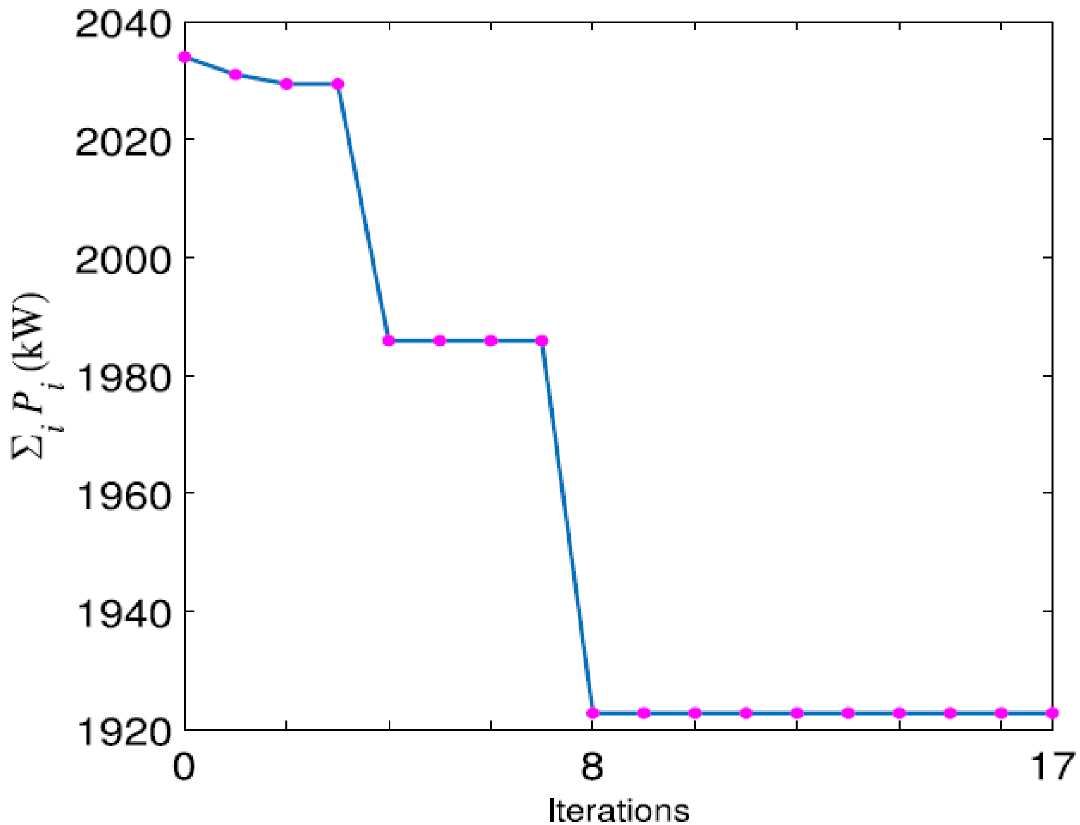


$P_D = 4,000$

Convergent Behaviors

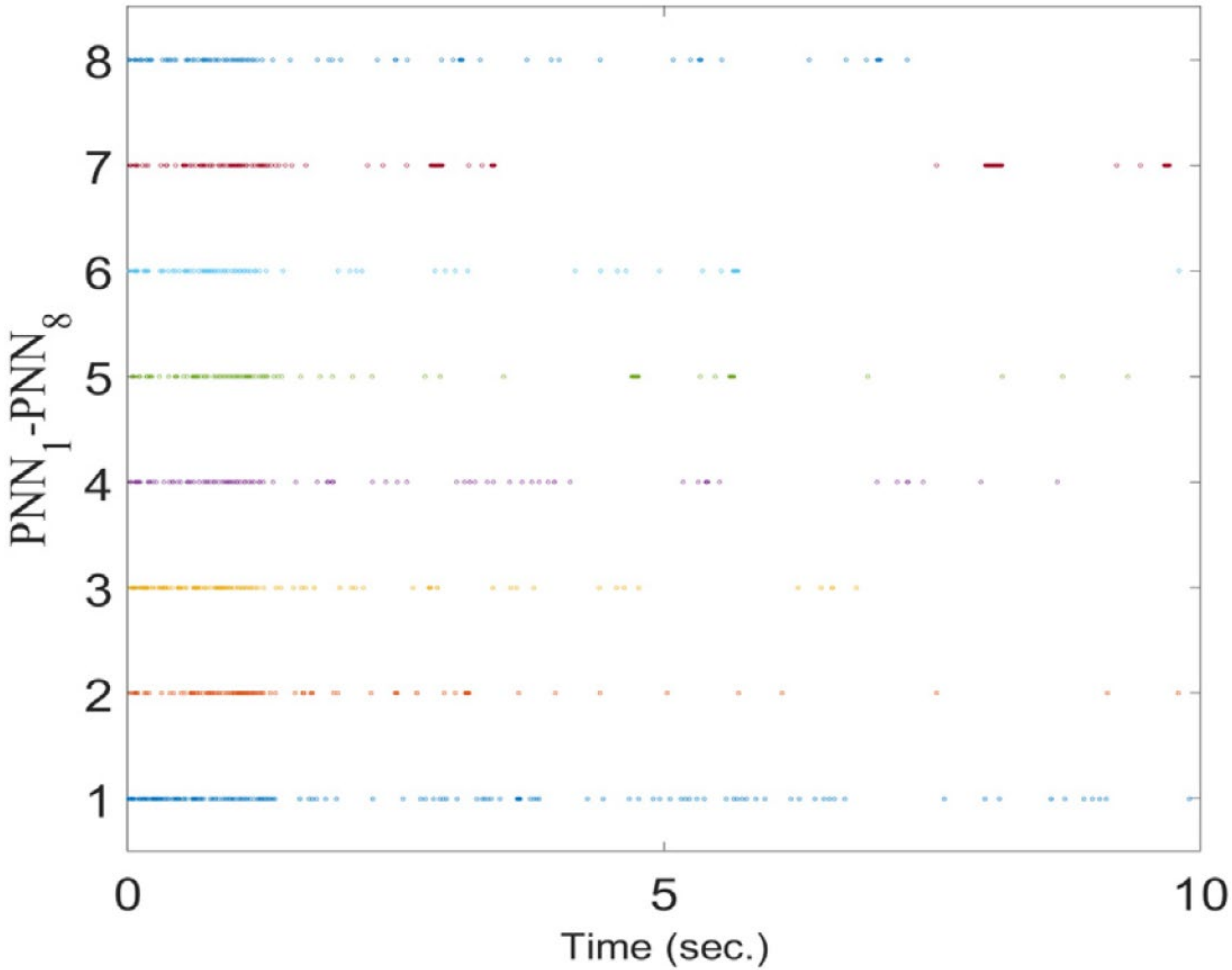


Convergent Behaviors (cont'd)



Z. Xia, Y. Liu, and **J. Wang**, "An event-triggered collaborative neurodynamic approach to distributed global optimization," *Neural Networks*, vol. 169, pp.181-190, January 2024.

