A new variable spatial regularized FxLMS algorithm for diffusion active noise control

A B S T R A C T Distributed multi-channel active noise control (ANC) systems attract a lot of attention due to the reduced computational complexity than centralized control methods and improved stability than decentralized control methods. However, the combination of controllers within a neighborhood in a diffusion manner introduces an estimation bias and may degrade the control accuracy. This is because the secondary sources and error microphones of an ANC system are usually physically placed at different locations and the optimal solution to each controller is different. In this paper, a new diffusion filtered-x least mean squares algorithm (Diff-FxLMS) has been developed that balances combination strength and estimation bias via a variable spatial regularization. The mean squares error criterion subject to a bias constraint is used such that the spatial regularization parameter could be adapted according to the penalized Lagrangian. A detailed performance analysis is carried out, based on which user parameters can be selected automatically. Performance of the proposed variable spatial regularized Diff-FxLMS (VSR-Diff-FxLMS) algorithm and theoretical analysis is verified by simulations.

1. Introduction

Active noise control (ANC) systems have been widely used for noise reduction at lower frequencies [1–5]. For multichannel ANC systems, distributed operation of networked controllers could benefit in saving computational burden of the centralized control and enhancing stability of the decentralized control as shown in recent study [6]. In the context of strongly coupled acoustic systems, where cross channels cause large inferences, distributed ANC algorithms with incremental structures [78] have been employed. The incremental implementation is equivalent to the centralized control and hence has a high computational complexity. A recent study tries to implement the centralized control within a neighborhood, which has a performance comparable to the centralized ANC algorithm at a reduced complexity [9]. To further reduce the complexity of distributed control algorithms, diffusion (Diff) ANC algorithms [10] can be an alternative. A diffusion structure usually has superior properties in flexibility and stability than its incremental counterparts [11]. One problem of the Diff algorithms is that they are developed on the assumption that the optimal coefficients of each controller are the same. For ANC systems, however, the secondary sources (loudspeakers) and error microphones are usually physically distributed at different locations such that the optimal responses deviate from each other. This results in the so called multitask problem in ANC systems, which may degrade the noise reduction performance of diffusion control methods [12]. To illustrate the effect of the multitask problem on Diff ANC algorithms, we first present in Fig. 1 (cited from [12]) the kth channel of a multichannel ANC system using the filtered-x (Fx) based algorithm [13]. The undesirable sound $\{d0k(n)\}$ from the source signal $\{x(n)\}$ through the kth primary path $\{pk(n)\}$ is to be cancelled by the anti-sound $\{yk(n)\}$, generated from the controller $\{wk(n)\}$. A microphone is used to pick up the residual signal $\{ek(n)\}$ at time instant n. $\{ek(n)\}$ includes $\{\eta k(n)\}$ the background noise, and $\{\gamma k(n)\}$ the interference from the other channels. The paths from the controllers to the error microphones are called the channels. To cancel out the noise $\{d0k(n)\}$ at the kth error microphone without considering the cross-talk signals, the controller $\{wk(n)\}$ aims to approximate $\{-pk(n)\}$ after cascading with $\{skk(n)\}$. Therefore, the solution1 to each controller for cancelling out the noise from the primary source deviates from each other if the nodes are located at different places. The centralized and decentralized control methods have been used to find these distinct solutions. The multiple error FxLMS (MEFxLMS) algorithm [13] has been frequently used, where the centralized control optimizes a global cost function over all the residual signal $\{ek(n)\}$. This kind of algorithm has been proved to be steady and efficient, but is quite computationally consuming due to the large amount of (K^2) secondary-path modeling [14]. For weakly coupled systems, where the cross-talk signals are weak and could be ignored, the decentralized control optimizes the local error signal such that it reduces the complexity of multichannel ANC systems at a cost of higher risk of instability [15]. Recently, a lot of effort has been taken to derive a sufficient stability condition for such systems [1617] or improve the convergence performance by controlling the eigenvalues that paly an important role on the adaptation process [18]. As an alternative, the diffusion control is developed, which updates the local controllers as in the decentralized control and then combines these local estimates using a set of combination weights. The diffusion control increases the stability of the decentralized control system through the communication between nodes while maintains a comparable complexity. However, the combination strategy introduces an estimation bias since the local estimates have different true values as aforementioned. In a very few studies addressing multitask problems of diffusion ANC systems, the effect of diffusion strategy is theoretically analyzed [1219]. It shows that the diffusion control enforces stability of the ANC system at the cost of the increased estimation bias. An approach that uses the variable spatial regularization (SR) is proposed to alleviate the estimation bias problem [20]. Compared to the conventional Diff-FxLMS algorithm that uses a set of fixed combination weights, a variable SR parameter is more flexible to adjust diffusion strength. In this paper, a new FxLMS algorithm with variable SR is proposed. Different from our previous work [20], where the variable SR is derived from the steady-state estimation error, the new algorithm uses a rule to adjust the diffusion strength automatically according to the instantaneous estimation error in time-varying acoustic environments. The theoretical analysis includes the transient update process of VSR, which also differs from our previous work that focuses on the steady state only. Various attempts to explore the general multitask problem in diffusion networks have been carried out in current literature. One method decomposes the unknown space-varying parameters into basic functions [21] so that conventional Metropolis or uniform rules [22] can be used to generate the fixed combination weights. More generally, penalized optimization approaches can benefit tracking of multitask and interrelated optima [2324]. For instance, the nodes are clustered and updated according to the distance between estimates [25]; the estimation bias and variance analysis framework of the spatial regularization is developed [2324] that provides guidance for diffusion adaptive algorithm design in multitask environments. In this paper, a regularized local cost function subject to a smoothness constraint and a penalty term has been minimized to derive a new variable spatial regularized Diff-FxLMS (VSR-Diff-FxLMS) algorithm. The resulting VSR rule arises naturally from the constraint and adjusts the combination strength automatically. The proposed algorithm varies the SR parameter to tradeoff the conflicting requirements of strong combination to reduce noise variance and weak combination to minimize estimation bias. The mean and mean squares performance of the proposed algorithm is analyzed theoretically. Different from the analysis in our previous work [19], which characterizes the estimation deviation and specifies the convergence condition, the performance analysis in this paper focuses on the bias and variance analysis and gains insight into the mechanism of the VSR strategy. Formulas for selecting the optimal SR parameter are obtained from the theoretical analysis. The proposed algorithm is examined by computer simulations. It shows that the proposed algorithm efficiently improves the noise reduction performance in the multitask environment.

