Journal of Beijing Institute of Technology, 2026, Vol. 35, No. 2

Robust Graph Diffusion for Multi-Task Learning in Ultra-High Voltage Direct Current Monitoring Systems

Shaojie Wang[1], Chuliang Xue[1], Hui Li[1], Rui Zhu[1], Xiaoping Ren[2], Limei Hu[2]✉

(1. State Grid Shanghai Direct Current Company, Shanghai 200235, China; 2. College of Artificial Intelligence, Southwest University, Chongqing 400715, China)

Abstract: In this work, we investigate the problem of multi-task learning (MTL) in ultra-high voltage direct current (UHVDC) monitoring systems. Considering the measurements are affected by wireless channel impairments, typically characterized by block fading and link noise. Such channel imperfections significantly degrade the performance of distributed estimation in real-world power system environments. Based on the graph signal processing method, we propose the multi-task robust decoupled diffusion least mean square algorithm (MT-RDDLMS). Specifically, a decoupled adapt-then-combine strategy is introduced to reduce the influence of wireless channels on data exchange among measurement units. Moreover, an average estimation method with an adaptive smoothing factor is developed to further suppress link noise and enhance estimation accuracy. Simulation results confirm the robustness and effectiveness of the proposed algorithm under realistic wireless channel conditions.

Keywords: graph signal processing; distributed multitask learning; ultra-high voltage direct current (UHVDC) systems; average estimation

1 Introduction

With the rapid development of fields such as smart grids, social networks [1] and wireless sensor networks [2], graph signal processing (GSP) has received increasing attention. Graph signals are typically defined on the nodes of a graph and are characterized by the topological structure of the graph. These signals represent data associated with each node over the network [3]. GSP provides a set of tools for analyzing and processing signals on graphs, primarily through the graph shift operator (GSO) and the graph Fourier transform (GFT).

As traditional signal processing theories are increasingly applied in the graph domain, numerous graph filter structures have been proposed in [4–6]. Concurrently, adaptive filtering algorithms have been extended to the field of graph signal processing, including diffusion least mean square (LMS) [7], diffusion preconditioned LMS (PLMS) [8], and diffusion recursive least square (RLS) [9]. These algorithms leverage dynamic streaming graph signals for distributed adaptive estimation. Depending on the relationships among the target parameter vectors, distributed estimation tasks can be categorized into single-task learning [10] and multi-task learning (MTL) [11,12]. In single-task learning, all nodes estimate the same parameter vector, whereas in MTL, the network is divided into clusters, with nodes within each cluster learning the same target parameters. The nodes cooperate by exploiting the similarities between tasks to improve net-

Manuscript received Oct. 18, 2025; revised Nov. 12, 2025; accepted Nov. 30, 2025. The associate editor coordinating the review of this manuscript was Dr. Lijuan Jia.
✉ Corresponding author. Email: hlm0903@swu.edu.cn
DOI: 10.15918/j.jbit1004-0579.2025.071

work performance. For example, in ultra-high voltage direct current (UHVDC) systems, distributed measurement boards can be deployed to monitor electrical parameters such as voltage and current across converter stations. Each board leverages the similarity between its own measurements and those of neighboring boards to collaboratively estimate the operational state [13].

It is important to note that in many practical applications, such as wireless sensor networks and smart grids, data transmission over wireless channels is subject to imperfections, including block fading and link noise [14]. These impairments can distort the information exchanged between neighboring nodes, degrading the estimation accuracy and slowing the convergence of the algorithms [14–16]. The single-task learning problem with wireless channels has been studied in [14], which explores the impact of wireless channels on the performance of distributed estimation. In [17], the multi-task learning problem in wireless channels was considered. In [18] and [12], the authors considered that the quality of information can affect the performance of the estimate. The authors proposed a smart cooperation strategy based on an information-theoretic criterion, which excludes cooperation between nodes with different tasks or poor-quality links. In [15], a fixed moving average filter is applied to process imprecise exchange weights caused by link noise, which improved estimation accuracy but at the expense of slower convergence.

In particular, the measurement boards in UHVDC are important in acquiring data for real-time equipment monitoring. The quality of measurements directly determine the system safety through their operational stability. With the widespread deployment of UHVDC monitoring systems, the application of graph-based signal analysis has become crucial for monitoring and maintenance. The current maintenance model, which relies on manual inspection and post-fault repair, suffers from deficiencies such as the lack of early fault identification, strong dependence on expert experience, and high risks of sudden failures. Therefore, it is essential to build an intelligent predictive maintenance system based on advanced algorithms.