Communication Frequency



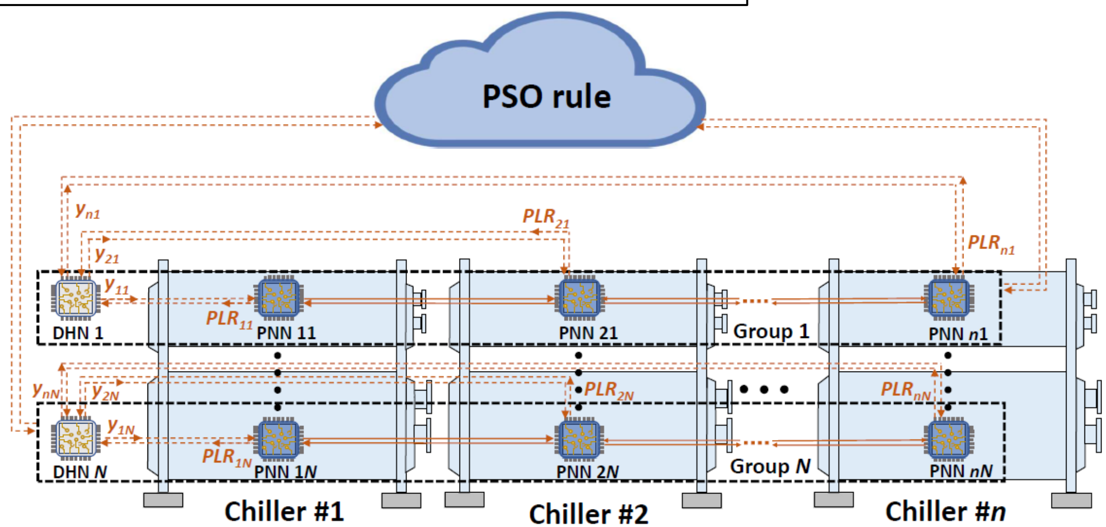
CityU

Distributed Chiller Loading II

$$\min_{PLR,y} \sum_{i=1}^{n} P_i(PLR_i)y_i$$
$$\text{s.t.} \sum_{i=1}^{n} \overline{P}_i PLR_i y_i - P_D = 0,$$
$$\underline{PLR}_i \le PLR_i \le \overline{PLR}_i,$$
$$y_i \in \{0,1\}, \; i = 1, \ldots, n;$$

$$\min_{y} \sum_{i=1}^{n} P_i\left(PLR_i\right) y_i$$
$$\text{s.t.} \sum_{i=1}^{n} \overline{P}_i PLR_i y_i - P_D \ge 0,$$
$$y_i \in \{0,1\}, i = 1, \ldots, n.$$

$$\min_{PLR} \sum_{i \in \mathcal{I}} P_i\left(PLR_i\right)$$
$$\text{s.t.} \sum_{i \in \mathcal{I}} \bar{P}_i PLR_i - P_D = 0$$
$$\underline{PLR}_i \le PLR_i \le \overline{PLR}_i, i \in \mathcal{I}.$$



Distributed Chiller Loading II (cont'd)

An eight-chiller system

Method	Type	P_D=8000 RT best / worst	mean	SD	average time	P_D=7000 RT best / worst	mean	SD	average time	P_D=6000 RT best / worst	mean	SD	average time
GA [9]	Cen.	4753.76 / 4937.83	4820.09	41.49	4.89	4007.62 / 4306.64	4188.72	41.47	4.74	3405.03 / 3762.39	3625.51	104.95	4.75
PSO [9]	Cen.	4734.77 / 5952.85	5022.53	256.79	1.59	3935.20 / 4736.15	4147.30	200.63	1.58	3216.29 / 3976.31	3336.61	166.04	1.58
DE [11]	Cen.	4834.23 / 5544.16	5145.81	157.69	4.27	4014.22 / 4997.32	4425.73	210.09	4.30	3273.93 / 4106.53	3690.33	204.26	4.29
IFA [12]	Cen.	4735.71 / 4735.83	4735.75	0.02	28.21	3963.71 / 4112.02	4087.98	46.04	28.09	3296.50 / 3714.69	3486.36	114.06	28.06
CNO-CL [14]	Cen.	4734.01 / 4734.01	4734.01	0.00	15.95	3935.19 / 3935.19	3935.19	0.00	36.25	3216.29 / 3216.29	3216.29	0.00	60.14
CNO-DCL	Dis.	4734.01 / 4734.01	4734.01	0.00	2.84	3935.19 / 3935.19	3935.19	0.00	26.73	3216.29 / 3216.29	3216.29	0.00	25.25

Method	Type	P_D=5000 RT best / worst	mean	SD	average time	P_D=4000 RT best / worst	mean	SD	average time	P_D=3000 RT best / worst	mean	SD	average time
GA [9]	Cen.	2766.31 / 3280.86	2994.23	123.97	4.96	2123.48 / 2837.19	2368.83	155.69	4.96	1533.06 / 2131.79	1751.24	153.42	5.10
PSO [9]	Cen.	2587.76 / 3239.43	2706.77	163.63	1.57	1922.78 / 2272.56	1984.96	97.14	1.68	1363.33 / 2089.65	1498.63	147.25	1.72
DE [11]	Cen.	2587.76 / 2650.48	2589.64	10.75	4.19	1952.11 / 2762.98	2308.12	192.71	4.02	1363.33 / 1515.19	1377.59	20.77	4.02
IFA [12]	Cen.	2557.33 / 3578.22	2939.19	185.82	28.08	1922.79 / 2936.13	2364.70	196.63	27.97	1408.76 / 2297.29	1790.33	177.22	27.69
CNO-CL [14]	Cen.	2557.33 / 2557.33	2557.33	0.00	74.42	1922.78 / 1922.78	1922.78	0.00	73.13	1363.33 / 1363.33	1363.33	0.00	82.72
CNO-DCL	Dis.	2557.33 / 2557.33	2557.33	0.00	3.98	1922.78 / 1922.78	1922.78	0.00	17.23	1363.33 / 1363.33	1363.33	0.00	29.33

A 20-chiller system

Method	Type	P_D=13050 RT best / worst	mean	SD	average time	P_D=11600 RT best / worst	mean	SD	average time	P_D=10150 RT best / worst	mean	SD	average time
GA [9]	Cen.	9293.78 / 9401.31	9326.04	19.29	7.50	7293.64 / 7361.76	7322.59	15.50	7.51	5910.94 / 6024.68	5947.46	22.22	7.40
PSO [9]	Cen.	9307.41 / 10650.73	9734.65	351.31	4.68	7942.34 / 10992.43	9149.29	658.79	4.76	6269.72 / 7563.12	6799.53	220.73	4.74
DE [11]	Cen.	10188.39 / 11267.59	10746.82	219.46	7.09	7917.54 / 9502.17	8921.39	307.28	7.11	6956.12 / 8048.28	7458.87	306.82	7.05
IFA [12]	Cen.	9286.71 / 9287.49	9286.98	0.17	134.02	7278.37 / 7278.58	7278.45	0.04	129.78	5890.74 / 5965.45	5896.54	19.51	129.47
CNO-CL [14]	Cen.	9286.49 / 9286.49	9286.49	0.00	62.48	7278.32 / 7278.32	7278.32	0.00	48.41	5890.69 / 5890.69	5890.69	0.00	57.45
CNO-DCL	Dis.	9286.49 / 9286.49	9286.49	0.00	0.53	7278.32 / 7278.32	7278.32	0.00	0.65	5890.69 / 5890.69	5890.69	0.00	1.22

