Supplementary Materials

Proof of Theorem 1

For $t \geq 1$, we first define some notations in a compact form as $\boldsymbol{w}^{\tau}(t-1) = [\boldsymbol{w}_1^{\tau}(t-1)^T, \ldots, \boldsymbol{w}_K^{\tau}(t-1)^T]^T$; $\boldsymbol{w}^{R}(t-1) = [\boldsymbol{w}_1^{R}(t-1)^T, \ldots, \boldsymbol{w}_K^{R}(t-1)^T]^T$; $s(t-1) = [s_{J_{1,t}}(\boldsymbol{w}_1^{\tau}(t-1))^T, \ldots, s_{J_{K,t}}(\boldsymbol{w}_K^{\tau}(t-1))^T]^T$. In addition, according to Assumption 1 and (5), the subgradients are also upper bounded, and thus we let $\|s(t-1)\|_{\infty} \leq C_s$. The average estimates vector is defined as $\bar{\boldsymbol{w}}(t-1) = \left[(\frac{1}{K}\mathbf{1}_K\mathbf{1}_K^T) \otimes I_{D+1}\right] \boldsymbol{w}^{\tau}(t-1)$, where I_{D+1} and $\otimes$ denote the $(D+1) \times (D+1)$ identity matrix and the Kronecker product, respectively. Then, the consensus error and the quantization error are defined as $\boldsymbol{\delta}(t-1) = \boldsymbol{w}^{\tau}(t-1) - \bar{\boldsymbol{w}}(t-1)$ and $\boldsymbol{e}(t-1) = \boldsymbol{w}^{\tau}(t-1) - \boldsymbol{w}^{R}(t-1)$, respectively. We further define the normalized consensus error and quantization error as $\tilde{\boldsymbol{\delta}}(t-1) = \frac{\boldsymbol{\delta}(t-1)}{\mu(t-1)}$ and $\tilde{\boldsymbol{e}}(t-1) = \frac{\boldsymbol{e}(t-1)}{\mu(t-1)}$, respectively.

Based on the above definitions, (8) and (10), the input of the quantizer is expressed by

$$
\begin{aligned}
&\Psi(t-1)\\
=&\frac{\boldsymbol{w}^{\tau}(t)-\boldsymbol{w}^{R}(t-1)}{\mu(t-1)}\\
=&\frac{1}{\mu(t-1)}[\boldsymbol{w}^{\tau}(t-1)-(h\boldsymbol{L}\otimes I_{D+1})\boldsymbol{w}^{R}(t-1)\\
&-\boldsymbol{w}^{R}(t-1)-h\mu(t-1)s(t-1)]\\
=&\frac{1}{\mu(t-1)}\left[(I_K+h\boldsymbol{L})\otimes I_{D+1}\right]\left[\boldsymbol{w}^{\tau}(t-1)-\boldsymbol{w}^{R}(t-1)\right]\\
&-\frac{1}{\mu(t-1)}(h\boldsymbol{L}\otimes I_{D+1})\left[\boldsymbol{w}^{\tau}(t-1)-\bar{\boldsymbol{w}}^{\tau}(t-1)\right]\\
&-hs(t-1)\\
=&\left[(I_K+h\boldsymbol{L})\otimes I_{D+1}\right]\tilde{\boldsymbol{e}}(t-1)-(h\boldsymbol{L}\otimes I_{D+1})\tilde{\boldsymbol{\delta}}(t-1)\\
&-hs(t-1),
\end{aligned}
\tag{S1}
$$

where $(h\boldsymbol{L}\otimes I_{D+1})\bar{\boldsymbol{w}}^{\tau}(t-1)=0$. Then, by employing (8) and (S1), we obtain

$$
\begin{aligned}
\tilde{\boldsymbol{e}}(t)=&\frac{\boldsymbol{w}^{\tau}(t)-\boldsymbol{w}^{R}(t)}{\mu(t)}\\
=&\frac{1}{\mu(t)}\left(\boldsymbol{w}^{\tau}(t)-\boldsymbol{w}^{R}(t-1)\right)\\
&-\frac{\mu(t-1)}{\mu(t)}Q\left[\frac{1}{\mu(t-1)}\left(\boldsymbol{w}^{\tau}(t)-\boldsymbol{w}^{R}(t-1)\right)\right]\\
=&\gamma(t-1)^{-1}\left\{\Psi(t-1)-Q\left[\Psi(t-1)\right]\right\}.
\end{aligned}
\tag{S2}
$$