2. Review of Diff-FxLMS algorithms

Recalling the diagram of the multichannel ANC system in Fig. 1 and defining the convolution sign "*", the unwanted noise signal, output of the controller, and interference from other secondary sources can be expressed as $d_{0k}(n) = p_k(n)*x(n)$, $y_k(n) = x(n)*w_k(n)*s_{kk}(n)$, and $\gamma_k(n) = \sum_{l\neq k,l=1}^{K} x(n)*w_l(n)*s_{lk}(n)$. Assuming a stationary environment, i. e. the acoustic paths and controllers are evolving slowly, the commutative property can be used to derive the Fx-type algorithm and $x_{s_{lk}}(n) = x(n)*s_{lk}(n)$ can be used as the input of the controller. Consequently, the error signal reads $e_k(n) = d_{0k}(n)+\boldsymbol{x}_k^T(n)\boldsymbol{w}_k+\eta_k(n)+\gamma_k(n)$ (1) where $\boldsymbol{w}_k = [w_{k,1}...w_{k,L}]^T$ is the kth controller of length *L*, and **x***k*(*n*) = [*xskk*(*n*)...*xskk*(*n* □ *L* + 1)]*T* is the filtered reference signal. To achieve a quiet zone, the controllers should be updated in response to all the available error signals so as to minimize a global cost function [13] *J* = Σ*K k*=1 *Jk* = *E*[Σ*K k*=1 *e*2*k*(*n*)] (2) where *Jk* = *E*[*e*2*k*(*n*)] (3) is the local cost function. The global LMS solution to (2) is not distributed since every node needs to have access to global information in order to compute the new estimate. A general formulation of distributed estimators based on diffusion strategies has been proposed [22] that derives diffusion filters by optimizing the regularized local cost function at each node *Jdis k* = *Jk* +*ξk* Σ *l*∈N*k*/{*k*}*blk*||***wk*** □ ***ψl***‖2 (4) where ***ψl*** is an estimate of the optima ***wl*** that is available at each node parameter. {*blk*} is a set of non-negative real coefficients that give different weights to different neighbours. They have the following properties [22] *blk* = 0 if *l* ∉ N*k*, **1***T**B*** = **1***T* (5) where **1** is the column vector with unit entries and {*blk*} forms the combination matrix ***B***. It can be seen that {*blk*} are zero when the *l*th and *k*th nodes are not connected electrically, i.e. they do not share information between each other. Moreover, each column of the matrix adds up to one. Taking the derivative of *Jdis k* in terms of ***wk*** and following the incremental solutions in [22], the so called Adapt-Then-Combine (ATC) FxLMS algorithm is specialized ***ψk***(*n* + 1) = ***wk***(*n*) □ *μk*∇***wk*** *e*2*k*(*n*) = ***wk***(*n*) □ *μkek*(*n*)ˆ**x***k*(*n*) (6) ***wk***(*n* + 1) = ***ψk***(*n* + 1) □ *ξk* Σ *l*∈N*k*/{*k*}*blk*(***ψk***(*n* + 1) □ ***ψl***(*n* + 1)) (7) where the stochastic approximationˆ*Jk* = *e*2*k*(*n*) has been used for *Jk*, and {*μk*, *ξk*} are positive parameters denoting the step-size and the SR parameter. It should be noted that the filtered reference signal has been replaced by its estimateˆ*xslk*(*n*) = *x*(*n*)**slk*(*n*) andˆ**x***k*(*n*) = [ˆ*xskk*(*n*)...ˆ*xskk*(*n* □ *L* + 1)]*T*, since {*slk*(*n*)} is unknown and estimated as {ˆ*slk*(*n*)}. Defining the nonnegative scalars. *akk* = 1 □ (1 □ *bkk*)*ξk*, *alk* = *ξkblk* for *l* ≠ *k*, Eq. (7) reduces to ***wk***(*n* + 1) =Σ*l*∈N*kalk**ψl***(*n* + 1) (8) It can be seen that the cluster averaged estimate ***wk***(*n*) combines the local solutions ***ψk***(*n*) over *l* ∈ N*k* via {*alk*}. The adaptation and combination steps can be exchanged, which formulates the Combine-Then- Adapt (CTA) CTA-FxLMS algorithm.

3. The proposed VSR-ATC-FxLMS algorithm

Comparing the global and diffusion cost functions in (2) and (4), it can be seen that if the optimal solutions {***wk***,0 for *k* = 1,…,*K*} at each node do not equal to each other, i.e. ***wm***,0 ≠ ***wn***,0 if *m* ≠ *n*, an extra error term arises from the second term of the cost function (4), which reads *σ*2***wk*** = Σ *l*∈N*k*/{*k*}*blk*||***wk***,0 □ ***wl***,0‖2 (9) This estimation bias is due to the multitask problem. The specific expression of {***wk***,0} will be presented in (24) Section 3, where the performance analysis is carried out. In the sequel, it shows how the unwanted system error is used to select combination coefficients in an adaptive way. We consider the Lagrangian for minimizing *Jk* subject to a constraint of *σ*2***wk*** *J*(1) *k* = *Jdis k* +*λk*(Σ *l*∈N*k*/{*k*}*blk*||***wk*** □ ***ψl***‖2 □ *σ*2***wk***) (10) for some Lagrangian multipliers *λk*. Due to the convex property of the local cost function, the solution to (10) should be unique. However, *λk* may not be unique considering that *λk* may degenerate to 0. To avoid this problem, we subtract a penalty term *κkλ*2*k* (*κk*>0) from *J*(1) *k* such that the multiplier could be maximized much easier as a quadratic function. Consequently, we have *J*(2) *k* = *Jdis k* +*JCk* (11) where *JCk* = *κkλk* [(Σ *l*∈N*k*/{*k*}*blk*||***wk*** □ ***ψl***‖2) □ *σ*2***wk***] □ *κkλ*2*k*(12) Using the instant system errors and applying the stochastic root-finding rules [26] to (11), we obtain an adaptive estimate to determine {***wk***, *λk*} ***wk***(*n* + 1) = ***wk***(*n*) □ (*μk*∇***wk*** *Jdis k* + *ξk*∇***wk*** *JCk*) (13) *λk*(*n* + 1) = *λk*(*n*)+*βk*∇*λk JCk* (14) where ∇***u*** denotes the derivative operator in terms of ***u*** and *βk* is a positive parameter. Substituting *Jdis k* into (11) and recalling the derivation of the ATC-FxLMS algorithm from (6)-(8), the incremental solution to (13) can be obtained in a similar way. Absorbing the penalty parameter *κk* into *ξk*, an extra term (1+*κkλk*(*n*)) is multiplied to the SR parameter *ξk* ***ψk***(*n* + 1) = ***wk***(*n*) □ *μk*∇***wk*** *e*2*k*(*n*) (15) ***wk***(*n* + 1) = ***ψk***(*n* + 1) □ *ξk*(1 + *κkλk*(*n*)) Σ *l*∈N*k*/{*k*}*blk*(***ψk***(*n* + 1) □ ***ψl***(*n* + 1)). (16) Compared to the conventional Diff-FxLMS algorithm in (7), the extra term (1+*κkλk*(*n*)) works as a measurement of the system error and makes the SR adjustable according to the convergence status. After redefining a set of variable combination coefficients. *akk*(*n*) = 1 □ (1 □ *bkk*)*αk*(*n*), *alk* = *αk*(*n*)*blk* for *l* ≠ *k* (17) the variable parameter can be written as. *αk*(*n*) = *ξk*(1 + *κkλk*(*n*)), *s.t.* 0 ≤ *αk*(*n*) ≤ 1

