A. Proof of Theorem 2

In the beginning, we define several auxiliary variables, which will be used in this proof.

Let $\bar{z}(m) = \frac{1}{n} \sum_{i=1}^{n} z_i(m)$ and $\bar{g}(m) = \frac{1}{n} \sum_{i=1}^{n} \hat{g}_i(m).$ Then, we define

$$F_{m+1}(x) = \eta \bar{z}(m + 1) \top x + \|x\|_2^2$$

and $\bar{x}(m + 1) = \arg \min_{x \in \mathcal{K}_s} \bar{F}_{m+1}(x).$ Similarly, let $\hat{x}_i(m) = \arg \min_{x \in \mathcal{K}_s} \eta z_i(m) \top x + \|x\|_2^2.$

Moreover, we introduce the following two lemmas with respect to the theoretical guarantees of $\delta$-smoothed function.

Lemma 8 (Lemma 2.6 in Hazan (2016)) Let $f(x) : \mathbb{R}^d \rightarrow \mathbb{R}$ be convex and $G$-Lipschitz over a convex and compact set $\mathcal{K} \subset \mathbb{R}^d.$ Then, $f_\delta(x)$ is convex and $G$-Lipschitz over $\mathcal{K}_\delta$, and it holds that $|f_\delta(x) - f(x)| \leq \delta G$ for any $x \in \mathcal{K}_\delta.$

Lemma 9 (Lemma 4 in Garber & Koren (2019)) Let $f(x) : \mathbb{R}^d \rightarrow \mathbb{R}$ be convex and suppose that all subgradients of $f$ are upper bounded by $G$ in $\ell_2$-norm over a convex and compact set $\mathcal{K} \subset \mathbb{R}^d.$ For any $x \in \mathcal{K}_\delta$, $\|\nabla f_\delta(x)\|_2 \leq G.$

We first assume that for all $i \in V$ and $m = 1, \cdots, B,$

$$\|\bar{g}_i(m)\|_2 \leq \beta.$$

Let $x^* \in \arg \min_{x \in \mathcal{K}} \sum_{t=1}^{T} f_t(x)$ and $\bar{x}^* = (1 - \delta/r)x^*.$ For any $i, j \in V$, we have

$$\sum_{t=1}^{T} f_{t,j}(\bar{y}_i(t)) - \sum_{t=1}^{T} f_{t,j}(x^*) \leq \sum_{t=1}^{T} (f_{t,j}(x_i(m_t)) + \delta u_i(t)) - \sum_{t=1}^{T} f_{t,j}(x^*)$$

$$\leq \sum_{t=1}^{T} (f_{t,j}(x_i(m_t)) + G\|u_i(t)\|_2) - \sum_{t=1}^{T} (f_{t,j}(\bar{x}^*) - G\|\bar{x}^* - x^*\|_2)$$

$$\leq \sum_{t=1}^{T} f_{t,j}(x_i(m_t)) - \sum_{t=1}^{T} f_{t,j}(\bar{x}^*) + \delta GT + \frac{\delta GRT}{r}$$

where the first inequality is due to Assumption 1 and the third inequality is due to Lemma 8.

Then, similar to the proof of Theorem 1, we derive an upper bound of $\|x_i(m) - \bar{x}(m)\|_2$ by further introducing the following lemma.

Lemma 10 Let $\bar{x}_i(m) = \arg \min_{x \in \mathcal{K}_s} F_{m,i}(x)$, for $m \in [B].$ Assume $\|\bar{g}_i(m)\|_2 \leq \beta$ for any $i \in V$ and $m \in [B].$ Algorithm 3 with $\epsilon \leq 8R^2$ and $L = \frac{\alpha^2 R^2}{\beta^2}$ has

$$F_{m,i}(x_i(m)) - F_{m,i}(\bar{x}_i(m)) \leq \epsilon$$

for any $i \in V$ and $m \in [B]$, where $\alpha = \frac{1 + \sigma(x)}{1 - \sigma(x)} \sqrt{n} + 1.$

Applying Lemma 2 with $\|\bar{g}_i(m)\|_2 \leq \beta$, we have

$$\|z_i(m) - \bar{z}(m)\|_2 \leq \alpha' \beta$$ 

(17)
where $\alpha' = \frac{\sqrt{n}}{1-\sigma_2^2(\mathbf{P}^T)}$.