In this work, we model the UHVDC monitoring systems as an undirected graph, and the problem of multi-task learning in graph signal processing under wireless channel impairments is considered, with a specific focus on applications in UHVDC measurement board monitoring. To the best of our knowledge, this has not been explored in the existing literature. To tackle this problem, a robust graph decoupled diffusion multi-task LMS algorithm called MT-RDDLMS is proposed. The main contributions of this paper are as follows:

1) To reduce the degradation of estimation performance caused by wireless channels in UHVDC monitoring systems, the decoupled adapt-then-combine method is adopted into graph signal multi-task learning.

2) To suppress the impact of link noise on the estimation process, an average estimation method with an adaptive smoothing factor is proposed.

3) Numerical simulations, including scenarios inspired by real-world UHVDC signal processing, demonstrate the robustness and effectiveness of the proposed MT-RDDLMS algorithm in wireless channel environments.

The structure of this work is organized as follows. Section 2 presents the preliminaries and problem statement. In Section 3, a robust graph-decoupled diffusion multi-task LMS algorithm is proposed. Section 4 provides the simulation results.

2 System Model and MT-DLMS Algorithm

2.1 System Model

In typical UHVDC monitoring systems, the physical components like converter stations, transmission towers, and the measurement boards in UHVDC can be represented by the nodes in graph. The connections between power lines and communication links can be characterized by edges. An undirected graph is appropriate because these connections are inherently bidirectional; a fault or signal in one component can directly influence another, regardless of direction. This structure effectively captures the system's mutual dependencies and facilitates analysis. Therefore, an undirected graph $\mathcal{G}=(\mathcal{J},\mathcal{E})$ with vertex set $\mathcal{J}$ and edge set $\mathcal{E}$. If $(i,j)\in\mathcal{E}$, nodes i and j are neighbors, meaning they can communicate with each other. The neighborhood of node k, denoted by $\mathcal{N}_k$, includes node k itself. Additionally, $\mathcal{N}_k^-$ indicates that it does not include node k itself.

The input graph signal $\boldsymbol{x}(i)$ on $\mathcal{G}$ is defined as a mapping from vertex set $\mathcal{J}$ at time i, where $\boldsymbol{x}(i)=[x_1(i),\cdots,x_k(i)]^T$ and $z_k(i)$ represents the signal value of node k at time i. The GSO is defined as a weighted combination of the neighborhood, and the GSO at time step i is denoted by $\boldsymbol{S}^i$. The shift-invariant graph filter is defined by

$$\boldsymbol{H}=\sum_{m=0}^{L-1}h_m^o\boldsymbol{S}^m \tag{1}$$

where $\boldsymbol{h}^o=[h_0,h_1,\cdots,h_{L-1}]^{\mathrm{T}}$ denotes the coefficient vector of the graph filter, and L is the length of the graph filter. Let $\boldsymbol{Z}(i)$ represent the $N\times M$ matrix, which is given by

$$\boldsymbol{Z}(i)=\left[x(i),\boldsymbol{S}\boldsymbol{x}(i-1),\cdots,\boldsymbol{S}^{L-1}\boldsymbol{x}(i-L+1)\right]^{\mathrm{T}} \tag{2}$$

The filtered graph signal $\boldsymbol{y}(i)$ and the local filtered graph signal $\boldsymbol{y}_k(i)$ are respectively expressed as

$$\boldsymbol{y}(i)=\boldsymbol{Z}(i)\boldsymbol{h}^o+\boldsymbol{v}(i)=\boldsymbol{H}\boldsymbol{x}(i)+\boldsymbol{v}(i) \tag{3}$$

$$y_k(i)=\boldsymbol{z}_k(i)^{\mathrm{T}}\boldsymbol{h}_k^o+v_k(i) \tag{4}$$

where $\boldsymbol{v}(i)=[v_1(i),v_2(i),\cdots,v_N(i)]^{\mathrm{T}}$ denotes an independent and identically zero mean noise with variance $\sigma_{v,k}^2$ at each node k. In (4), $z_k(i)^{\mathrm{T}}$ represents the regression vector at node k, which is given by

$$\begin{aligned}\boldsymbol{z}_k(i)^{\mathrm{T}}=&\mathrm{col}\{[\boldsymbol{x}(i)]_k,[\boldsymbol{S}\boldsymbol{x}(i-1)]_k,\cdots,\\&[\boldsymbol{S}^{L-1}\boldsymbol{x}(i-L+1)]_k\}\end{aligned} \tag{5}$$

where $\boldsymbol{z}_k(i)^{\mathrm{T}}$ is the k-th row of $\boldsymbol{Z}(i)$.