(18) It shows that when the algorithm is far from convergence, a larger combination coefficient is used while when the algorithm is close to convergence, the combination strength decreases. Then, the update equations for controller coefficients (15)-(16) can be summarized as $\boldsymbol{\psi}_k(n+1) = \boldsymbol{w}_k(n)$ □ $\mu_k e_k(n)\hat{\mathbf{x}}_k(n)$ (19) $\boldsymbol{w}_k(n+1) = \Sigma_{l\in N_k} a_{lk}(n)\boldsymbol{\psi}_l(n+1)$ (20) It can be seen that the proposed diffusion algorithm has a set of variable combination parameters as defined in (17) and (18). The variable parameter $\alpha_k(n)$ arises naturally from the constraint, which are called the VSR parameter. It can be seen from (20) that the VSR automatically updates with $\lambda_k(n)$. According to (14), the update equation for $\lambda_k(n)$ can be obtained by taking the derivative of J_{Ck} in terms of $\lambda_k(n)$, and absorbing the parameter $2\kappa_k$ into β_k, i.e. $\lambda_k(n+1) = \lambda_k(n) + \beta_k\,[\frac{1}{2}(\delta_k(n)$ □ $\hat{\sigma}^2_{wk})$ □ $\lambda_k(n)]$ (21) where $\delta_k(n) = \Sigma_{l\in N_k/\{k\}} b_{lk}\|\boldsymbol{\psi}_k(n+1)$ □ $\boldsymbol{\psi}_l(n+1)\|^2$ is the instantaneous estimation bias, and $\hat{\sigma}^2_{wk}$ is an estimate of the theoretical value σ^2_{wk} as defined in (9). Thus, $\lambda_k(n)$ measures the distance between the instantaneous estimation bias and the theoretical value σ^2_{wk} in (9). The value of $\hat{\sigma}^2_{wk}$ can be specified via the performance analysis as shown in the next section. In case the instantaneous estimation bias is smaller than the theoretical value, $\lambda_k(n)$ could be negative. This means the combination might be too strong such that the controllers converge to a similar value. A negative $\lambda_k(n)$ then reduces the combination coefficients according to (18), which should be lower bounded by 0. We call (17)-(21) the VSR-*Y ATC-FxLMS algorithm.*