Furthermore, applying Lemma 3 with (17), we have
\[
\|\overline{x}_i(m) - \overline{x}(m)\|_2 \leq \eta\|z_i(m) - z(m)\|_2 \leq \eta\alpha' \beta
\]
which implies that
\[
\|x_i(m) - x(m)\|_2 \leq \|x_i(m) - \overline{x}_i(m)\|_2 + \|\overline{x}_i(m) - x(m)\|_2
\leq \sqrt{F_{m,i}(x_i(m))} - F_{m,i}(\overline{x}_i(m)) + \eta \alpha' \beta
\leq \sqrt{\epsilon} + \eta \alpha' \beta
\]
where the second inequality is due to the fact $F_{m,i}(x)$ is 2-strongly convex and (5), and the last inequality is due to Lemma 10.

For brevity, let $\epsilon' = \sqrt{\epsilon} + \eta \alpha' \beta$. Then, we can use (18) to bound the first term in the right side of (16) as
\[
\sum_{t=1}^{T} (\hat{f}_{t,j,\delta}(x_i(m_t)) - \hat{f}_{t,j,\delta}(\overline{x}))
\leq \sum_{t=1}^{T} (\hat{f}_{t,j,\delta}(\overline{x}(m_t)) - \hat{f}_{t,j,\delta}(\overline{x})) + \sum_{i=1}^{T} \|\overline{x}(m_t) - x_i(m_t)\|_2
\leq \sum_{t=1}^{T} (\hat{f}_{t,j,\delta}(x_j(m_t)) - \hat{f}_{t,j,\delta}(\overline{x})) + \sum_{i=1}^{T} \|\overline{x}(m_t) - x_j(m_t)\|_2 + GT \epsilon'
\leq \sum_{t=1}^{T} \nabla (\hat{f}_{t,j,\delta}(x_j(m_t)))^T (x_j(m_t) - \overline{x}) + 2GT \epsilon'
\leq \sum_{t=1}^{T} \|\nabla (\hat{f}_{t,j,\delta}(x_j(m_t)))\|_2 \|x_j(m_t) - \overline{x}\|_2 + \sum_{t=1}^{T} \|\overline{x}(m_t) - x_j(m_t)\|_2 + 2GT \epsilon'
\leq \sum_{t=1}^{T} \|\nabla (\hat{f}_{t,j,\delta}(x_j(m_t)))\|_2 \|\overline{x}(m_t) - x_j(m_t)\|_2 + \sum_{t=1}^{T} \|\overline{x}(m_t) - \overline{x}\|_2 + 2GT \epsilon'
\leq \sum_{t=1}^{T} \|\nabla (\hat{f}_{t,j,\delta}(x_j(m_t)))\|_2 \|\overline{x}(m_t) - \overline{x}\|_2 + 3GT \epsilon'
\]
where the third inequality is due to the convexity of $\hat{f}_{t,j,\delta}(x)$ and the fifth inequality is due to Lemma 9.

Combining (16), (19) and $\epsilon' = \sqrt{\epsilon} + \eta \alpha' \beta$, for any $i \in V$, we have
\[
\sum_{t=1}^{T} \sum_{j=1}^{n} f_{t,j}(y_i(t)) - \sum_{t=1}^{T} \sum_{j=1}^{n} f_{t,j}(x^*)
\leq \sum_{t=1}^{T} \sum_{j=1}^{n} \nabla (\hat{f}_{t,j,\delta}(x_j(m_t)))^T (\overline{x}(m_t) - \overline{x}^*) + 3\delta n GT + \frac{\delta n GT}{r} + 3nGT (\sqrt{\epsilon} + \eta \alpha' \beta).
\]

Moreover, to bound $\sum_{t=1}^{T} \sum_{j=1}^{n} \nabla (\hat{f}_{t,j,\delta}(x_j(m_t)))^T (\overline{x}(m_t) - \overline{x}^*)$, we introduce the following lemma.