Method	Type	P_D=8700 RT best / worst	mean	SD	average time	P_D=7250 RT best / worst	mean	SD	average time	P_D=5800 RT best / worst	mean	SD	average time
GA [9]	Cen.	4975.30 / 5079.74	5019.05	23.24	7.39	4157.11 / 4481.54	4258.88	58.81	7.69	3322.83 / 3563.15	3437.40	57.32	7.65
PSO [9]	Cen.	5080.86 / 5697.55	5407.83	126.59	4.87	4125.79 / 4699.08	4348.79	117.64	4.79	3245.64 / 3705.91	3508.28	105.54	4.78
DE [11]	Cen.	5679.43 / 6747.46	6285.46	256.47	7.03	4423.34 / 5773.08	5164.85	272.14	7.02	3634.22 / 4624.62	4015.83	189.90	7.00
IFA [12]	Cen.	4942.76 / 5005.12	4970.55	17.55	129.04	4103.48 / 4327.59	4181.83	51.55	130.60	3267.88 / 3503.05	3351.42	48.33	130.11
CNO-CL [14]	Cen.	4942.64 / 4942.64	4942.64	0.00	296.96	4074.55 / 4080.84	4075.28	1.95	672.30	3225.91 / 3233.22	3226.64	2.25	657.17
CNO-DCL	Dis.	4942.64 / 4942.64	4942.64	0.00	6.11	4074.55 / 4074.55	4074.55	0.00	9.99	3225.91 / 3225.91	3225.91	0.00	4.45

Z. Chen, J. Wang, and Q.L. Han, "Distributed Chiller Loading via Collaborative Neurodynamic Optimization with Heterogeneous Neural Networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 54, no 4, 2024.

Event-triggered Chiller Loading

$$\min_{\mathrm{PLR},y} \sum_{i=1}^{n} f_i(\mathrm{PLR}_i(t))$$

Power consumption

$$\text{s.t.} \sum_{i=1}^{n} \bar{P}_i \, \mathrm{PLR}_i(t) - P_D = 0,$$

Supply-demand constraint

$$\mathrm{PLR}_i \, y_i(t) \leq P_i(t) \leq \overline{\mathrm{PLR}}_i y_i(t),$$

Capacity constraints

$$\sum_{i=1}^{n} y_i(t) \leq k(P_D),$$

Cardinality constraint

$$y_i(t)(1 - y_i(t)) = 0.$$

Quadratic equations

Event-triggered Chiller Loading (cont'd)

The triggering instant $t_{m+1} = \min\{\tau > t_m \mid |P_D(\tau) - P_D(t_m)| \geq \sigma_e P_D(t_m)\}$

Method	$P_D(t)$ (kW)					Total power consumption (kW)	Total energy consumption (kJ)
	6858	6477	6096	5717	5334		
GAMS [8]	**4738.5753**	**4421.6486**	**4143.7064**	**3842.5532**	**3546.4375**	**20692.9210**	**74494515.6000**
AVL [2]	4916.93	4635.22	4358.71	4087.20	3821.34	21819.40	78549840.00
GA [2]	4766.33	4459.16	4185.87	3940.60	3706.22	21058.18	75809448.00
SA [3]	4776.98	4453.64	4178.80	3925.16	3675.18	21009.76	75635136.00
BGA [4]	4744.6512	4445.3493	4172.6155	3911.0079	3694.0319	20967.6558	75483560.8800
CGA [4]	4738.9645	4430.2444	4147.8178	3907.7607	3686.8597	20911.6471	75281929.5600
PSO [4]	4739.7845	4423.0534	4147.8055	3920.9642	3642.5786	20874.1862	75147070.3200
ES [5]	4738.76	4422.06	4144.12	3906.19	3627.46	20838.59	75018924.00
DCEDA [27]	4738.6600	4421.6739	4143.7297	3842.6020	3546.6437	20693.3093	74495913.4800
IFA [6]	4738.576	4421.649	**4143.706**	3960.573*	3627.760*	20892.264	75212150.400
DCSA [7]	**4738.575**	4421.649	**4143.706**	3960.560*	3627.987*	20892.477	75212917.200
ET-CNO-LD	**4738.5753**	**4421.6486**	**4143.7064**	**3842.5532**	**3546.4375**	**20692.9210**	**74494515.6000**

Z. Chen, J. Wang, and Q.L. Han, "Event-triggered cardinality-constrained cooling and electrical load dispatch based on collaborative neurodynamic optimization," *IEEE Trans. Neural Networks and Learning Systems*, vol. 34, pp. 5464-5475, 2023.

Receding-Horizon Chiller Loading

OCL planning over a multi-period

- Chillers should **NOT be frequently switched on or off** to avoid excess attrition and prolong the lifespan of chillers
- Once a chiller is switched on or off, it should be kept on its current on/off status for a while to warm up or cool down (**minimum-up/down-time constraints**)
- In most existing formulations, the minimum-up/down-time constraints are absent

$$\min_{PLR,y} \sum_{t=t_0}^{t_0+T-1} \sum_{i=1}^{n} P_i(PLR_i(t))y_i(t)$$

Total power consumption

$$\text{s.t.} \sum_{i=1}^{n} \overline{P}_i PLR_i(t) y_i(t) = P_D(t), \quad t \in \{t_0, \ldots, t_0+T-1\},$$

Supply–demand constraints

$$\sum_{\tau=1}^{\tau_{\text{on}}} (y_i(t+\tau-1) - (y_i(t) - y_i(t-1))) \geq 0, \quad i \in \{1,\ldots,n\}, t \in \{t_0 - T_{\text{on}}+1, \ldots, t_0+T-1\},$$

$$\sum_{\tau=1}^{\tau_{\text{off}}} (y_i(t+\tau-1) - 1 - (y_i(t) - y_i(t-1))) \leq 0, \quad i \in \{1,\ldots,n\}, t \in \{t_0 - T_{\text{off}}+1, \ldots, t_0+T-1\},$$

Minimum-up/down-time constraints

$$\underline{PLR}_i \leq PLR_i(t) \leq \overline{PLR}_i, \quad i \in \{1,\ldots,n\}, t \in \{t_0, \ldots, t_0+T-1\},$$

Capacity constraints

$$y_i(t) \in \{0,1\}, \quad i \in \{1,\ldots,n\}, t \in \{t_0, \ldots, t_0+T-1\},$$

Binary constraints

Receding-Horizon Chiller Loading (cont'd)

Subproblem (1)

$$\min_{PLR} \sum_{t\in\mathcal{I}_{\text{on}}} \sum_{i\in\mathcal{I}_{\text{on}}} P_i(PLR_i(t))$$

$$\text{s.t.} \sum_{i\in\mathcal{I}_{\text{on}}} \overline{P}_i PLR_i(t) = P_D(t),$$

$$t \in \{t_0, \ldots, t_0+T-1\},$$

$$\underline{PLR}_i \le PLR_i(t) \le \overline{PLR}_i; i, t \in \mathcal{I}_{\text{on}}.$$