2.2 MT-DLMS Algorithm

In a multi-task graph filter model in UHVDC, the nodes in the network are divided into Q clusters. Nodes within the same cluster learn the same graph filter coefficient, namely

$$h_k^{\circ}=h_{(\alpha)}^{\circ},\ \text{for each node } k\in\mathcal{C}_\alpha \tag{6}$$

where $\mathcal{C}_\alpha$ represents the α-cluster, in α-cluster, there is a filter coefficient vector $h_{(\alpha)}^{\circ}$ to estimate.

According to [8], MT-DLMS employs the well-known adapt-then-combine (ATC) strategy to minimize the objective function in a distributed manner, it can be written as follows

$$\boldsymbol{\psi}_k(i)=\boldsymbol{h}_k(i-1)+\mu_k\boldsymbol{z}_k(i)\left[y_k(i)-\boldsymbol{z}_k^{\mathrm{T}}\boldsymbol{h}_k(i-1)\right] \tag{7}$$

$$\boldsymbol{h}_k(i)=\sum_{\ell\in\mathcal{N}_{\|}}a_{\ell k}\boldsymbol{\psi}_\ell(i) \tag{8}$$

where μ_k is the step size for node k, $a_{\ell k}$ is the combination coefficient. The combination coefficient must satisfy the following constraint

$$\sum_{l=1}^{N}a_{\ell k}=1,a_{\ell k}\geqslant 0,a_{\ell k}=0,\ \text{if}\,\ell\notin\mathcal{N}_{\|} \tag{9}$$

After the clustering step, the combination coefficient $a_{\ell k}$ is non-zero only if node k and its neighboring node ℓ are assigned to the same cluster.

3 Robust Graph Decoupled Diffusion Multi-Task LMS Algorithm

To account for the impact of the wireless channel in UHVDC, a robust multi-task learning strategy is proposed. Specifically, we incorporate the decoupled adapt-then-combine method [19] into graph signal processing, resulting in the following formulation

$$\boldsymbol{\psi}_k(i) = \boldsymbol{\psi}_k(i-1) + \mu_k \boldsymbol{z}_k(i)\left[y_k(i) - \boldsymbol{z}_k^{\mathrm{T}}(i)\boldsymbol{\psi}_k(i-1)\right] \tag{10}$$

$$\boldsymbol{h}_k(i) = a_{kk}(i)\boldsymbol{\psi}_k(i) + \sum_{\ell \in \mathcal{N}_k^-} a_{\ell k}(i)\boldsymbol{h}_\ell(i-1) \tag{11}$$

Compared with (7), $\boldsymbol{h}_k(i-1)$ in (10) is replaced by $\boldsymbol{\psi}_k(i-1)$, which means that the combined graph filter weight is not feedback into the adaptation stage. This feedback can be beneficial in certain scenarios, such as an ideal case without link noise. In (10), each node functions as an independent filter during the adaptive stage and does not receive information from its neighbors. Consequently, the adaptive stage is not influenced by the wireless channel. In addition, this design helps prevent the algorithm from being affected by inaccuracies in the multi-task learning process.

Since it takes time for nodes to transmit information to their neighbors, the combined information in (11) consists of the intermediate estimate of node k at the current time i, along with the graph filter coefficients of its neighbors from the previous time $i-1$.

When nodes exchange information with wireless channels, the process is affected by additive noise and block fading, as illustrated in Fig. 1. The graph filter coefficient received by node k from its neighboring node ℓ is given as follows

$$\boldsymbol{h}_{\ell,k}(i) = \boldsymbol{g}_{\ell,k}(i)\boldsymbol{h}_\ell(i) + \boldsymbol{q}_{\ell,k}(i), \quad \ell \in \mathcal{N}_k^- \tag{12}$$

where $\boldsymbol{g}_{\ell,k}(i)$ represents the fading coefficient in the wireless channel, it remains constant within each data block [17]. The additive noise in the wireless channel is denoted by $\boldsymbol{q}_{\ell,k}(i)$, which representing the link noise between node ℓ and node k at time i.

To facilitate the analysis and explanation of the algorithm. The following assumptions for the wireless channel are adopted, which are widely accepted in [14, 16, 17].

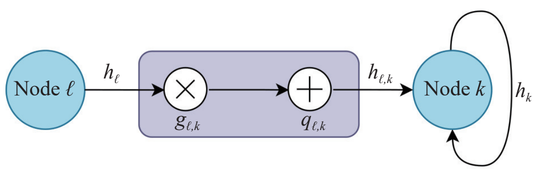

Fig. 1 Node k receives data from neighbor node ℓ

Assumption 1 The link noise signals $\boldsymbol{q}_{\ell,k}(i)$ are spatially and temporally independent zero-mean Gaussian random variables, and are uncorrelated with other signals.

Assumption 2 The fading coefficients $\boldsymbol{g}_{\ell,k}(i)$ are zero-mean Gaussian random variables, and they are independent and identically distributed over time and space.