**Lemma 11** Let $\overline{z}(m) = \frac{1}{n} \sum_{i=1}^{n} z_i(m)$ and $\overline{g}(m) = \frac{1}{n} \sum_{i=1}^{n} \overline{g}_i(m)$. Moreover, we define
\[
\overline{F}_{m+1}(x) = \eta \overline{z}(m+1)^T x + \|x\|_2^2
\]
and $\bar{x}(m+1) = \arg\min_{x \in K_i} \tilde{F}_{m+1}(x)$. Assume $\|\tilde{g}_i(m)\|_2 \leq \beta$ for any $i \in V$ and $m \in [B]$, with probability at least $1 - \gamma$. Algorithm 3 has

$$
\sum_{t=1}^{T} \sum_{j=1}^{n} \nabla f_{t,j}(x_{t,j}(m_i)) (\bar{x}(m_i) - \bar{x}^*) \leq 2nR(KG + \beta) \sqrt{2B \ln \frac{1}{\eta} + \frac{nR^2}{\eta} + n\eta B^2}
$$

where $\bar{x}^* = (1 - \delta/r)x^*$ and $x^* \in \arg\min_{x \in K} \sum_{t=1}^{T} f_t(x)$.

According to Lemma 11, assume that $\|\tilde{g}_i(m)\|_2 \leq \beta$ for any $i \in V$ and $m \in [B]$, with probability at least $1 - \gamma$, we have

$$
\sum_{t=1}^{T} \sum_{j=1}^{n} f_{t,j}(y_t(t)) - \sum_{t=1}^{T} \sum_{j=1}^{n} f_{t,j}(x^*) 
\leq 2nR(KG + \beta) \sqrt{2B \ln \frac{1}{\eta} + \frac{nR^2}{\eta} + n\eta B^2 + 3\delta n GT + \frac{\delta n G T}{r} + 3nG T \left( \sqrt{\epsilon + \eta \alpha' \beta} \right)}.
$$

Substituting $\eta = \frac{\epsilon R}{\alpha T d M} T^{-3/4}$, $\delta = cT^{-1/4}$, $\epsilon = 4R^2 T^{-1/2}$, $\beta = \alpha_T \frac{d M \sqrt{K}}{\delta}$, $KG$ and $K = T^{1/2}$ into the above inequality, we have

$$
R_{T,i} \leq 2nR \left( 2G + \frac{\alpha_T d M c}{c} \right) \sqrt{2 \ln \frac{1}{\gamma} T^{3/4} + \frac{\alpha_T n d M R}{c} T^{3/4}} + n \left( R + \frac{c RG}{\alpha_T d M} \right) \left( \frac{\alpha_T d M c}{c} + G \right) T^{3/4} + 3cnG T^{3/4} + \frac{cn G R}{r} T^{3/4} + 6nG T^{3/4}
\leq O \left( \alpha_T T^{3/4} \right).
$$

Let $\mathcal{A}$ denote the event of $\|\tilde{g}_i(m)\|_2 \leq \beta$, $\forall i \in V$, $m \in [B]$. Because we have used the event $\mathcal{A}$ as a fact, the above result should be formulated as

$$
\Pr \left( R_{T,i} \leq O \left( \alpha_T T^{3/4} \right), \mathcal{A} \right) \geq 1 - \gamma.
$$

Furthermore, we introduce the following lemma with respect to the probability of the event $\mathcal{A}$.

**Lemma 12** For all $i \in V$ and $m \in [B]$, Algorithm 3 has

$$
\|\tilde{g}_i(m)\|_2 \leq \left( 1 + \sqrt{8 \ln \frac{n B}{\gamma}} \right) \frac{d M \sqrt{K}}{\delta} + KG
$$

with probability at least $1 - \gamma$.

Then, applying Lemma 12 with $B = T / K = \sqrt{T}$, we have

$$
\Pr (\mathcal{A}) \geq 1 - \gamma.
$$

Combining (20) with (21), we complete the proof.