Subproblem (2)

$$\min_{y} \sum_{t=t_0}^{t_0+T-1} \sum_{i=1}^{n} P_i(PLR_i(t)) y_i(t)$$

$$\text{s.t. } P_D(t) - \sum_{i=1}^{n} \overline{P}_i PLR_i(t) y_i(t) \le 0,$$

$$t \in \{t_0, \ldots, t_0+T-1\},$$

$$\sum_{\tau=1}^{\mathcal{T}_{\text{on}}} (y_i(t+\tau-1) - (y_i(t) - y_i(t-1))) \ge 0,$$

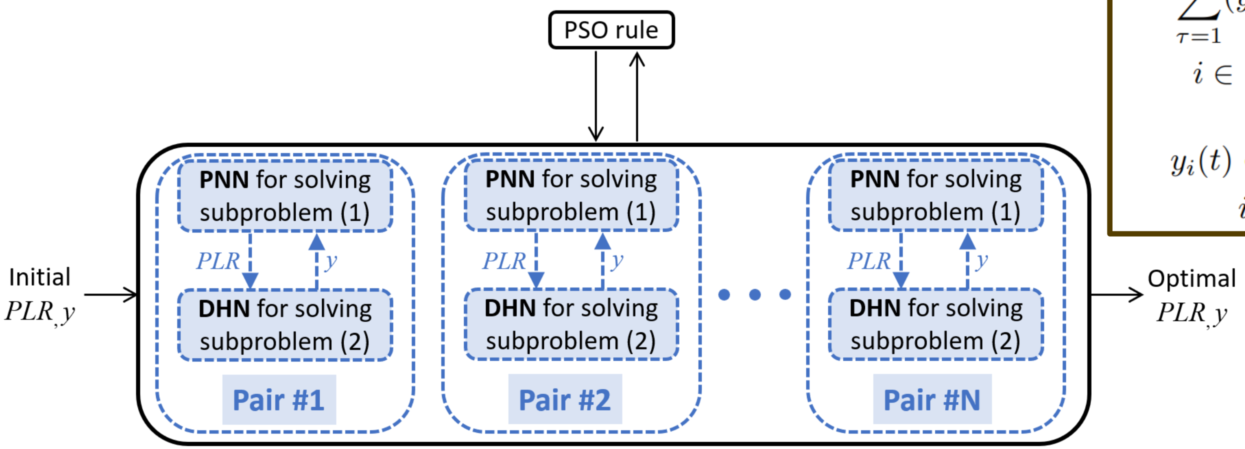
$$i \in \{1, \ldots, n\}, t \in \{t_0 - T_{\text{on}} + 1, \ldots, t_0+T-1\},$$

$$\sum_{\tau=1}^{\mathcal{T}_{\text{off}}} (y_i(t+\tau-1) - 1 - (y_i(t) - y_i(t-1))) \le 0,$$

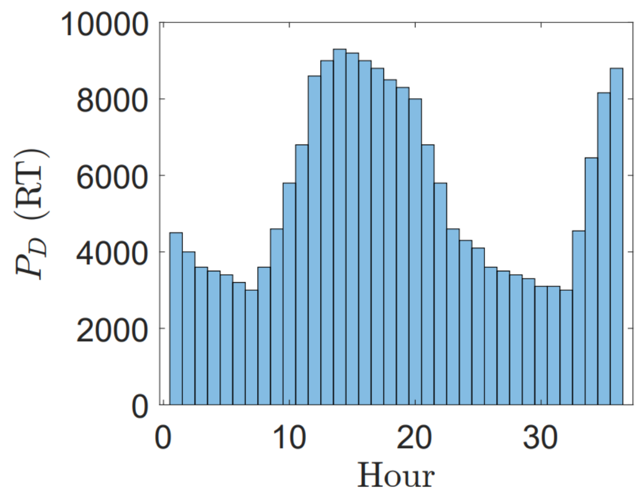
$$i \in \{1, \ldots, n\}, t \in \{t_0 - T_{\text{off}} + 1, \ldots, t_0+T-1\},$$

$$y_i(t) \in \{0, 1\},$$

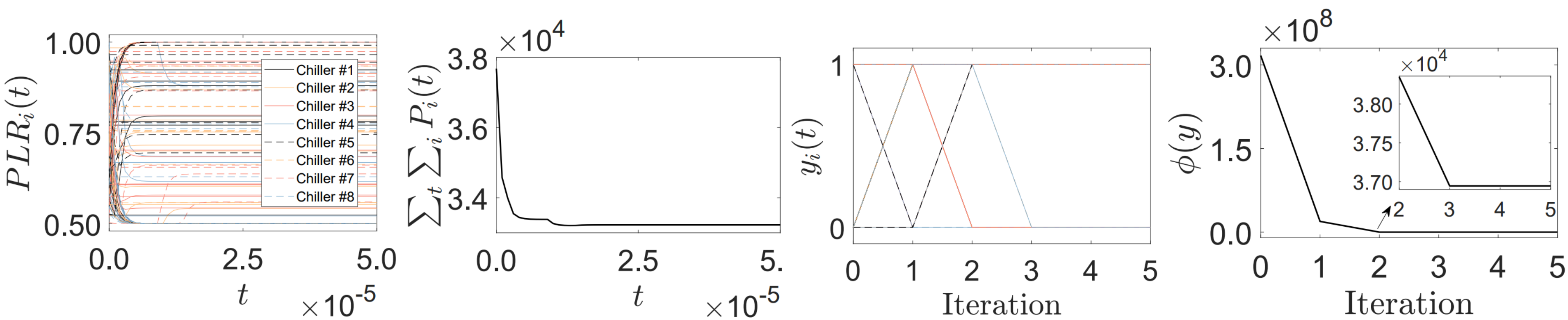
$$i \in \{1, \ldots, n\}, t \in \{t_0, \ldots, t_0+T-1\}.$$



Receding-Horizon Chiller Loading (cont'd)