4. *Performance analysis*

To make the performance analysis mathematically available, the following assumptions are made: (A1) the filtered input signal $\{\hat{x}_{sij}(n)\}$ is the zero-mean Gaussian random signals; (A2) $\mathbf{x}_k(n)$ (or $\hat{\mathbf{x}}_k(n)$) is statistically independent with $\mathbf{w}_k(n)$ or $\eta_k(n)$. Moreover, $\mathbf{w}_k(n)$ and $\eta_k(n)$ are statistically independent with each other; (A3) the multichannel system is weakly coupled such that the interference $\{\gamma_k(n)\}$ is small. To have a concise expression in the following analysis, we define the global representations in terms of the stochastic quantities that appear in the update equations. $\mathbf{e}(n) = col\{e_1(n)...e_K(n)\}(K\times 1)$, $\mathbf{d}(n) = col\{d_1(n)...d_K(n)\}$ $(K\times 1)$, $\boldsymbol{\psi}_C(n) = col\{\boldsymbol{\psi}_1(n)...\boldsymbol{\psi}_K(n)\}(KL\times 1)$, $\mathbf{w}_C(n) = col\{\mathbf{w}_1(n)...\mathbf{w}_K(n)\}$ $(KL\times 1)$, $\mathbf{X}(n) = diag\{\mathbf{x}_1(n)...\mathbf{x}_K(n)\}(KL\times K)$, $\hat{\mathbf{X}}(n) = diag\{\hat{\mathbf{x}}_1(n)...\hat{\mathbf{x}}_K(n)\}$ $(KL\times K)$. Here, col{} and diag{} format the elements as a column vector or block diagonal matrix. We also introduce the diagonal matrices of the global expression for step-sizes of dimension $KL\times KL$ $\mathbf{D}_\mu(n) = diag\{\mu_1\mathbf{I}_L...\mu_K\mathbf{I}_L\}$ According to (17), the global combination matrix becomes $\mathbf{G} = [\mathbf{I}_K$ □ $(\mathbf{I}_K$ □ $\mathbf{B})\mathbf{D}_\alpha(n)]\otimes\mathbf{I}_L$, where $\otimes$ is the Kronecker product, $\mathbf{I}_L$ the identity matrix of size L, and $\mathbf{B}$ the symmetric combining matrix formed by $\{b_{lk}(n)\}$. In the following, the behaviour of the VSR-ATC-FxLMS algorithm is studied such that the user parameters of the proposed algorithm can be selected. Different from our previous work [12], where the mean and mean squares performance of the conventional ATC-FxLMS algorithm is theoretically analysed, this paper focuses on effect of the variable SR rule and uses the bias-variance analysis to optimize the important user parameters that determine the SR strength of the proposed VSR-ATC-FxLMS. This work is not available in [12]. 4.1. Mean convergence analysis and constraint parameter selection Using the notations, we have the following difference equation from (19)(20) $\mathbf{w}_C(n+1) = \mathbf{G}^T[\mathbf{w}_C(n)$ □ $\mathbf{D}_\mu\hat{\mathbf{X}}(n)\mathbf{e}(n)]$ (22) The global solution can be obtained from (22) as the algorithm reaches the steady state $\mathbf{w}_\infty =$ □ $[\hat{\mathbf{R}}_{XX} + (\mathbf{G}\mathbf{D}_\mu)$□ $^1(\mathbf{I}_{LK}$ □ $\mathbf{G})]$□ $^1\hat{\mathbf{r}}_{Xd}$ (23) where $\hat{\mathbf{R}}_{XX} = E[\hat{\mathbf{X}}(n)\mathbf{X}^T(n)]$ and $\hat{\mathbf{r}}_{Xd} = E[\hat{\mathbf{X}}(n)\mathbf{d}(n)]$. It can be seen that if there is no communication between nodes, i.e. $\mathbf{G} = \mathbf{I}_{LK}$, the global solution reduces to $\mathbf{w}_0 =$ □ $\hat{\mathbf{R}}$□ $^1_{XX}\hat{\mathbf{r}}_{Xd}$. This solution stacks the bias-free estimates $\mathbf{w}_{k,0}$ at all nodes. Using the relationship $\boldsymbol{\psi}_0 = \mathbf{G}$□ $^1\mathbf{w}_0$, the user parameter can be selected according to (9). $\hat{\sigma}^2_{wk} = \Sigma_{l\in N_k/\{k\}} b_{lk}\|\boldsymbol{\psi}_{k,0}$ □ $\boldsymbol{\psi}_{l,0}\|^2$ (24) This parameter can be calculated offline from a section of training data. Next, the estimation bias $\Delta\mathbf{w} = \mathbf{w}_0$ □ $\mathbf{w}_\infty$ is to be specified. From (23) and the global solution $\mathbf{w}_0$, the bias term can be further written as $\Delta\mathbf{w} = (\mathbf{G}\mathbf{D}_\mu\hat{\mathbf{R}}_{XX})$□ $^1(\mathbf{I}_{LK}$ □ $\mathbf{G})\mathbf{w}_\infty$ (25) Suppose the global solution to each controller has a common part and write the stacked K common parts as $\mathbf{p}_0$ $(KL \times 1)$. Then, the deviation of $\mathbf{w}_\infty$ to $\mathbf{p}_0$ can be expressed as $\Delta\mathbf{p} = \mathbf{w}_\infty$ □ $\mathbf{p}_0$. Since the common part satisfies $\mathbf{p}_0 = \mathbf{G}\mathbf{p}_0$, the global system deviation has the property ($\mathbf{I}_{KL}$ − $\mathbf{G}$) $\mathbf{w}_\infty$ = ($\mathbf{I}_{KL}$ − $\mathbf{G}$)$\Delta\mathbf{p}$. Thus, the estimation bias can be expressed as $\Delta\mathbf{w}$ = ($\mathbf{G}\mathbf{D}_\mu\hat{\mathbf{R}}_{XX}$)$^{-1}$($\mathbf{I}_{KL}$ − $\mathbf{G}$)$\Delta\mathbf{p}$ (26) Defining the global weight error vector $\mathbf{v}(n)$ = $\mathbf{w}_\infty$ − $\mathbf{w}_C(n)$ and using (22), (18), and (21), the mean