### B. Proof of Lemma 10

For $m = 1$, because $x_i(1) = \hat{x}_i(1) = \arg\min_{x \in K_i} \|x\|_2^2$, we have

$$
F_{1,j}(x_i(1)) - F_{1,j}(\hat{x}_i(1)) = 0 \leq \epsilon.
$$
Then, for $m = 2$, we have
\[
F_{m,i}(x_i(m-1)) - F_{m,i}^{\hat{\theta}_i}(m) = F_{m-1,i}(x_i(m-1)) + \eta(z_i(m) - z_i(m-1))^\top x_i(m-1) \\
- F_{m-1,i}(\hat{\theta}_i(m)) - \eta(z_i(m) - z_i(m-1))^\top \hat{\theta}_i(m) \\
\leq F_{m-1,i}(x_i(m-1)) - F_{m-1,i}(\hat{\theta}_i(m)) \\
+ \eta(z_i(m) - z_i(m-1))^\top (x_i(m-1) - \hat{\theta}_i(m)) \\
\leq \epsilon + \eta||z_i(m) - z_i(m-1)||_2||x_i(m-1) - \hat{\theta}_i(m)||_2
\] (23)
where the first inequality is due to $\hat{\theta}_i(m-1) = \text{argmin}_{x \in \mathbb{K}_i} F_{m-1,i}(x)$ and the fourth inequality is due to that $F_{m-1,i}(x)$ is 2-strongly convex and (5).

Moreover, because for each $m = 1, \cdots, B$, $F_{m,i}(x)$ is 2-strongly convex, we also have
\[
||\hat{\theta}_i(m-1) - \hat{\theta}_i(m)||_2^2 \leq F_{m,i}(\hat{\theta}_i(m-1)) - F_{m,i}(\hat{\theta}_i(m)) = F_{m-1,i}(\hat{\theta}_i(m-1)) + \eta(z_i(m) - z_i(m-1))^\top \hat{\theta}_i(m-1) \\
- F_{m-1,i}(\hat{\theta}_i(m)) - \eta(z_i(m) - z_i(m-1))^\top \hat{\theta}_i(m) \\
= F_{m-1,i}(\hat{\theta}_i(m-1)) - F_{m-1,i}(\hat{\theta}_i(m)) \\
+ \eta(z_i(m) - z_i(m-1))^\top (\hat{\theta}_i(m-1) - \hat{\theta}_i(m)) \\
\leq \eta||z_i(m) - z_i(m-1)||_2||\hat{\theta}_i(m-1) - \hat{\theta}_i(m)||_2
\]
which further implies that
\[
||\hat{\theta}_i(m-1) - \hat{\theta}_i(m)||_2 \leq \eta||z_i(m) - z_i(m-1)||_2.
\] (24)
For $m \in [B]$, applying Lemma 6 with $\|\hat{\theta}_i(m)\|_2 \leq \beta$, we have
\[
||z_i(m+1) - z_i(m)||_2 \leq \alpha \beta.
\] (25)
Substituting (24) and (25) into (23), we have
\[
F_{m,i}(x_i(m-1)) - F_{m,i}(\hat{\theta}_i(m)) \leq \epsilon + \eta||z_i(m) - z_i(m-1)||_2\sqrt{\epsilon} + \eta^2||z_i(m) - z_i(m-1)||_2^2 \\
\leq \epsilon + \eta \alpha \beta \sqrt{\epsilon} + \eta^2 \alpha^2 \beta^2.
\]

According to Algorithm 3, we have $x_i(m) = \text{CGSC}(\mathbb{K}_i, \epsilon, L, F_{m,i}(x), x_i(m-1))$. Because $F_{m,i}(x)$ is 2-smooth and 2-strongly convex, $\epsilon \leq 8R^2$ and $L = \frac{16\alpha^2}{\epsilon^2}(\eta \alpha \beta \sqrt{\epsilon} + \eta^2 \alpha^2 \beta^2)$, applying Lemma 7 with $\mathcal{K}' = \mathcal{K}_i$, we have
\[
F_{m,i}(x_i(m)) - F_{m,i}(\hat{\theta}_i(m)) \leq \epsilon
\]
for $m = 2$. By induction, we can complete the proof for $m = 1, \cdots, B$.