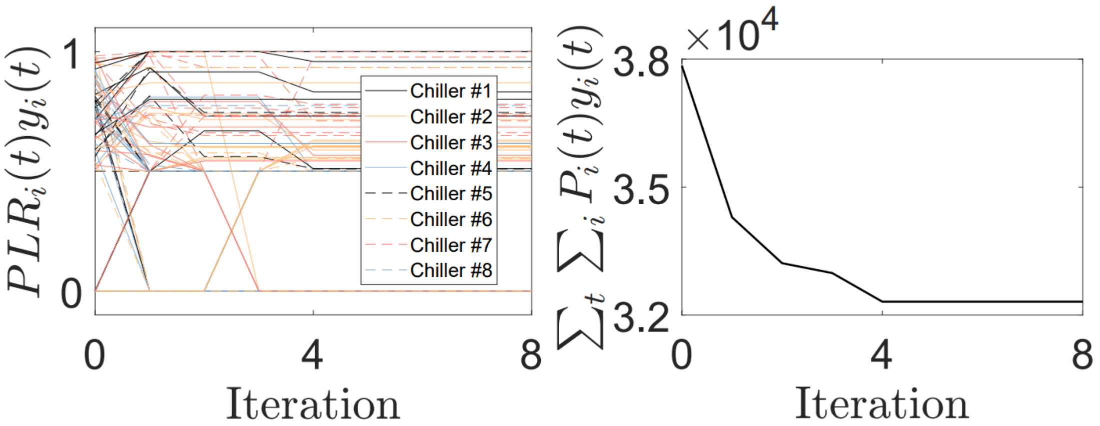


Cooling loads

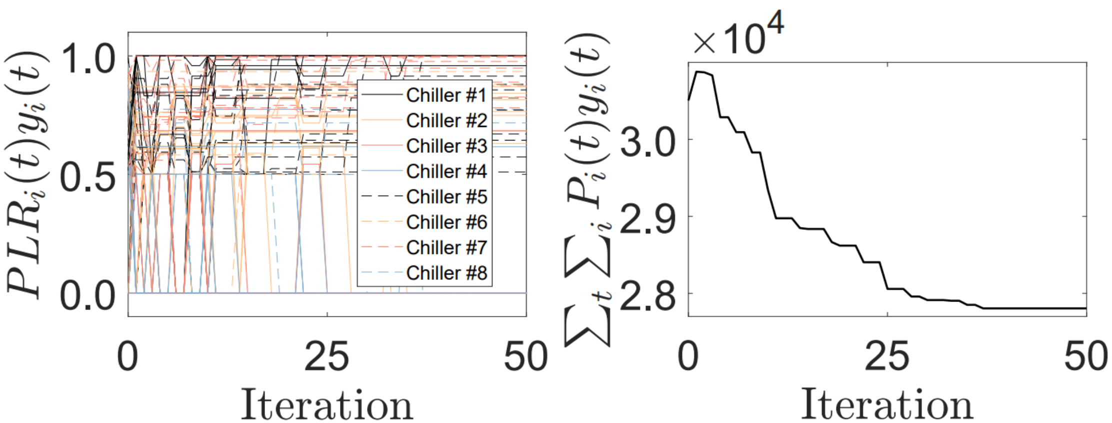


Convergent Behaviors (inner loop)

Receding-Horizon Chiller Loading (cont'd)



Convergent Behaviors (middle loop)



Convergent Behaviors (outer loop)

Receding-Horizon Chiller Loading (cont'd)

$T_{\text{on}} = 3$ hours
$T_{\text{off}} = 2$ hours

Hour	1	2	3	4	5	6	7	8
Chiller #1	1.00	1.00	0.96	1.00	1.00	1.00	1.00	0.96
Chiller #2	0.75	0.68	0.63	0.80	0.74	0.74	0.70	0.63
Chiller #3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chiller #4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chiller #5	0.86	0.64	0.51	0.00	0.00	0.82	0.70	0.51
Chiller #6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chiller #7	0.99	0.87	0.78	1.00	0.98	0.00	0.00	0.78
Chiller #8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\sum_i P_i(t)y_i(t)$	2187.47	1922.78	1751.20	1661.65	1599.51	1611.26	1504.17	1751.20

Hour	9	10	11	12	13	14	15	16
Chiller #1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Chiller #2	0.77	0.82	0.84	0.87	0.92	0.97	0.95	0.92
Chiller #3	0.00	0.00	0.00	0.68	0.73	0.76	0.75	0.73
Chiller #4	0.00	0.00	0.00	0.62	0.69	0.76	0.74	0.69
Chiller #5	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Chiller #6	0.00	0.82	0.88	0.93	1.00	1.00	1.00	1.00
Chiller #7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Chiller #8	0.00	0.00	0.72	0.77	0.87	0.95	0.92	0.87
$\sum_i P_i(t)y_i(t)$	2248.00	3056.32	3766.62	5211.04	5586.61	5912.30	5798.74	5586.61

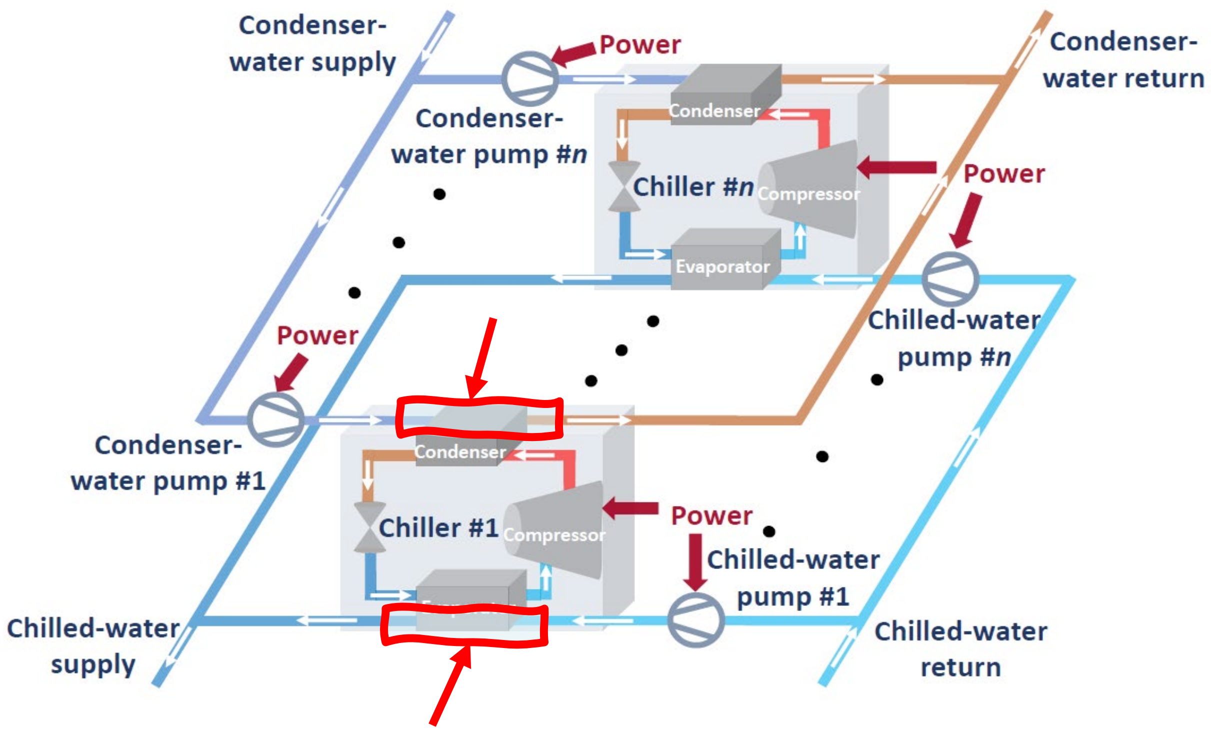
Hour	17	18	19	20	21	22	23	24
Chiller #1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Chiller #2	0.89	0.86	0.84	0.89	0.84	0.87	0.77	0.72
Chiller #3	0.70	0.68	0.66	0.71	0.00	0.00	0.00	0.00
Chiller #4	0.65	0.60	0.56	0.00	0.00	0.00	0.00	0.00
Chiller #5	1.00	1.00	1.00	1.00	1.00	1.00	0.91	0.77
Chiller #6	0.97	0.91	0.87	0.98	0.88	0.00	0.00	0.00
Chiller #7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95
Chiller #8	0.82	0.75	0.71	0.82	0.72	0.77	0.00	0.00
$\sum_i P_i(t)y_i(t)$	5392.19	5125.03	4961.54	4734.01	3766.62	3100.12	2248.00	2074.42

Z. Chen, J. Wang, and Q.L. Han, "Receding-Horizon Chiller Operation Planning via Collaborative Neurodynamic Optimization," *IEEE Trans. on Smart Grid*, vol. 15, no. 2, pp. 2321-2331, 2024.