difference equations for the controller, $\alpha_k(n)$, and $\lambda_k(n)$ can be written as $E[\mathbf{v}(n + 1)] = \mathbf{G}(\mathbf{I}_{LK} - \mathbf{D}_\mu\hat{\mathbf{R}}_{XX})E[\mathbf{v}(n)]$ (27) $E[\alpha_k(n)] = \xi_k(1 + \kappa_k E[\lambda_k(n)])$ (28) $E[\lambda_k(n + 1)] = (1 - \beta_k)E[\lambda_k(n)]+\frac{1}{2}\beta_k(E[\delta_k(n)] - \hat{\sigma}^2_{wk})$ (29) Since the radius of $\mathbf{G}(\mathbf{I}_{LK} - \mathbf{D}_\mu\hat{\mathbf{R}}_{XX})$ is generally smaller than that of $(\mathbf{I}_{LK} - \mathbf{D}_\mu\hat{\mathbf{R}}_{XX})$, the network topology enforces robustness of the mean convergence of (27) compared to the decentralized control. This conclusion is similar to that in [12]. 4.2. Mean squares performance analysis and EMSE To study the mean squares performance, we define the error covariance matrix $\boldsymbol{\Xi}(n) = E[\mathbf{v}(n)\mathbf{v}^T(n)]$. From (22) and (27), the difference equation can be obtained $\boldsymbol{\Xi}(n + 1) = \mathbf{G}\boldsymbol{\Xi}(n)\mathbf{G} - \mathbf{G}[\mathbf{D}_\mu\hat{\mathbf{R}}_{XX}\boldsymbol{\Xi}(n) + \boldsymbol{\Xi}(n)\mathbf{R}_{X\hat{X}}\mathbf{D}_\mu]\mathbf{G} + \mathbf{G}\mathbf{D}_\mu\hat{\mathbf{R}}_{XX}\mathbf{v}(n)\mathbf{w}^T_\infty(\mathbf{G} - \mathbf{I}_{LK})^T + (\mathbf{G} - \mathbf{I}_{LK})\mathbf{w}_\infty\mathbf{v}^T(n)\mathbf{R}_{X\hat{X}}\mathbf{D}_\mu\mathbf{G} + \mathbf{G}\mathbf{D}_\mu E[\mathbf{Q}(n) - \mathbf{R}_y]\mathbf{D}_\mu\mathbf{G}$ (30) where $\mathbf{R}_y = \hat{\mathbf{R}}_{XX}\Delta\mathbf{w}\Delta\mathbf{w}^T\mathbf{R}_{X\hat{X}}$ and $\mathbf{Q}(n) = E[\mathbf{X}(n)\mathbf{e}(n)\mathbf{e}^T(n)\mathbf{X}^T(n)]$. It is worth noting that the relationship $\mathbf{w}_\infty = \mathbf{G}\mathbf{w}_\infty$ no longer holds, which results in the 3rd and 4th terms on the right hand side of (30) that contain the combination matrix ($\mathbf{G} - \mathbf{I}_{LK}$). According to the assumptions and after proper algebraic operations to deal with $\mathbf{Q}(n)$ (see Appendix A in [12]), Eq. (30) can be rewritten as $\boldsymbol{\Xi}(n + 1) = \mathbf{G}[\boldsymbol{\Xi}(n) - \mathbf{D}_\mu\hat{\mathbf{R}}_{XX}\boldsymbol{\Xi}(n) - \boldsymbol{\Xi}(n)\mathbf{R}_{X\hat{X}}\mathbf{D}_\mu]\mathbf{G} + \mathbf{G}\mathbf{D}_\mu\hat{\mathbf{R}}_{XX}[2\boldsymbol{\Xi}(n) + \Delta\mathbf{w}\Delta\mathbf{w}^T + \boldsymbol{\Xi}_\Sigma(n)]\mathbf{R}_{X\hat{X}}\mathbf{D}_\mu\mathbf{G} + \{Tr[(\boldsymbol{\Xi}(n) + \boldsymbol{\Xi}_\Sigma(n))\mathbf{R}_{XX}] + \sigma^2_y\}\mathbf{G}\mathbf{D}_\mu\mathbf{R}_{\hat{X}\hat{X}}\mathbf{D}_\mu\mathbf{G} + \mathbf{G}\mathbf{D}_\mu\mathbf{D}_\Sigma\mathbf{R}_{XX}\mathbf{D}_\mu\mathbf{G}$ (31) where $\boldsymbol{\Xi}_\Sigma(n) = E[\mathbf{v}(n)]\Delta\mathbf{w}^T + \Delta\mathbf{w}E[\mathbf{v}^T(n)]$, $\sigma^2_y = Tr(\mathbf{R}^{-1}_{XX}\mathbf{R}_y)$, $\mathbf{R}_{\hat{X}\hat{X}} = E[\hat{\mathbf{X}}(n)\hat{\mathbf{X}}^T(n)]$, and $\mathbf{R}_{XX} = E[\mathbf{X}(n)\mathbf{X}^T(n)]$. Moreover, a diagonal matrix is defined $\mathbf{D}_\Sigma = diag\{\sigma^2_{\Sigma,1}\mathbf{I}_L...\sigma^2_{\Sigma,K}\mathbf{I}_L\}$, where $\sigma^2_{\Sigma,k}$ represents the variance of the summarized noise $\eta_{\Sigma k}(n) = \Delta_k(n)+\eta_k(n)+\gamma_k(n)$ with the modeling error $\Delta_k(n) = d_{0k}(n)+\mathbf{x}^T_k(n)\mathbf{w}_{k,0}$ [12]. We now examine the square error after the cooperative learning process reaches a steady state. The averaged global steady-state EMSE is defined as [22] $J = \frac{1}{K}Tr(\mathbf{R}_{xx}\boldsymbol{\Xi}(\infty))$ (32) where $\mathbf{R}_{xx} = diag\{\mathbf{R}_{xx,1}...\mathbf{R}_{xx,K}\}$ and $\mathbf{R}_{xx,k} = E[\mathbf{x}_k(n)\mathbf{x}^T_k(n)]$ are, respectively, the global and local input covariance matrices. At the steady state, we take trace of (31) multiplied with an arbitrary matrix Σ, and rewrite it using vector notations, e.g. $Tr(\Sigma\boldsymbol{\Xi}(\infty)) = vec\{\boldsymbol{\Xi}(\infty)\}^T vec\{\Sigma\}$. By defining $vec(\Sigma) = \mathbf{F}^{-1}vec(\mathbf{R}_{xx})$ and using the Kronecker product properties $vec(\mathbf{X}\Sigma\mathbf{Y}) = (\mathbf{Y} \otimes \mathbf{X}^T)vec(\Sigma)$ and $Tr(\mathbf{X}\mathbf{Y}) = vec(\mathbf{X}^T)^T vec(\mathbf{Y})$ for some arbitrary matrices $\mathbf{X}$ and $\mathbf{Y}$, the steady-state EMSE can be obtained as $J = \frac{1}{K}v^T_M\mathbf{F}^{-1}vec(\mathbf{R}_{xx})/[1 - v^T_R\mathbf{F}^{-1}vec(\mathbf{R}_{xx})]$ (33) where $\mathbf{F} = \mathbf{I}_{L^2K^2} - (\mathbf{I}_{L^2K^2} - \mathbf{R}_F)(\mathbf{G} \otimes \mathbf{G})$ with $\mathbf{R}_F = \mathbf{I}_{LK} \otimes \mathbf{D}_\mu\hat{\mathbf{R}}_{XX} + \mathbf{R}_{X\hat{X}}\mathbf{D}_\mu \otimes \mathbf{I}_{LK} + 2\mathbf{R}_{X\hat{X}}\mathbf{D}_\mu \otimes \hat{\mathbf{R}}_{XX}\mathbf{D}_\mu$, $v_M = (\mathbf{D}_\mu \otimes \mathbf{D}_\mu)(\mathbf{G}\otimes \mathbf{G})vec(\mathbf{R}_y + \sigma^2_y\mathbf{R}_{\hat{X}\hat{X}} + \mathbf{D}_\Sigma\mathbf{R}_{XX})$, and $v_R = (\mathbf{D}_\mu \otimes \mathbf{D}_\mu)(\mathbf{G}\otimes \mathbf{G})vec(\mathbf{R}_{\hat{X}\hat{X}})$. 