**C. Proof of Lemma 11**

We first introduce the classical Azuma’s inequality (Azuma, 1967) for martingales in the following lemma.
Lemma 13 Suppose $D_1, \ldots, D_p$ is a martingale difference sequence and
\[
|D_j| \leq c_j
\]
almost surely. Then, we have
\[
\Pr \left( \sum_{j=1}^{\infty} D_j \geq \Delta \right) \leq \exp \left( \frac{-\Delta^2}{2 \sum_{j=1}^{\infty} c_j^2} \right).
\]
To apply Lemma 13, with $\mathcal{T}_m = \{(m-1)K + 1, \ldots, mK\}$, we define
\[
D_m = \sum_{t \in \mathcal{T}_m} \sum_{j=1}^{n} \left( \nabla \hat{f}_{t,j,\delta}(x_j(m)) - \hat{g}_j(t) \right)^\top (x(m) - x^*)
\]
(26)
\[
= \sum_{j=1}^{n} \left( \sum_{t \in \mathcal{T}_m} \nabla \hat{f}_{t,j,\delta}(x_j(m)) - \hat{g}_j(m) \right)^\top (x(m) - x^*).
\]
According to Algorithm 3 and Lemma 1, we have
\[
\mathbb{E} \left[ D_m | x_1(m), \ldots, x_n(m), x(m) \right] = 0
\]
which further implies that $D_1, \ldots, D_B$ is a martingale difference sequence with
\[
|D_m| = \left| \sum_{j=1}^{n} \left( \sum_{t \in \mathcal{T}_m} \nabla \hat{f}_{t,j,\delta}(x_j(m)) - \hat{g}_j(m) \right)^\top (x(m) - x^*) \right|
\]
\[
\leq \sum_{j=1}^{n} \left\| \sum_{t \in \mathcal{T}_m} \nabla \hat{f}_{t,j,\delta}(x_j(m)) - \hat{g}_j(m) \right\|_2 \left\| (x(m) - x^*) \right\|_2
\]
\[
\leq 2R \sum_{j=1}^{n} \left( \left\| \sum_{t \in \mathcal{T}_m} \nabla \hat{f}_{t,j,\delta}(x_j(m)) \right\|_2 + \left\| \hat{g}_j(m) \right\|_2 \right)
\]
\[
\leq 2R \sum_{j=1}^{n} \sum_{t \in \mathcal{T}_m} \left\| \nabla \hat{f}_{t,j,\delta}(x_j(m)) \right\|_2 + 2nR\beta
\]
\[
\leq 2nR\ln G + 2nR\beta
\]
where the last inequality is due to Lemma 9.

Then, applying Lemma 13 with $\Delta = 2nR(KG + \beta) \sqrt{2B \ln \frac{1}{\gamma}}$, with probability at least $1 - \gamma$, we have
\[
\sum_{m=1}^{B} D_m \leq \Delta = 2nR(KG + \beta) \sqrt{2B \ln \frac{T}{\gamma}}.
\]
(27)

Additionally, combining (26) with $\hat{g}(m) = \frac{1}{n} \sum_{i=1}^{n} \hat{g}_i(m)$, we further have
\[
\sum_{t=1}^{T} \sum_{j=1}^{n} \nabla \hat{f}_{t,j,\delta}(x_j(m_t))^\top (x(m_t) - \bar{x}^*) = \sum_{m=1}^{B} D_m + n \sum_{m=1}^{B} \hat{g}(m)^\top (x(m) - \bar{x}^*).
\]
(28)

Therefore, we still need to bound $\sum_{m=1}^{B} \hat{g}(m)^\top (x(m) - \bar{x}^*)$. According to Assumption 4, it is easy to verify that
\[
\bar{z}(m+1) = \frac{1}{n} \sum_{i=1}^{n} z_i(m+1) = \frac{1}{n} \sum_{i=1}^{n} \left( \sum_{j \in N_i} P_{ij} z_j(m) + \hat{g}_i(m) \right) = \bar{z}(m) + \bar{g}(m).
\]
Moreover, according to the definition, we have
\[ \tilde{x}(m + 1) = \arg\min_{x \in K_m} \tilde{F}_{m+1}(x) = \arg\min_{x \in K_m} \eta \tilde{z}(m+1)^\top x + \|x\|^2. \]