Assumption 3 For all nodes and all time instances, the fading coefficient $\boldsymbol{g}_{\ell,k}(i)$ and the link noise $\boldsymbol{q}_{\ell,k}(i)$ are mutually independent.

It should be noted that in practical networks, the fading coefficient $\boldsymbol{g}_{\ell,k}(i)$ is not known in advance [14, 17]. Therefore, nodes acquire their own channel state information and compensate for the fading effect using local equalizer gains, which could be given by

$$\gamma_{\ell,k}(i) = \frac{\boldsymbol{g}_{\ell,k}(i)^*}{|\boldsymbol{g}_{\ell,k}(i)|^2} \tag{13}$$

After compensation using the local equalizer gains $\gamma_{\ell,k}(i)$, (11) can be rewritten as

$$\begin{aligned}\boldsymbol{h}'_{\ell,k}(i) =& \gamma_{\ell,k}(i)\boldsymbol{g}_{\ell,k}(i)\boldsymbol{h}_\ell(i) + \gamma_{\ell,k}(i)\boldsymbol{q}_{\ell,k}(i) = \\ & \boldsymbol{h}_\ell(i) + \boldsymbol{q}'_{\ell,k}(i)\end{aligned} \tag{14}$$

To mitigate the impact of link noise, we use the well-known moving average method [20] to denoise the transmitted data. In (14), $\boldsymbol{h}'_{\ell,k}(i)$ is replaced by the denoised estimate $\hat{\boldsymbol{h}}_{\ell,k}(i)$, which is given by

$$\hat{\boldsymbol{h}}_{\ell,k}(i) = (1-\beta_{k,i})\hat{\boldsymbol{h}}_{\ell,k}(i-1) + \beta_{k,i}\boldsymbol{h}'_{\ell,k}(i) \tag{15}$$

where $\beta_{k,i}$ is the smoothing factor, and $0 < \beta_{k,i} << 1$. When $\beta_{k,i}$ is close to 0, $\hat{\boldsymbol{h}}_{k,l}(i)$ provides limited information in initial state [16]. Thus, in this work, $\beta_{k,i}$ is designed as an adaptive smoothing factor.

Specifically, the sigmoid function is used to constrain $\beta_{k,i}$ within the range (0, 1), and is defined as

$$\beta_{k,i}=\frac{1}{1+e^{-\theta_{k,i}}} \tag{16}$$

where $\theta_{k,i}$ is an auxiliary variable to facilitate adaptive updates based on the following loss function

$$J_n^{\mathrm{MSE}}=\frac{1}{2}\sum_{n=1}^{N}\mathbb{E}\{(y_k(i)-\boldsymbol{z}_k^{\mathrm{T}}(i)\boldsymbol{h}_k(i-1))\} \tag{17}$$

By minimizing (17) using the gradient descent method, the update rule for $\theta_{k,i}$ is derived as

$$\begin{aligned}\theta_{k,i}=&\theta_{k,i-1}+\mu_{\theta_k}\beta_{k,i-1}(1-\beta_{k,i-1})\times\\&(y_k(i)-\boldsymbol{z}_k^{\mathrm{T}}(i)\boldsymbol{h}_k(i-1))\times\\&\boldsymbol{z}_k^{\mathrm{T}}(i)(\boldsymbol{p}_k(i-1)-\boldsymbol{\psi}_k(i-2))\end{aligned} \tag{18}$$

where μ_{θ_k} is a step-size parameter and

$$\boldsymbol{p}_k(i-1)=\boldsymbol{h}_k(i-1)-a_{kk}(i-1)\boldsymbol{\psi}_k(i-1) \tag{19}$$

To ensure the continuous learning of (18), the value of $\theta_{k,i}$ is constrained within the interval $[-v^{+},v^{+}]$ [16].

In this work, nodes are assumed to have no prior knowledge of the underlying clusters. In other words, nodes cannot determine which neighbors share the same learning objective. Following the approach in [8, 11], we facilitate clustering by computing the normalized ℓ_2-norm distance between the graph filter coefficients of node k and its neighboring node $\ell\in\mathcal{N}_k$. To enable online identification of clustering structure, we introduce the following Boolean variable

$$d_{\ell,k}(i)=\begin{cases}1, & \text{if } \dfrac{\|\hat{\boldsymbol{h}}_\ell(i-1)-\boldsymbol{h}_k(i-1)\|^2}{\|\boldsymbol{h}_k(i-1)\|^2}\leqslant\eta,\ \ell\in\mathcal{N}_k\\ 0, & \text{otherwise}\end{cases} \tag{20}$$

where $\eta\in(0,1)$ is a predefined similarity-related threshold. If the l_2-norm distance between the graph filter estimates of node k and node ℓ is less than η, the two nodes will be assigned to the same cluster. It is important to note that the l_2-norm distance is calculated between $\hat{\boldsymbol{h}}_\ell(i-1)$ and $\boldsymbol{h}_k(i-1)$. The reason for choosing $\hat{\boldsymbol{h}}_\ell(i-1)$ is that the denoised information leads to better clustering. The reason for choosing $\boldsymbol{h}_k(i-1)$ is that it combines the local estimate $\boldsymbol{\psi}_k(i-1)$ with the information $\hat{\boldsymbol{h}}_\ell(i-2)$ from its neighbors.