4.3. Bias-variance analysis and SR parameter selection The effect of SR on the ATC-FxLMS algorithm includes both the estimation bias in (26) and variance in (33). These expressions are still implicit. To make it clear, we consider a uniform setting of $\xi_k = \xi_0$ for $k = 1…K$, and assume that $\delta_k(n)$ converges to $\delta_k(\infty) = \hat{\sigma}^2_{wk}$ in the steady state. Then the limit $\lambda_{k,\infty} = \frac{1}{2}(\delta_k(\infty) - \hat{\sigma}^2_{wk}) \cong 0$ can be obtained such that $\alpha_k(n) = \xi_k(1 + \kappa_k\lambda_k(n)) \cong \xi_0$ and $\mathbf{G}_\infty = \mathbf{I}_{KL} - \xi_0\mathbf{B}_\Delta$ with $\mathbf{B}_\Delta = (\mathbf{I}_K - \mathbf{B})\otimes \mathbf{I}_L$. In the following, we discuss a special case of a small and uniform setting of the step-size, i.e. $\mu_k = \mu_0$, for $k = 1…K$. Using these limits and assumptions, the estimation bias reduces to $\Delta\mathbf{w} = (\mathbf{G}_\infty\mathbf{D}_\mu\hat{\mathbf{R}}_{XX})^{-1}(\mathbf{I}_{KL} - \mathbf{G})\Delta\mathbf{p}$. Applying the first order approximation to $\mathbf{G}_\infty$, $\Delta\mathbf{w}$ reduces to $\Delta\mathbf{w} \approx \frac{1}{\mu_0}\xi_0\hat{\mathbf{R}}^{-1}_{XX}\mathbf{B}_\Delta(\mathbf{I}_{KL} - \mathbf{G})\Delta\mathbf{p}$ (34) On the other hand, the steady-state EMSE finds from Appendix A $J \approx \frac{1}{K}\mu_0 vec(\mathbf{D}_\Sigma\mathbf{R}_{XX})^T(\mathbf{I}_{L^2K^2} - \xi_0\mathbf{B}_\Delta \oplus \mathbf{B}_\Delta)^T \times[\mathbf{I}_{L^2K^2} - \xi_0\mathbf{R}^{-1}_F(\mathbf{B}_\Delta \oplus \mathbf{B}_\Delta)](\mathbf{R}_{X\hat{X}} \oplus \hat{\mathbf{R}}_{XX})^{-1}vec(\mathbf{R}_{xx})$ (35) Using (34)(35) and defining the total MSE $J_R = \frac{1}{K}\Delta\mathbf{w}^T\mathbf{R}_{xx}\Delta\mathbf{w} + J$, we have $J_R = \frac{1}{K}(a\xi^2_0 - b\xi_0 + c)$ (36) where $a = \frac{1}{\mu^2_0}Tr(\mathbf{R}_{xx}\hat{\mathbf{R}}^{-1}_{XX}\mathbf{B}_\Delta\Delta\mathbf{p}\Delta\mathbf{p}^T\mathbf{B}^T_\Delta\hat{\mathbf{R}}^{-T}_{XX}) + \mu^2_0 vec(\mathbf{D}_\Sigma\mathbf{R}_{XX})^T \times (\mathbf{B}_\Delta \oplus \mathbf{B}_\Delta)^T\mathbf{R}^{-1}_F(\mathbf{B}_\Delta \oplus \mathbf{B}_\Delta)\mathbf{R}^{-1}_F vec(\mathbf{R}_{xx})$ $b = \mu_0 vec(\mathbf{D}_\Sigma\mathbf{R}_{XX})^T[(\mathbf{B}_\Delta \oplus \mathbf{B}_\Delta)^T + \mathbf{R}^{-1}_F(\mathbf{B}_\Delta \oplus \mathbf{B}_\Delta)](\mathbf{R}_{X\hat{X}} \oplus \hat{\mathbf{R}}_{XX})^{-1}vec(\mathbf{R}_{xx})$, and $c = \mu_0 vec(\mathbf{D}_\Sigma\mathbf{R}_{XX})^T(\mathbf{R}_{X\hat{X}} \oplus \hat{\mathbf{R}}_{XX})^{-1}vec(\mathbf{R}_{xx})$ It is clear that the estimation bias is proportional to regularization while the estimation variance is inverse proportional to ξ_0. The proposed algorithm should adjust the parameter $\alpha_k(n)$ automatically so that it uses strong combination towards estimate in common to reduce estimation variance and weak combination when nodes converge to distinct local optima to constrain the estimation bias. To minimize J_R, the derivative of (36) is set to 0 to get the optimal SR parameter $\xi_{opt} = \frac{1}{2}b/a$, $\xi_{opt} \leq 1$ (37) This formula helps to narrow the range for SR parameter selection. For *a special case of white source noise, Eq. (37) reduces to $\xi(s)$ opt = $12\mu 0\tau 1(1 + 1\ 2\mu 0\sigma 2xs)\sigma 2x\sigma 20\ 1\ \mu 0\sigma 2x\|\boldsymbol{B}\Delta\Delta\boldsymbol{p}\|22 + 1\ 2\sigma 2xs\tau 2\sigma 2x\sigma 20$, $0 \leq \xi(s)$ opt ≤ 1 (38) where $\sigma 2x$, $\sigma 2xs$, and $\sigma 20$ are, respectively, the source, the filtered input, and summarized noise power, $\tau 1 = vec(\boldsymbol{R}XX)T(\boldsymbol{B}\Delta \oplus \boldsymbol{B}\Delta)\ (\boldsymbol{R}X\hat{}X \oplus \boldsymbol{R}\hat{}XX)\square\ 1vec(\boldsymbol{R}XX)$ and $\tau 2 = vec(\boldsymbol{R}XX)T\ (\boldsymbol{B}\Delta \oplus \boldsymbol{B}\Delta)T\ (\boldsymbol{B}\Delta \oplus \boldsymbol{B}\Delta)(\boldsymbol{R}X\hat{}X \oplus \boldsymbol{R}\hat{}XX)\square\ 1vec(\boldsymbol{R}XX)$ It can be seen that the optimal SR parameter is proportional to the step-size, input power, and system noise power, and inverse proportional to the deviation of the optimal controllers $\|\boldsymbol{B}\Delta\Delta\boldsymbol{p}\|22$.*