In addition, using (10), we have

$$\begin{aligned}\tilde{\boldsymbol{\delta}}(t) =& \frac{\boldsymbol{w}^{\tau}(t) - \bar{\boldsymbol{w}}(t)}{\mu(t)} \\ =& \frac{1}{\mu(t)}[\boldsymbol{w}^{\tau}(t-1) - h\boldsymbol{L} \otimes I_{D+1}\boldsymbol{w}^{R}(t-1) \\ & - h\mu(t-1)s(t-1) - \bar{\boldsymbol{w}}(t-1) \\ & - \frac{h\mu(t-1)}{K}\mathbf{1}_K\mathbf{1}_K^T \otimes I_{D+1}s(k-1) \\ & + h\boldsymbol{L} \otimes I_{D+1}\boldsymbol{w}^{\tau}(t-1) - h\boldsymbol{L} \otimes I_{D+1}\boldsymbol{w}^{\tau}(t-1)] \\ =& \gamma(t-1)^{-1}[(I_K - h\boldsymbol{L}) \otimes I_{D+1}\tilde{\boldsymbol{\delta}}(t-1) \\ & + h\boldsymbol{L} \otimes I_{D+1}\tilde{\boldsymbol{e}}(t-1) - h\boldsymbol{Z} \otimes I_{D+1}s(t-1)],\end{aligned} \tag{S3}$$

where $\boldsymbol{Z} = I_K - \frac{1}{K}\mathbf{1}_K\mathbf{1}_K^T$.

Note that the adopted quantizer will not be saturated if the input $\Psi(t-1) \leq N + \frac{1}{2}$. Then, we prove that the quantizer will never be saturated during the iteration by mathematical induction.

For $t = 1$, by taking the infinite norm on both sides of (S1) and using (16), we have

$$\begin{aligned}\|\Psi(0)\|_{\infty} =& \|\frac{\boldsymbol{w}^{\tau}(0)}{\mu(0)} - hs(0)\|_{\infty} \\ \leq& \frac{C_w}{\mu(0)} + hC_s \\ \leq& N + \frac{1}{2}.\end{aligned} \tag{S4}$$

Then, the quantizer is unsaturated for $\Psi(0)$, and thus its quantization error is upper bounded by $\frac{1}{2}$. From (S2) and (S4), we further obtain

$$\begin{aligned}\|\tilde{\boldsymbol{e}}(1)\|_{\infty} =& \|\gamma(0)^{-1}\left\{\Psi(0) - Q\left[\Psi(0)\right]\right\}\|_{\infty} \\ \leq& \frac{1}{2\gamma(0)}.\end{aligned} \tag{S5}$$

We suppose $\sup\limits_{0\leq j\leq t-1} \|\Psi(j)\|_{\infty} \leq N + \frac{1}{2}$, and thus we have $\sup\limits_{0\leq j\leq t-1} \|\tilde{\boldsymbol{e}}(j+1)\|_{\infty} \leq \frac{1}{2\gamma(j)} \leq \frac{1}{2\gamma(0)}$. Then, we try to prove $\|\Psi(t)\|_{\infty} \leq N + \frac{1}{2}$.

Since the Laplacian matrix $\boldsymbol{L}$ is a real symmetric one, and thus it can be orthogonally decomposed as

$$\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T, \tag{S6}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}\{0, \lambda_2(\boldsymbol{L}), \ldots, \lambda_K(\boldsymbol{L})\}$ and $\boldsymbol{U} = \left[\mathbf{1}_K/\sqrt{K}, \phi_2, \ldots, \phi_K\right]$ is a orthogonal matrix with $\phi_k^T\boldsymbol{L} = \lambda_k(\boldsymbol{L})\phi_k^T$ $(k = 2, \ldots, K)$. We denote $\bar{\boldsymbol{\delta}}(t-1) = (\boldsymbol{U} \otimes I_{D+1})^{-1}\tilde{\boldsymbol{\delta}}(t-1) = \boldsymbol{U}^T \otimes I_{D+1}\tilde{\boldsymbol{\delta}}(t-1)$, which can be decomposed as

$$\left[\bar{\boldsymbol{\delta}}_1(t-1)^T, \bar{\boldsymbol{\delta}}_2(t-1)^T\right]^T = [\mathbf{1}_K/\sqrt{K}, \boldsymbol{\phi}]^T \otimes I_{D+1}\tilde{\boldsymbol{\delta}}(t-1), \tag{S7}$$