Hybrid Model Predictive Control

Motivation

- Existing model-based control schemes do NOT consider thermodynamics in chillers, making them less realistic



Hybrid Model Predictive Control (cont'd)

Min Total power consumption

s.t. Supply–demand constraints

Energy-conservation constraints

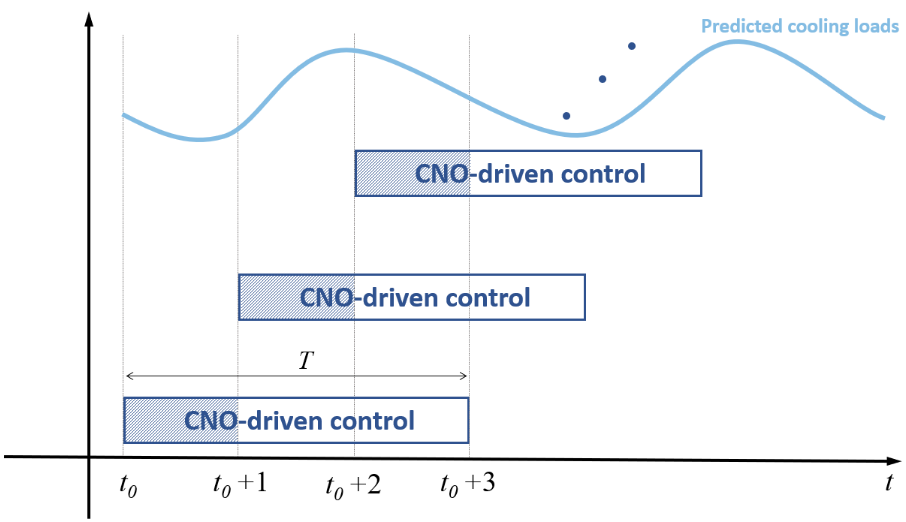
Thermodynamics constraints

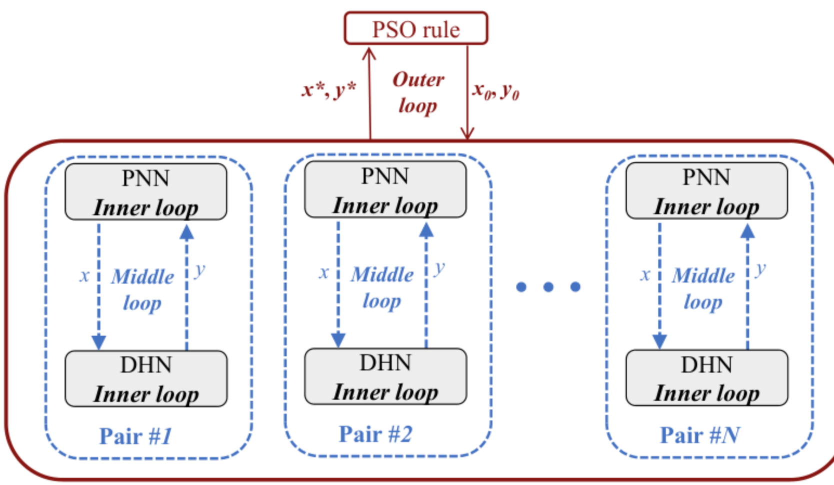
Bound constraints

Binary constraints

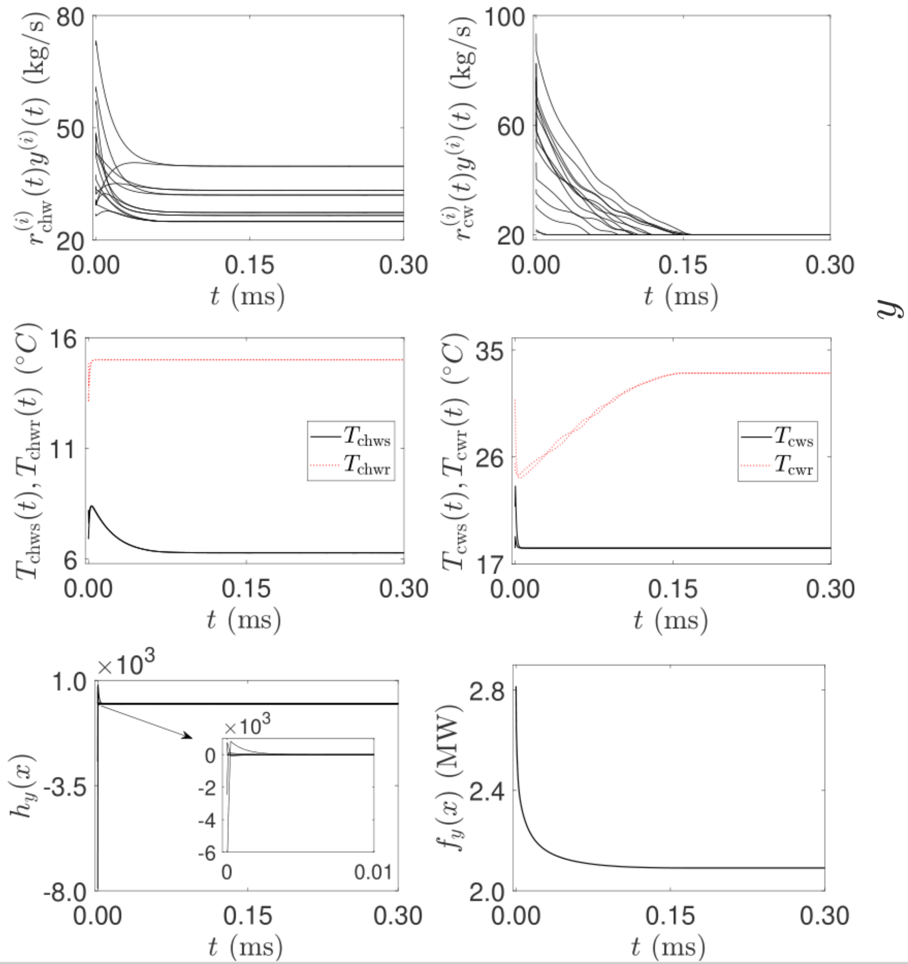
$$\rho_w C_p \sum_{i=1}^{n} V_e^{(i)} \frac{dT_{\text{chws}}}{dt} = \sum_{i=1}^{n} r_{\text{chw}}^{(i)} y^{(i)} C_p (T_{\text{chwr}} - T_{\text{chws}}) - P_D,$$