5. Simulation results

In this section, the effect of the proposed VR strategy on the multitask ANC system has been investigated and the theoretical analysis is verified. The algorithms under test include the decentralized FxLMS, conventional ATC-FxLMS, and the proposed VSR-ATC-FxLMS with full ($\hat{\sigma}2wk=\sigma 2wk$) or null ($\hat{\sigma}2wk=0$) knowledge of channels. The Metropolis combination is used for all the diffusion algorithms. All the simulation results are averaged over 100 independent trials if not specified. 5.1. Evaluation of the mean squares performance analysis In the first experiment, we consider a system of 10 nodes. As illustrated in Fig. 2, the secondary loudspeakers and error microphones (denoted as blue dots) are located on two circles centered at (0, 0). The radiuses for loudspeakers and microphones are, respectively, rL and rM. The primary noise source is located on the x-z plane with a radius of rS and an angle of φ to z-axis. The simulation is conducted in a free field. The radiuses are set to rM = 1 and rL = 1.1. The sampling rate is set to 2 KHz to focus on lower frequencies. Thus, the primary and secondary paths have a length of 30. The length of the controllers is set to 100. This setting that describes an acoustic free-field has been used in Sections 5.1 and 5.2 to compare noise control performance at different primary noise positions and using different control strategies. Fig. 3(a) shows the simulated and theoretical results for the averaged EMSE curves [see Eq. (31), where the VR strategy is modeled by Eqs. (17), (27)-(29)] under different settings. First, the power ratio between the source signal $\{x(n)\}$ and the background noise $\{\eta k(n)\}$, called signal to noise ratio (SNR), is set to 10 dB. The angle φ is set to 0◦ (two lines at bottom), 2◦ (two lines in the middle), and 5◦ (two lines on top) to form different multitask problems. The user parameters $\{\mu k, \xi k, \beta k, \kappa k\}$ for VSR-ATC-FxLMS are set to {0.0005, 0.01, 0.01, 100}. It can be seen that the theoretical and simulated EMSE curves generally agree well with each other. The steady-state EMSE increases as the angle φ when the primary source is away from the z-axis and the multitask problem becomes more serious [see Eq. (25)]. Next, the above experiment is repeated under different settings of SNRs. Similar performance can be observed in Fig. 3(b). 5.2. Evaluation of the VSR-ATC-FxLMS algorithm Computer simulations of a 10-node network as in Fig. 2 have been carried out to validate the proposed algorithm. The radiuses are still set to rM = 1 and rL = 1.1. The primary path is located on the x-z plan with rS = 2, and φ = 0◦ and 5◦. The primary and secondary paths have a length of 100. The length of the controllers is set to 250. In the simulations, several control methods are compared in Fig. 4. The SNR is set to 10 dB. The step-size for ATC-FxLMS and decentralized FxLMS is set to 0.00015. The user parameters for VSR-ATC-FxLMS are set to $\{\mu k, \beta k, \kappa k\}$ = {0.00015, 0.01, 1000}. The other settings are identical to those in the previous experiment if not specified. Different values for the angle φ are used. First, φ is set to 0◦. In this case, the parameter ξk is set to 1. It can be seen from Fig. 4(a) that the ATC-FxLMS algorithm outperforms the decentralized version, since ATC cooperation combats noise interference via signal combination. The proposed VSR-ATC-FxLMSs have similar performance with ATC-FxLMS. In the second case, φ is set to 5◦. The parameter ξk is set to 0.01. It turns out that the decentralized FxLMS converges to a lower steady-state EMSE than the conventional ATC-FxLMS. The degradation of the ATC-FxLMS is due to the estimation bias introduced by the combination of different local estimates. The proposed VSR-ATC-FxLMS adjusts the SR parameter and have a similar performance with the decentralized algorithm. To further evaluate the performance of ANC systems, the noise reduction values of the proposed VSR-ATC-FxLMS algorithm have been calculated and compared with the conventional ATC-FxLMS algorithm. Testing is performed using the 10-node network in the previous experiment at $\varphi = 5^{\circ}$. The noise spectra before and after implementing ANC have been shown in Fig. 5 and the resulting noise reduction levels (difference of averaged noise power before and after implementing ANC) have been listed in Table 1. It can be seen that the proposed algorithm achieves a higher noise reduction level than the conventional ATC-FxLMS algorithm. 5.3. Evaluation of the SR parameter selection In this experiment, the simulated and theoretical total MSE results using different SR parameters are compared. In this simulation, a more general model is used to study the relationship between multitask level and the SR parameter. The primary paths $\mathbf{p}i$ are assumed to distributed on a circle of radius r centered at $\mathbf{p}0$, i.e. $\mathbf{p}i = \mathbf{p}0 + r\mathbf{g}i$, for $i = 1, 2,\ldots$, and $\mathbf{g}i$ is a Gaussian sequence of unit norm. This random walk model could generally describe changes in adaptive filtering [2123], where the radius controls the multitask problem level. The radius r is set to 0.1 and 0.01, respectively, to simulate strong and weak multitask problems. Short primary paths of length 20 have been used to evaluate the theoretical analysis. The secondary sources are assumed to be close to the corresponding loudspeakers and the paths from the ith loudspeaker to the ith microphone have a length of 5. The cross-channels have a length of 20. In the simulations, SNR is set to 10 dB. The step-size of the ATC-FxLMS algorithm is set to 0.001. The other settings are identical to those in the first experiment if not specified. Different SR parameters and radius are applied and the simulated and theoretical total MSE curves are shown in Fig. 6. It can be seen that the simulated and theoretical results generally agree well with each other under different settings. JR is slightly overestimated by (36) due to the approximations used. Comparing the optimal SR parameters in Fig. 6(a) and (b), where the radius is set to $r = 0.1$ and 0.01 respectively, it can be seen that the optimal value decreases as radius. When the radius is small as shown in Fig. 6(b), the system MSE rarely changes as the SR parameter. This does not mean that the diffusion control is useless. The advantage of the proposed algorithm is to be presented in the next experiment, where different methods are compared.

6. Conclusions

In this paper, we propose a Diff-FxLMS algorithm that adjusts the spatial regularization automatically. The VSR rule is derived from the penalized Lagrangian method. The theoretical analysis discloses the mechanism of variable diffusion control that balances between strong and weak combination under different conditions. Simulations verify the theoretical analysis and evaluate the performance of the proposed algorithm under different multitask problems.

Declaration of Competing Interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability The data that has been used is confidential.

Acknowledgment The work described in this study was supported by National Natural Science Foundation of China (61901174), Guangdong Basic and Applied Basic Research Foundation (2023A1511010771, 2019A1515010771), the State Key Laboratory of Mechanical System and Vibration (MSV202202), and the Fundamental Research Funds for the Central Universities (2022ZYGXZR104).