where $\boldsymbol{\phi} = [\phi_2, \ldots, \phi_K]$ and $\bar{\boldsymbol{\delta}}_1(t-1) = \mathbf{1}_K/\sqrt{K} \otimes I_{D+1}\tilde{\boldsymbol{\delta}}(t-1) = \mathbf{0}_{D+1}$. By multiplying $\boldsymbol{U}^T \otimes I_{D+1}$ on both sides of (S3) and using (S7), we obtain

$$\begin{aligned}
\bar{\boldsymbol{\delta}}_2(t) = &\left(\frac{I_{K-1} - h\bar{\boldsymbol{\Lambda}}}{\gamma(t-1)}\right) \otimes I_{D+1}\bar{\boldsymbol{\delta}}_2(t-1) \\
&+ \gamma(t-1)^{-1}h\left(\boldsymbol{\phi}^T\boldsymbol{L}\right) \otimes I_{D+1}\tilde{\boldsymbol{e}}(t-1) \\
&- \gamma(t-1)^{-1}h\left(\boldsymbol{\phi}^T\boldsymbol{Z}\right) \otimes I_{D+1}s(t-1),
\end{aligned} \tag{S8}$$

where $\bar{\boldsymbol{\Lambda}} = \mathrm{diag}\left\{\lambda_2(\boldsymbol{L}), \ldots, \lambda_K(\boldsymbol{L})\right\}$. Denoting $P_h(t-1) = \left(I_{K-1} - h\bar{\boldsymbol{\Lambda}}\right)/\gamma(t-1)$ and $\Phi(t, t-r) = P_h(t-1) \times \cdots \times P_h(t-r)$ with $\Phi(t,t) = I_K$, we rewrite (S8) recursively as

$$\begin{aligned}
\bar{\boldsymbol{\delta}}_2(t) =&\Phi(t,0)\boldsymbol{\phi}^T \otimes I_{D+1}\tilde{\boldsymbol{\delta}}(0) \\
&+ \gamma(0)^{-1}h\left(\Phi(t,1)\boldsymbol{\phi}^T\boldsymbol{L}\right) \otimes I_{D+1}\tilde{\boldsymbol{e}}(0) \\
&+ \sum_{r=0}^{t-2} h\left(\Phi(t,t-r)\boldsymbol{\phi}^T\boldsymbol{L}\right) \otimes I_{D+1}\frac{\tilde{\boldsymbol{e}}(t-1-r)}{\gamma(t-1-r)} \\
&- \sum_{r=0}^{t-1} h\left(\Phi(t,t-r)\boldsymbol{\phi}^T\boldsymbol{Z}\right) \otimes I_{D+1}\frac{s(t-1-r)}{\gamma(t-1-r)}.
\end{aligned} \tag{S9}$$

Multiplying $\boldsymbol{U} \otimes I_{D+1}$ on both sides of (S7), we have

$$\tilde{\boldsymbol{\delta}}(t) = \boldsymbol{\phi} \otimes I_{D+1}\bar{\boldsymbol{\delta}}_2(t). \tag{1}$$

Then, by multiplying $\boldsymbol{\phi} \otimes I_{D+1}$ on both sides of (S9), we obtain

$$\begin{aligned}
\tilde{\boldsymbol{\delta}}(t) = & \left(\boldsymbol{\phi}\Phi(t,0)\boldsymbol{\phi}^T\right) \otimes I_{D+1}\tilde{\boldsymbol{\delta}}(0) \\
& + \gamma(0)^{-1}h\left(\boldsymbol{\phi}\Phi(t,1)\boldsymbol{\phi}^T\boldsymbol{L}\right) \otimes I_{D+1}\tilde{\boldsymbol{e}}(0) \\
& + \sum_{r=0}^{t-2} h\left(\boldsymbol{\phi}\Phi(t,t-r)\boldsymbol{\phi}^T\boldsymbol{L}\right) \otimes I_{D+1}\frac{\tilde{\boldsymbol{e}}(t-1-r)}{\gamma(t-1-r)} \\
& - \sum_{r=0}^{t-1} h\left(\boldsymbol{\phi}\Phi(t,t-r)\boldsymbol{\phi}^T\boldsymbol{Z}\right) \otimes I_{D+1}\frac{s(t-1-r)}{\gamma(t-1-r)}.
\end{aligned} \tag{S10}$$