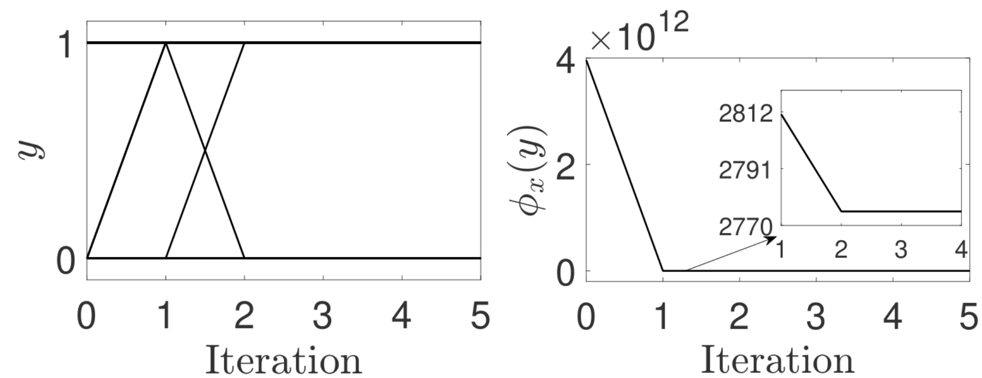
$$\rho_w C_p \sum_{i=1}^{n} V_c^{(i)} \frac{dT_{\text{cwr}}}{dt} = \sum_{i=1}^{n} r_{\text{cw}}^{(i)} y^{(i)} C_p (T_{\text{cwr}} - T_{\text{cws}}) - P_D - \sum_{i=1}^{n} P_{\text{ch}}^{(i)}(\cdot) y^{(i)},$$





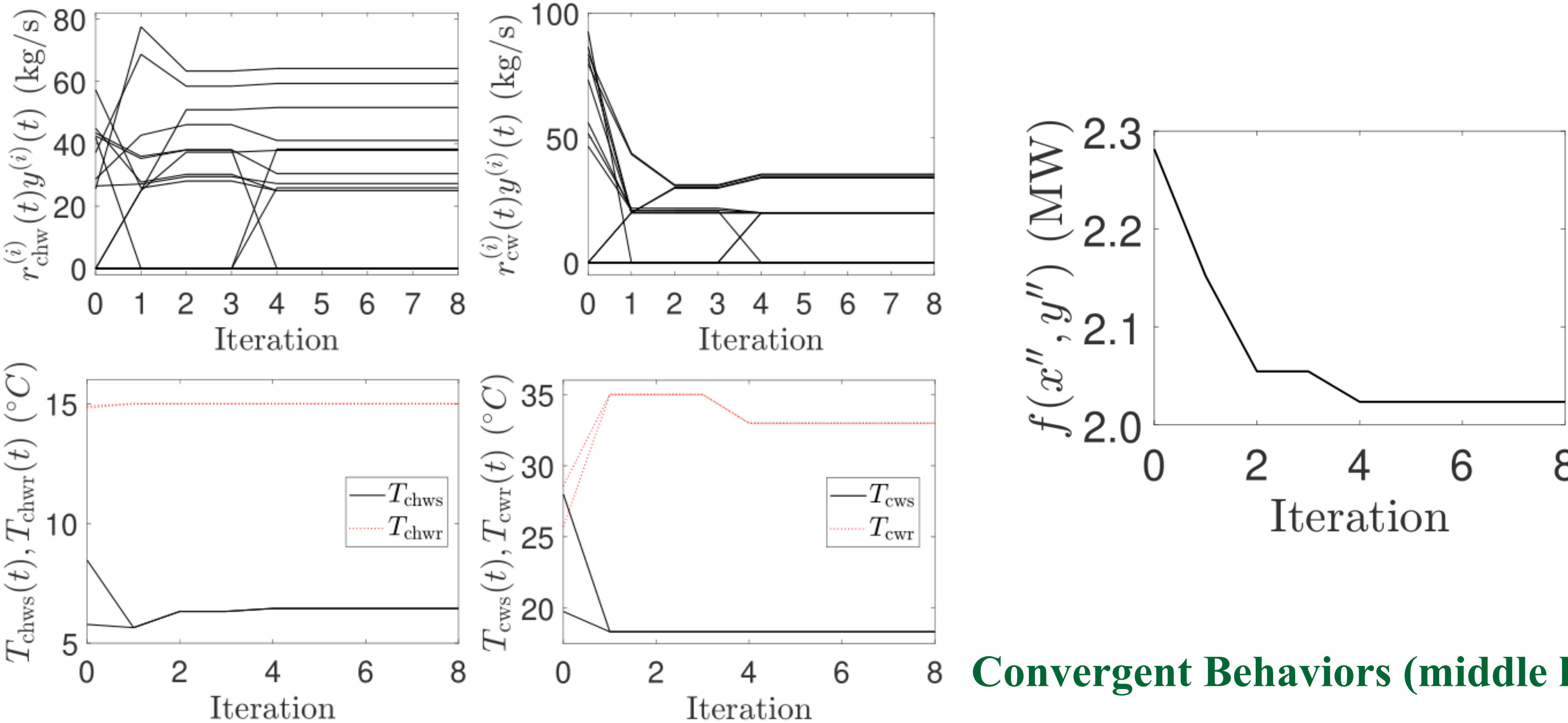
Hybrid Model Predictive Control (cont'd)





Convergent Behaviors (inner loop)

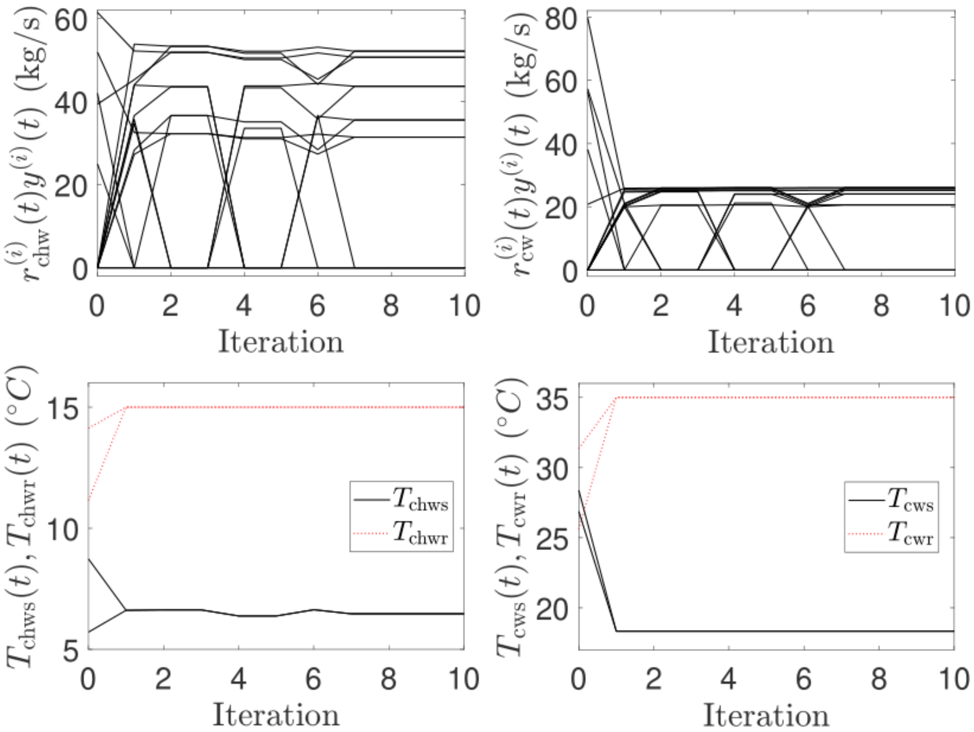
Hybrid Model Predictive Control (cont'd)