To further reduce the impact of link noise, a convex combination strategy is employed as follows

$$t_{\ell,k}(i)=\delta t_{\ell,k}(i-1)+(1-\delta)b_{\ell,k}(i) \tag{21}$$

where $\delta\in(0,1)$ is a forgetting factor, which is used to balance past and present clustering results. Here, $t_{\ell,k}(i)$ is a trust level that assists the clustering process. Based on the trust level $t_{\ell,k}(i)$, the clustering matrix $\mathcal{C}(i)$ can be computed as follows

$$\mathcal{C}_{\ell,\|}(\rangle)=\begin{cases}1, & \text{if}\quad t_{\ell,k}(i)>\chi\\ 0, & \text{otherwise}\end{cases} \tag{22}$$

where $\chi\in[0.5,1)$ is a predefined threshold parameter. Using the clustering matrix $\mathcal{C}$, each node k determines its cluster and identifies its neighbors within the cluster. Additionally, the clustering matrix can also be used to evaluate the accuracy of multi-task learning.

Once the clustering matrix $\mathcal{C}$ is obtained, the combination coefficients can be computed. For simplicity, the averaging criterion [21] is adopted

$$a_{\ell k}(i)=\frac{1}{\sum\limits_{l\in\mathcal{N}_\|}\mathcal{C}_{\ell,\|}(\rangle)} \tag{23}$$

At last, each node k aggregates the local estimate $\boldsymbol{\psi}_k(i)$ and the denoised graph filter coefficient $\hat{\boldsymbol{h}}_{\ell,k}(i-1)$ using the combiner $a_{\ell k}$

$$\boldsymbol{h}_k(i)=a_{kk}(i)\boldsymbol{\psi}_k(i)+\sum_{\ell\in\mathcal{N}_\|^-}a_{\ell k}(i)\hat{\boldsymbol{h}}_{\ell,k}(i-1) \tag{24}$$

A summary of the proposed MT-RDDLMS algorithm is presented in Algorithm 1.

Algorithm 1 MT-RDDLMS Algorithm

Initialization: $\boldsymbol{h}_{k,0}=\boldsymbol{\varphi}_{k,0}=\boldsymbol{0}$, μ_k, μ_{θ_k}, η, δ, χ, v.

for time $i=1,2,3,\ldots$ **do**

for each node $k=1,2,3,\ldots,N$ **do** **1) Decoupled adaptation step:** $\boldsymbol{\psi}_k(i)=\boldsymbol{\psi}_k(i-1)+\mu_k\boldsymbol{z}_k(i)\left[y_k(i)-\boldsymbol{z}_k^{\top}(i)\boldsymbol{\psi}_k(i-1)\right]$.

2) Adaptive smoothing denoising: $\boldsymbol{h}'_{\ell,k}(i-1)=$

$\gamma_{\ell,k}(i-1)\boldsymbol{g}_{\ell,k}(i-1)\boldsymbol{h}_{\ell}(i-1)+\gamma_{\ell,k}(i-1)\boldsymbol{q}_{\ell,k}(i-1)$,
$\hat{\boldsymbol{h}}_{\ell,k}(i-1)=(1-\beta_{k,i})\hat{\boldsymbol{h}}_{\ell,k}(i-2)+\beta_{k,i}\boldsymbol{h}'_{\ell,k}(i-1)$.

3) Clustering and computing the combiners: $a_{\ell k}(i)=\frac{1}{\sum_{l\in\mathcal{N}_k}\mathcal{C}_{\ell,k}(i)}$.

4) Decoupled combination step: $\boldsymbol{h}_k(i)=a_{kk}(i)\boldsymbol{\psi}_k(i)+\sum_{\ell\in\mathcal{N}_k^-}a_{\ell k}(i)\hat{\boldsymbol{h}}_{\ell,k}(i-1)$

4 Simulation

In this section, to evaluate the performance of the proposed algorithm, we conduct extensive Monte Carlo simulations and provide a comprehensive analysis of the results.