Obviously, the 2-norm of $\tilde{\boldsymbol{\delta}}(t)$ is upper bounded as

$$\begin{aligned}
& \|\tilde{\boldsymbol{\delta}}(t)\|_2 \\
\leq & \|\left(\boldsymbol{\phi}\Phi(t,0)\boldsymbol{\phi}^T\right) \otimes I_{D+1}\tilde{\boldsymbol{\delta}}(0)\|_2 \\
& + \|\gamma(0)^{-1}h\left(\boldsymbol{\phi}\Phi(t,1)\boldsymbol{\phi}^T\boldsymbol{L}\right) \otimes I_{D+1}\tilde{\boldsymbol{e}}(0)\|_2 \\
& + \left\|\sum_{r=0}^{t-2} h\left(\boldsymbol{\phi}\Phi(t,t-r)\boldsymbol{\phi}^T\boldsymbol{L}\right) \otimes I_{D+1}\frac{\tilde{\boldsymbol{e}}(t-1-r)}{\gamma(t-1-r)}\right\|_2 \\
& + \left\|\sum_{r=0}^{t-1} h\left(\boldsymbol{\phi}\Phi(t,t-r)\boldsymbol{\phi}^T\boldsymbol{Z}\right) \otimes I_{D+1}\frac{s(t-1-r)}{\gamma(t-1-r)}\right\|_2.
\end{aligned} \tag{S11}$$

Based on Assumptions 1, 3, (15) and (16), $\|\tilde{\boldsymbol{\delta}}(t)\|_2$ can be further upper bounded with some modifications:

$$
\begin{aligned}
\|\tilde{\boldsymbol{\delta}}(t)\|_2 \leq & \frac{\sqrt{K(D+1)}C_\delta}{\mu(0)}\left(\frac{\rho_h}{\gamma(0)}\right)^t \\
& + \frac{\sqrt{K(D+1)}h\lambda_K(\boldsymbol{L})C_w}{\mu(0)\rho_h}\left(\frac{\rho_h}{\gamma(0)}\right)^t \\
& + \frac{\sqrt{K(D+1)}h\lambda_K(\boldsymbol{L})}{2\gamma(0)(\gamma(0)-\rho_h)}\left(1-\left(\frac{\rho_h}{\gamma(0)}\right)^t\right) \\
& + \frac{\sqrt{K(D+1)}hC_s}{\gamma(0)-\rho_h}\left(1-\left(\frac{\rho_h}{\gamma(0)}\right)^t\right) \\
\leq & \max\left\{\frac{\sqrt{K(D+1)}\left(C_\delta\rho_h + h\lambda_K(\boldsymbol{L})C_w\right)}{\mu(0)\rho_h},\right. \\
& \left.\frac{\sqrt{K(D+1)}h\left(\lambda_K(\boldsymbol{L})+2\gamma(0)C_s\right)}{2\gamma(0)\left(\gamma(0)-\rho_h\right)}\right\} \\
= & \frac{\sqrt{K(D+1)}h\left(\lambda_K(\boldsymbol{L})+2\gamma(0)C_s\right)}{2\gamma(0)\left(\gamma(0)-\rho_h\right)}.
\end{aligned}
\tag{S12}
$$

Recalling the definition of $\tilde{\boldsymbol{\delta}}(t)$ and using (S12), we obtain

$$
\begin{aligned}
\|\boldsymbol{w}^\tau(t)-\bar{\boldsymbol{w}}(t)\|_2 = & \|\mu(t)\tilde{\boldsymbol{\delta}}(t)\|_2 \\
& \leq \mu(t)\frac{\sqrt{K(D+1)}h\left(\lambda_K(\boldsymbol{L})+2\gamma(0)C_s\right)}{2\gamma(0)\left(\gamma(0)-\rho_h\right)}.
\end{aligned}
\tag{S13}
$$

Combining $\|\tilde{\boldsymbol{e}}(t)\|_\infty \leq \frac{1}{2\gamma(0)}$, $\|s(t-1)\|_\infty \leq C_s$, (S1) and (S12), we have

$$
\begin{aligned}
& \|\Psi(t)\|_\infty \\
\leq & \frac{1+2h|\mathcal{N}^*|}{2\gamma(0)} + hC_s \\
& + h\lambda_K(\boldsymbol{L})\frac{\sqrt{K(D+1)}h\left(\lambda_K(\boldsymbol{L})+2\gamma(0)C_s\right)}{2\gamma(0)\left(\gamma(0)-\rho_h\right)}.
\end{aligned}
\tag{S14}
$$

Obviously, if (17) hold, we obtain $\|\Psi(t)\|_\infty \leq N + \frac{1}{2}$, which implies the quantizer is not saturated. Therefore, by induction, the adopted quantizer will never be saturated during the process, which can be mathematically expressed by

$$
\sup_{t\geq 1}\|\Psi(t-1)\|_\infty \leq N + \frac{1}{2}.
$$

Then, Theorem 1 is proven.