So, applying Lemma 5 with the linear loss functions \( \{g(m)^\top x\}_{m=1}^B \), the decision set \( K = K_\delta \) and the regularizer \( \mathcal{R}(x) = \frac{\|x\|^2}{\eta} \), we have
\[ \sum_{m=1}^B g(m)^\top (\bar{x}(m) - \bar{x}) \leq \frac{\|\bar{x}\|^2}{\eta} - 0 + \sum_{m=1}^B g(m)^\top (\bar{x}(m) - \bar{x}(m+1)) \leq \frac{R^2}{\eta} + \sum_{m=1}^B \|g(m)\|_2 \|\bar{x}(m) - \bar{x}(m+1)\|_2. \]  
\[ (29) \]

Then, it is easy to verify that \( \tilde{F}_{m+1}(x) \) is 2-strongly convex, which implies that
\[ \|\bar{x}(m) - \bar{x}(m+1)\|^2 \leq \tilde{F}_{m+1}(\bar{x}(m)) - \tilde{F}_{m+1}(\bar{x}(m+1)) \]
\[ = \tilde{F}_m(\bar{x}(m)) + \eta g(m)^\top \bar{x}(m) - \tilde{F}_m(\bar{x}(m+1)) - \eta g(m)^\top \bar{x}(m+1) \]
\[ = \tilde{F}_m(\bar{x}(m)) - \tilde{F}_m(\bar{x}(m+1)) + \eta g(m)^\top (\bar{x}(m) - \bar{x}(m+1)) \]
\[ \leq \eta \|g(m)\|_2 \|\bar{x}(m) - \bar{x}(m+1)\|_2. \]
\[ (30) \]

The above inequality can be simplified as
\[ \|\bar{x}(m) - \bar{x}(m+1)\|_2 \leq \eta \|g(m)\|_2. \]
\[ (31) \]

Substituting (30) into (29), we have
\[ \sum_{m=1}^B g(m)^\top (\bar{x}(m) - \bar{x}) \leq \frac{R^2}{\eta} + \eta \sum_{m=1}^B \|g(m)\|^2 \]
\[ = \frac{R^2}{\eta} + \eta \sum_{m=1}^B \left\| \frac{1}{n} \sum_{i=1}^n \hat{g}_i(m) \right\|^2 \]
\[ \leq \frac{R^2}{\eta} + \eta \sum_{m=1}^B \sum_{i=1}^n \|\hat{g}_i(m)\|^2 \]
\[ = \frac{R^2}{\eta} + \eta B \beta^2. \]  
\[ (31) \]

Finally, substituting (27) and (31) into (28), we complete the proof.

**D. Proof of Lemma 12**

According to Algorithm 3, for any \( i \in V \) and \( m = 1, \cdots, B \), conditioned on \( x_i(m) \),
\[ g_i((m-1)K+1), \cdots, g_i(mK) \]
are \( K \) independent random vectors.

For brevity, for \( j = 1, \cdots, K \), let
\[ X_j = g_i(t_j) \]
where \( t_j = (m-1)K + j \), and let \( N = \left\| \sum_{j=1}^K X_j \right\|_2 \), \( \bar{S}_j = \sum_{k \neq j} X_k \) and \( \bar{X}_j \) be the set
\[ \{X_1, \cdots, X_{j-1}, X_{j+1}, \cdots, X_K\}. \]
To bound $N$ by using Lemma 13, we define $X_0 = \emptyset$, $X_j = \{X_1, \cdots, X_j\}$ for $j \geq 1$ and a sequence $D_1, \cdots, D_K$ as

$$D_j = \mathbb{E}[N|X_j] - \mathbb{E}[N|X_{j-1}].$$

It is not hard to verify that

$$\mathbb{E}[D_j|X_{j-1}] = \mathbb{E}[(\mathbb{E}[N|X_j] - \mathbb{E}[N|X_{j-1}])|X_{j-1}] = 0$$

which implies that $D_1, \cdots, D_K$ is a martingale difference sequence.