We consider the UHVDC graph composed of $N=60$ nodes, where the graph signal $\boldsymbol{x}(i)$ is an independent and identically distributed zero-mean Gaussian vector with covariance $\boldsymbol{R}_x=\text{diag}\{\sigma_{x,k}^2\}_{k=1}^N$. The variance $\sigma_{x,k}^2$ is generated randomly from a uniform distribution $\mathcal{U}(2,2.5)$. The order of the graph filter is set to $L=3$. The noise vector of the link follows a Gaussian vector with zero mean with covariance matrices $Q_{lk}=0.0005I_M$. The fading coefficients $\boldsymbol{g}_{\ell,k}(i)$ are Gaussian variables of zero mean with variances $\sigma_{h,\ell k}^2=1$.

The parameters for the adaptive clustering mechanism are set as follows $\{\eta=0.01,,\ \delta=0.95,\ \chi=0.5,\ \mu_{\theta_k}=0.01,\ v=4\}$. The network topology is generated using the GSPBOX toolbox [22], where each node is connected to its 6 nearest neighbors. The nodes are divided into three clusters: $\mathcal{C}_1=\{1,2,\ldots,20\}$, $\mathcal{C}_2=\{21,22,\ldots,40\}$, $\mathcal{C}_3=\{41,42,\ldots,60\}$. Each cluster is assigned a distinct optimal graph filter coefficient vector

$$\boldsymbol{h}_k^o=\begin{cases}[0.5,1.4,1.9]^{\mathrm{T}}, & \text{for } k\in\mathcal{C}_1\\ [1.3,0.1,0.4]^{\mathrm{T}}, & \text{for } k\in\mathcal{C}_2\\ [1.9,1.3,1.7]^{\mathrm{T}}, & \text{for } k\in\mathcal{C}_3\end{cases} \quad (25)$$

In addition to the multi-task scenario, we also evaluate performance under a single-task scenario, where all nodes share a common target vector $\boldsymbol{h}^o=[0.1,1.6,1.2]^{\mathrm{T}}$.

In simulation, the following algorithms are compared: the multi-task diffusion LMS (MT-DLMS) in ideal environment, the algorithm MT-DLMS with wireless channel, the Non-cooperative mode, the diffusion average estimate bias-compensated LMS (D-ABC-LMS) [15], the MT-DLMS-correntropy [17], and the proposed MT-RDDLMS algorithm are compared. Specifically, MT-DLMS in the ideal environment does not consider the influence of the wireless channel, and it assumes that the cluster information is known. Therefore, it can serve as a reference for optimal performance. The coefficient of the moving average filter in D-ABC-LMS is set to 0.05.

The network mean square deviation (MSD) is used as a performance metric, which is defined as $\text{MSD}(i)\triangleq 10\log_{10}\frac{1}{N}\sum_{k=1}^{N}\|\boldsymbol{h}^o-\boldsymbol{h}_k(i)\|_2^2$. The clustering heatmap is used as a metric to evaluate the performance of multi-task learning. All simulated results were averaged over 200 Monte Carlo runs.

4.1 Multi-Task Scenario

In this experiment, all algorithms employ the same step size of $\mu=0.01$. The transient MSD curves are shown in Fig. 2. As observed, MT-DLMS with wireless channel degradation performs worse than even the non-cooperative strategy. The MT-DLMS-correntropy algorithm demonstrates superior performance over the non-cooperative strategy. The D-ABC-LMS algorithm, which utilizes an average filter, enhances performance, but at the cost of slower convergence. In contrast, the proposed MT-RDDLMS algorithm achieves both fast convergence and low steady-state error. Its performance closely approaches that of the ideal MT-DLMS baseline.

Fig. 3(a) depicts the topology of the graph represented by the adjacency matrix. Fig. 3(b) illustrates the clusters inferred by MT-RDDLMS, which match the ground truth clusters $\mathcal{C}_1-\mathcal{C}_3$. Fig. 4 presents a heatmap comparison of the clustering matrices obtained by MT-RDDLMS and D-ABC-LMS for nodes 51–60 after 20 000 itera-