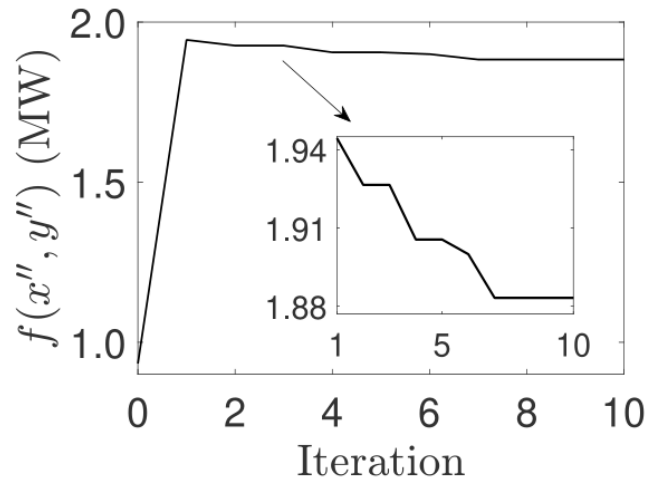


Convergent Behaviors (middle loop)

Z. Chen, J. Wang, and Q.L. Han, "Hybrid model predictive control of chiller systems via collaborative neurodynamic optimization," *IEEE Transactions on Industrial Informatics,* in press.

Hybrid Model Predictive Control (cont'd)

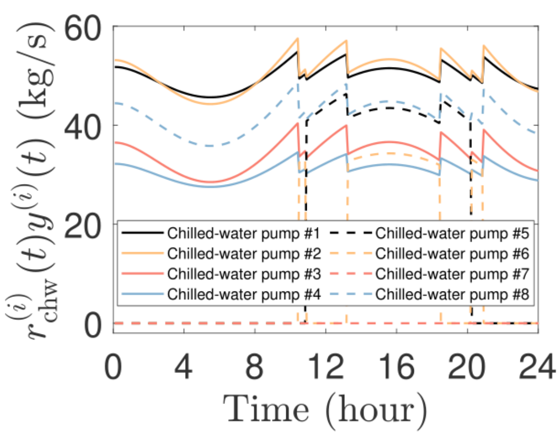


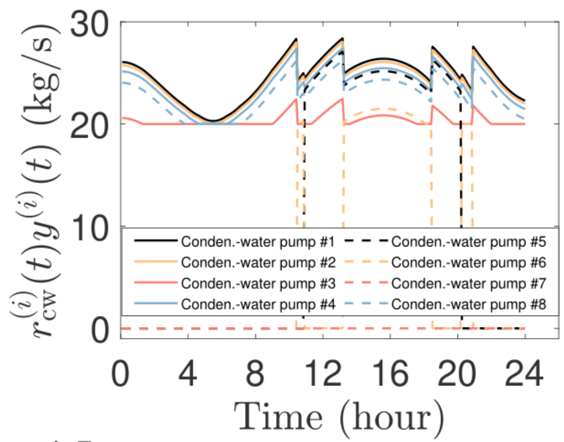


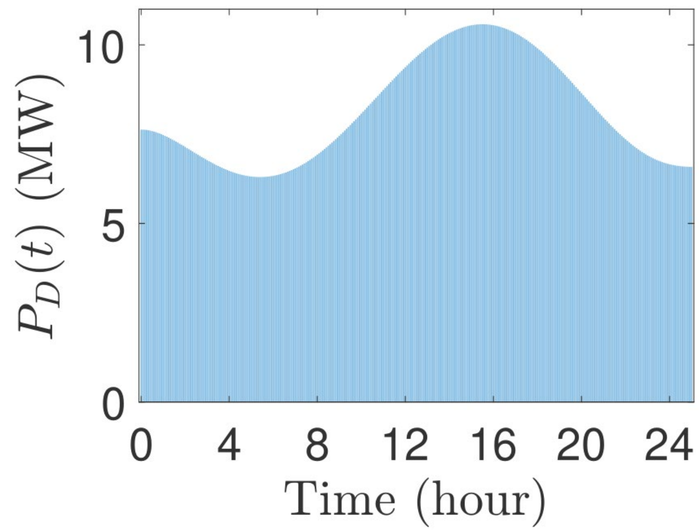
Convergent Behaviors (outer loop)

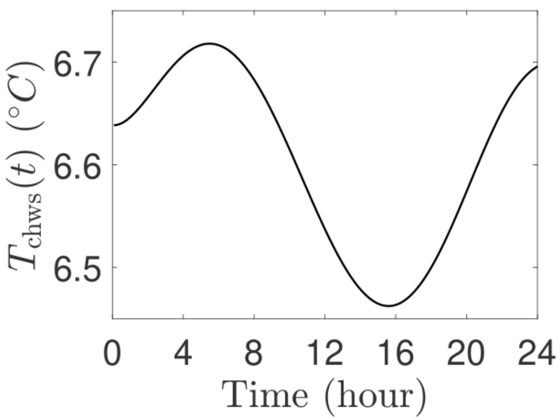
Z. Chen, J. Wang, and Q.L. Han, "Hybrid model predictive control of chiller systems via collaborative neurodynamic optimization," *IEEE Transactions on Industrial Informatics,* in press.

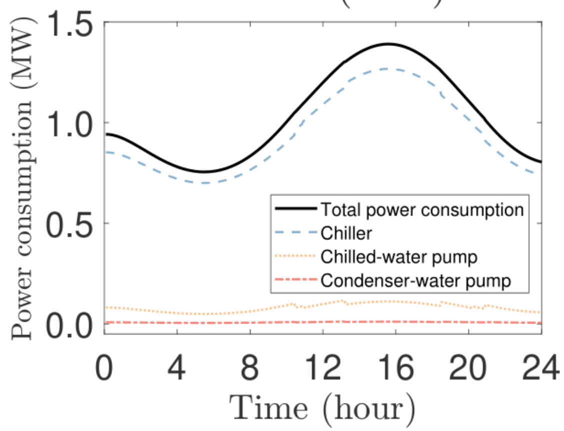
Hybrid Model Predictive Control (cont'd)











Within 24 hours:

- **control variables**
- **chilled-water supply temperature**
- **power consumption function values**

Z. Chen, J. Wang, and Q.L. Han, "Hybrid model predictive control of chiller systems via collaborative neurodynamic optimization," *IEEE Transactions on Industrial Informatics,* in press.

Concluding Remarks

- Collaborative neurodynamic optimization is a biologically and socially plausible approach.
- It has the desirable properties of almost-sure global convergence.
- It serves as a bridge between neurodynamic optimization and other natural-inspired optimization methods toward hybrid intelligence.
- Collaboration is the key to success.
- It plays an instrumental role in many applications where optimization problem-solving is imperative.