Appendix A:. Steady-State EMSE analysis

In this appendix, the steady-state EMSE is derived. To have a comprehensive result, some assumptions are made for the terms that determine the steady-state EMSE, i.e. we use the limit $\delta k(\infty) = \hat{\sigma}2wk$, and the small and uniform settings for the step-size and SR parameter. Recall that $\mathbf{G}\infty = \mathbf{I}KL$ □ $\xi 0\mathbf{B}\Delta$ and $\mathbf{B}\Delta = (\mathbf{I}K$ □ $\mathbf{B}) \otimes \mathbf{I}L$ under these assumptions. Then, we have $\mathbf{F} = \mathbf{R}F + (\mathbf{I}L2K2$ □ $\mathbf{R}F)(\xi 0\mathbf{B}\Delta \oplus \mathbf{B}\Delta$ □ $\xi 20\mathbf{B}\Delta \otimes \mathbf{B}\Delta)$, where $\oplus$ denotes the Kronecker sum. For small step-sizes *and SR parameters, **R**F is approximately equal to **R**F ≈ μ0**R**XˆX ⊕ **R**ˆXX and the matrix **F** can be written as **F** ≈ **R**F + ξ0**B**Δ ⊕ **B**Δ. Using the Woodbury Formula, it finds **F**□ 1 ≈ {**R**F[**I**L2K2 + ξ0**R**□ 1 F (**B**Δ ⊕ **B**Δ)]}□ 1 ≈ 1 μ0[**I**L2K2 □ ξ0**R**□ 1 F (**B**Δ ⊕ **B**Δ)](**R**XˆX ⊕ **R**ˆXX)□ 1 (A1) Next, using the small step-size assumption, νM is rewritten as νM = μ20(**I**L2K2 □ ξ0**B**Δ ⊕ **B**Δ + ξ20**B**Δ ⊗ **B**Δ)vec(ξ20 μ20**R**y +**D**Σ**R**XX) (A2) where Ry = BΔΔpΔpTBTΔ + Tr□ R□1XXBΔΔpΔpTBTΔ)RˆXˆX. Substituting (A.1)(A.2) into (32) and ignoring the denominator, it finds J ≈ 1Kμ0vec(**D**Σ**R**XX)T (**I**L2K2 □ ξ0**B**Δ ⊕ **B**Δ)T [**I**L2K2 □ ξ0**R**□ 1 F (**B**Δ ⊕ **B**Δ)](**R**XˆX ⊕ **R**ˆXX)□ 1vec(**R**xx) (A3) where the small SR parameter assumption has been used and the second order terms of ξ0 have been omitted.*

References

[1] Zhang Z, Wu M, Yin L, Gong C, Wang J, Zhou S, et al. Robust feedback controller combined with the remote microphone method for broadband active noise control in headrest. Appl Acoust 2022;195:108815. [2] Wu L, Wang L, Sun S, Sun X. Hybrid active and passive noise control in ventilation duct with internally placed microphones module. Appl Acoust 2022;188:108525. [3] An F, Liu B. Cascade biquad controller design for feedforward active noise control headphones considering incident noise from multiple directions. Appl Acoust 2022; 185:108430. [4] Niu F, Qiu XJ, Zhang D. Effects of active noise cancelling headphones on speech recognition. Appl Acoust 2020;165:107335. [5] Pradhan S, Qiu XJ. A 5-stage active control method with online secondary path modelling using decorrelated control signal. Appl Acoust 2020;164:107252. [6] Gong C, Wu M, Guo J, Cao Y, Zhang Z, Yang J. Modified narrowband active noise control system with frequency mismatch tolerance. Appl Acoust 2022;189:108598. [7] Ferrer M, Diego M, Pinero G, Gonzalez A. Active noise control over adaptive distributed networks. Signal Process 2015;107:82–95. [8] Ferrer M, Gonzalez A, De Diego M, Pi˜nero G. Distributed affine projection algorithm over acoustically coupled sensor networks. IEEE Trans Signal Process 2017;65(24):6423–34. [9] T. Li, H. Duan, S. Zhao, J. Lu, and I. S. Burnett, Distributed active noise control system based on a block diffusion FxLMS algorithm with bidirectional communication, arXiv:2212.13777, (2022). [10] Chu YJ, Mak CM. Diffusion control for multi-channel ANC systems using filtered-X algorithms. Proc. 24th Int. Cong. Sound Vib., London, UK, 23–27, Jul. 2017. 2017. [11] Sayed A. Adaptive networks. Proc IEEE 2014;102(4):460–97. [12] Chu YJ, Mak CM, Zhao Y, Chan SC, Wu M. Performance analysis of a diffusion control method for ANC systems and the network design. J Sound Vib 2020;475: 115273. [13] Elliott S, Stothers I, Nelson PA. A multiple error LMS algorithm and its application to the active control of sound and vibration. IEEE Trans Speech Signal Process 1987;35(10):1423–34. [14] Deleg`a R, Bernasconi G, Piroddi L. A novel cost-effective parallel narrowband ANC system with local secondary-path estimation. J Sound Vib 2017;401:311–25. [15] Zhang L, Tao J, Qiu X. Performance analysis of decentralized multi-channel feedback systems for active noise control in free space. Appl Acoust 2013;74(1): 181–8. [16] An F, Cao Y, Liu B. Optimized decentralized adaptive control of noise and vibration for periodic disturbances. J Acoust Soc Am 2018;144(4):EL275–80. [17] Zhang G, Tao J, Qiu X, Burnett I. Decentralized two-channel active noise control for single frequency by shaping matrix eigenvalues. IEEE Trans Audio Speech Lang Process 2019;27:44–52. [18] Pradhan S, Zhang G, Zhao S, Niwa K, Kleijn WB. On eigenvalue shaping for two-channel decentralized active noise control systems. Appl Acoust 2023;205:109260. [19] Chu YJ, Chan SC, Zhao Y, Wu M. Performance analysis of diffusion filtered-x algorithms in multitask ANC systems. Pro. Eusipco2019, La Coruna, Spain, Sep. 2–6. 2019. [20] Chu YJ, Chan SC, Mak CM, Wu M. A Diffusion FxLMS algorithm for multi-channel active noise control and variable spatial smoothing. In: Pro. ICASSP 2021, Toronto, Canada; 2021. p. 4695–9. [21] Chen J, Richard C, Sayed AH. Diffusion LMS over multitask networks. IEEE Trans Signal Process 2015;63(11):2733–48. [22] Cattivelli FS, Sayed AH. Diffusion LMS strategies for distributed estimation. IEEE Trans Signal Process 2010;58(3):1035–48. [23] Chen J, Richard C, Sayed AH. Multitask diffusion adaptation over networks. IEEE Trans Signal Process

2014;62(16):4129–44. [24] Nassif R, Vlaski S, Richard C, Sayed AH. A regularization framework for learning over multitask graphs. IEEE Signal Process Lett 2019;26(2):297–301. [25] Koppel A, Sadler BM, Ribeiro A. Proximity without consensus in online multiagent optimization. IEEE Trans Signal Process 2017;65(12):3062–77. [26] Wei Y, Gelfand SB, Krogmeier JV. Noise-constrained least mean squares algorithm. IEEE Trans Signal Process 2001;49(9):1961–70.