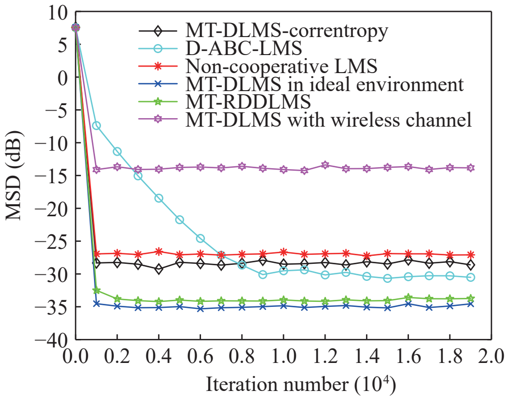

Fig. 2 Transient MSD performance of different algorithms in multi-task scenario

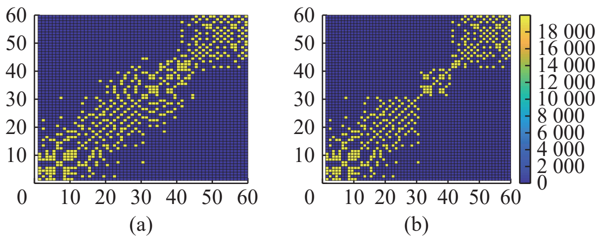

Fig. 3 The performance of cluster: (a) adjacency matrix; (b) inferred clusters

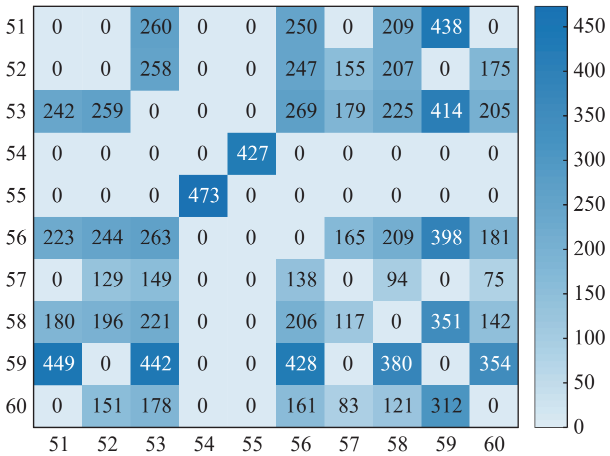

Fig. 4 Heatmap of the cluster difference between MT-RDDLMS and D-ABC-LMS

tions. The proposed method results in significantly more accurate clustering, achieving more than 200 correctly inferred associations on average.

4.2 Single-task scenario

In this experiment, all nodes learn the same graph filter parameters, and all algorithms use the same step size $\mu = 0.01$. As shown in Fig. 5, D-ABC-LMS and MT-DLMS-correntropy both

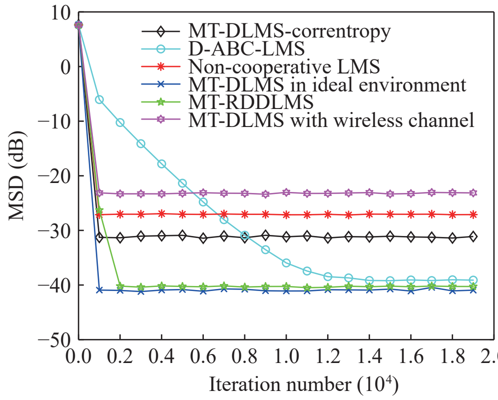

Fig. 5 Transient MSD performance of different algorithms in single-task scenario

outperform the non-cooperative approach due to their robustness to wireless distortions. However, MT-RDDLMS not only maintains superior estimation accuracy but also converges faster than both D-ABC-LMS and MT-DLMS-correntropy. Its performance close to the ideal MT-DLMS, confirming its effectiveness in the single-task scenario.

5 Conclusion

In this work, we mode the UHVDC system as an undirected graph, and the measurement boards are represented by nodes in the graph. A robust multi-task learning algorithm called MT-RDDLMS is proposed to estimate multiple graph filter parameters in the presence of wireless channel. The decoupled ATC strategy is adopted to ensure that each node will not be affected by the wireless channel during the adaptive stage, enhancing robustness to link disturbances. Furthermore, an adaptive smoothing scheme based on a time-varying factor was designed to suppress the influence of link noise. The simulation results verify the effectiveness and robustness of the algorithms.

References:

[1] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808-828, 2018.

[2] I. Jabłoński, "Graph signal processing in applica-

tions to sensor networks, smart grids, and smart cities," *IEEE Sens. J*, vol. 17, no. 23, pp. 7659-7666, 2017.

[3] X. Dong, D. Thanou, L. Toni, M. Bronstein, and P. Frossard,"Graph signal processing for machine learning: A review and new perspectives," *IEEE Signal Process Mag*, vol. 37, no. 6, pp. 117-127, 2020.

[4] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs: Graph filters," in *2013 international conference on acoustics, speech, and signal processing*, pp. 6163–6166, 2013.

[5] S. Segarra, A. G. Marques, and A. Ribeiro,"Optimal graph-filter design and applications to distributed linear network operators," *IEEE Trans. Signal Process*, vol. 65, no. 15, pp. 4117-4131, 2017.

[6] M. J. M. Spelta and W. A. Martins,"Normalized lms algorithm and data-selective strategies for adaptive graph signal estimation," *Signal Process*, vol. 167, pp. 107326, 2020.