Moreover, we have

$$|D_j| = |\mathbb{E}[N|X_j] - \mathbb{E}[N|X_{j-1}]| \leq \sup_{K_j} |N - \mathbb{E}[N|\tilde{X}_j]|. \quad (32)$$

Using the triangle inequality, we have

$$N \leq \|\tilde{S}_j\|_2 + \|X_j\|_2 \quad \text{and} \quad N \geq \|\tilde{S}_j\|_2 - \|X_j\|_2. \quad (33)$$

According to the Algorithm 3, we have

$$\|X_j\|_2 = \left\| \frac{d}{\delta} \int_{t_j,i}(y_i(t_j))u_i(t_j) \right\|_2 \leq \frac{dM}{\delta}. \quad (34)$$

Let $\Delta = \sqrt{\frac{KdM}{\delta}} \sqrt{8 \ln \frac{nB}{\gamma}}$. Then, applying Lemma 13, with probability at least $1 - \frac{1}{\Delta^2}$, we have

$$N - \mathbb{E}[N] = \mathbb{E}[N|X_K] - \mathbb{E}[N|X_0] = \sum_{j=1}^{K} D_j \leq \frac{\sqrt{KdM}}{\delta} \sqrt{8 \ln \frac{nB}{\gamma}}$$

which implies that

$$\|\tilde{g}_i(m)\|_2 = N \leq \frac{\sqrt{KdM}}{\delta} \sqrt{8 \ln \frac{nB}{\gamma}} + \mathbb{E}[N] \leq \frac{\sqrt{KdM}}{\delta} \sqrt{8 \ln \frac{nB}{\gamma}} + \sqrt{\mathbb{E}[N^2]}. \quad (35)$$

It is easy to provide an upper bound of $\mathbb{E}[N^2]$ by following the proof of Lemma 5 in Garber & Kretzu (2019). We include the detailed proof for completeness.

According to the definition, we have

$$\mathbb{E}[N^2] = \mathbb{E} \left[ \mathbb{E} \left[ \sum_{j=1}^{K} X_j^T X_j \left| x_i(m) \right. \right] \right] + \mathbb{E} \left[ \mathbb{E} \left[ \sum_{j=1}^{K} \sum_{k \neq j} X_j^T X_k \left| x_i(m) \right. \right] \right]$$

$$\leq \mathbb{E} \left[ \mathbb{E} \left[ \sum_{j=1}^{K} \|X_j\|_2^2 \left| x_i(m) \right. \right] \right] + \mathbb{E} \left[ \mathbb{E} \left[ \sum_{j=1}^{K} \sum_{k \neq j} \mathbb{E} \left[ X_j \left| x_i(m) \right. \right]^T \mathbb{E} \left[ X_k \left| x_i(m) \right. \right] \right] \right]$$

$$\leq \mathbb{E} \left[ \mathbb{E} \left[ \sum_{j=1}^{K} \|X_j\|_2^2 \left| x_i(m) \right. \right] \right] + \mathbb{E} \left[ \sum_{j=1}^{K} \sum_{k \neq j} \|\mathbb{E} \left[ X_j \left| x_i(m) \right. \right]\|_2 \|\mathbb{E} \left[ X_k \left| x_i(m) \right. \right]\|_2 \right]$$

$$\leq K \left( \frac{dM}{\delta} \right)^2 + \mathbb{E} \left[ \sum_{j=1}^{K} \sum_{k \neq j} \|\mathbb{E} \left[ X_j \left| x_i(m) \right. \right]\|_2 \|\mathbb{E} \left[ X_k \left| x_i(m) \right. \right]\|_2 \right]$$

$$\leq K \left( \frac{dM}{\delta} \right)^2 + (K^2 - K)G^2$$

$$\leq K \left( \frac{dM}{\delta} \right)^2 + K^2 G^2$$
where the third inequality is due to Lemmas 1 and 9.

Combining the above inequality with (35), with probability at least \(1 - \frac{\gamma}{nB}\), we have

\[
\|\hat{g}_i(m)\|_2 \leq \left(1 + \sqrt{8 \ln \frac{nB}{\gamma}}\right) \frac{dM \sqrt{K}}{\delta} + KG.
\]

Finally, using the union bound, we complete the proof for all \(i \in V\) and \(m = 1, \cdots, B\).