[7] R. Nassif, C. Richard, J. Chen, and A. H. Sayed, "A graph diffusion lms strategy for adaptive graph signal processing," in *51st Asilomar Conf. Signals Syst. Comput., Pacific Grove, CA, USA,*, pp. 1973–1976, 2017.

[8] F. Hua, R. Nassif, C. Richard, H. Wang, and A. H. Sayed,"Online distributed learning over graphs with multitask graphfilter models," *IEEE Trans. Signal Inf. Process. Networks*, vol. 6, pp. 63-77, 2020.

[9] Y. Lai, F. Chen, M. Feng, and J. Kurths,"Online distributed admm algorithm with rls-based multitask graph filter models," *IEEE Trans. Network Sci. Eng*, vol. 9, no. 6, pp. 4115-4128, 2022.

[10] H. Zhao, W. Xiang, and S. Lv,"A variable parameter lms algorithm based on generalized maximum correntropy criterion for graph signal processing," *IEEE Trans. Signal Inf. Process. Networks*, vol. 9, pp. 140-151, 2023.

[11] R. Nassif, S. Vlaski, C. Richard, J. Chen, and A. H. Sayed,"Multitask learning over graphs: An approach for distributed, streaming machine learning," *IEEE Signal Process Mag*, vol. 37, no. 3, pp. 14-25, 2020.

[12] F. Chen, J. Li, Y. Zhang, X. Shao, and S. Duan,"Distributed clustering for secure multitask estimation based on adaptive task switching," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 6192-6204, 2023.

[13] R. Abdolee, B. Champagne, and A. H. Sayed,"Estimation of space-time varying parameters using a diffusion lms algorithm," *IEEE Trans. Signal Process*, vol. 62, no. 2, pp. 403-418, 2013.

[14] R. Abdolee, B. Champagne, and A. H. sayed,"Diffusion adaptation over multiagent networks with wireless link impairments," *IEEE Trans. Mob. Comput*, vol. 15, no. 6, pp. 1362-1376, 2015.

[15] S. Zhang and H. C. So,"Diffusion average-estimate bias-compensated lms algorithms over adaptive networks using noisy measurements," *IEEE Trans. Signal Process*, vol. 68, pp. 4643-4655, 2020.

[16] S. Zhang, F. Zhang, H. Chen, R. Merched, and A. H. Sayed,"Frequency-domain diffusion adaptation over networks," *IEEE Trans. Signal Process*, vol. 69, pp. 5419-5430, 2021.

[17] M. Hajiabadi, G. A. Hodtani, and H. Khoshbin, "Robust learning over multitask adaptive networks with wireless communication links," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 66, no. 6, pp. 1083-1087, 2018.

[18] L. Hu, D. Deng, F. Chen, P. Wu, and Z. Chen, "Timeliness analysis for real-time updates across different service time," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 6, pp. 8054-8068, 2024.

[19] J. Fernandez-Bes, J. Arenas-Garcia, M. T. Silva, and L. A. Azpicueta-Ruiz,"Adaptive diffusion schemes for heterogeneous networks," *IEEE Trans. Signal Process*, vol. 65, no. 21, pp. 5661-5674, 2017.

[20] S. Zhang and J. Zhang,"Enhancing the tracking capability of recursive least p-norm algorithm via adaptive gain factor," *Digital Signal Process*, vol. 30, pp. 67-73, 2014.

[21] F. S. Cattivelli and A. H. Sayed,"Diffusion lms strategies for distributed estimation," *IEEE Trans. Signal Process*, vol. 58, no. 3, pp. 1035-1048, 2009.

[22] N. Perraudin, J. Paratte, D. Shuman, L. Martin, V. Kalofolias, P. Vandergheynst, and D. K. Hammond, "Gspbox: A toolbox for signal processing on graphs," *arXiv preprint,* arXiv: 1408.5781, 2014.

Shaojie Wang is with State Grid Shanghai DC Company, focusing on high-voltage direct current control and protection technologies.

Chuliang Xue , a senior engineer, is with State Grid Shanghai Direct Current Company, primarily focusing on research and testing of high-voltage technology.

Hui Li is a senior engineer at State Grid Shanghai Direct Current Company, primarily engaged in research and testing of high-voltage technology.

Rui Zhu , a senior engineer, is with State Grid Shanghai Direct Current Company, primarily focusing on research and testing of high-voltage technology.

Xiaoping Ren is with the College of Artificial Intelligence, Southwest University. His current research interests include distributed signal processing and deep learning.

Limei Hu is currently with the College of Artificial Intelligence, Southwest University. Her current research interests include signal processing and deep learning.