Appendix

A Preliminary Lemmas

In this section, we give preliminary lemmas which are also used in the proof the proofs of Lemmas 1 and 2. While Lemmas 4 and 5 depend on the optimality conditions of subproblems, Lemmas 3, 6 and 7 require Assumption 1.

Lemma 3. *It holds that* $\forall z_1, z_2 \in R^q$,

$$h(z_1) \leq h(z_2) + \nabla h(z_2)^T(z_1 - z_2) + (H/2)\|z_1 - z_2\|^2, \ -h(z_1) \leq -h(z_2) - \nabla h(z_2)^T(z_1 - z_2) + (H/2)\|z_1 - z_2\|^2.$$

Proof. Because $h(z)$ is Lipschitz differentiable by Assumption 1, so is $-h(z)$. Therefore, this lemma is proven exactly as same as Lemma 2.1 in [4].

Lemma 4. *It holds that* $y^k = \nabla h(z^k)$ *for all* $k \in \mathbb{N}$.

Proof. The optimality condition for the problem with regard to z^k gives rise to

$$\nabla h(z^k) - y^{k-1} - \rho(\sum\nolimits_{i=1}^{n} A_i x_i^k - z^k) = 0.$$

Because $y^k = y^{k-1} + \rho(\sum A_i x_i^k - z^k)$, we have $y^k = \nabla h(z^k)$.

Lemma 5. *It holds that for* $\forall k \in \mathbb{N}$,

$$L_\rho(\cdots, x_{i-1}^{k+1}, x_i^k, \cdots) - L_\rho(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots) \geq (\rho/2)\|A_i x_i^k - A_i x_i^{k+1}\|_2^2. \tag{8}$$

Proof.

$$\begin{aligned}
&L_\rho(\cdots, x_{i-1}^{k+1}, x_i^k, \cdots) - L_\rho(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots)\\
&= f(\cdots, x_{i-1}^{k+1}, x_i^k, \cdots) - f(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots)\\
&+ (y^k)^T(A_i x_i^k - A_i x_i^{k+1}) + (\rho/2)\|\sum\nolimits_{j=1}^{i-1} A_j x_j^{k+1} + \sum\nolimits_{j=i}^{n} A_j x_j^k - z^k\|_2^2 - (\rho/2)\|\sum\nolimits_{j=1}^{i} A_j x_j^{k+1} + \sum\nolimits_{j=i+1}^{n} A_j x_j^k - z^k\|_2^2\\
&= f(\cdots, x_{i-1}^{k+1}, x_i^k, \cdots) - f(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots)\\
&+ (y^k)^T(A_i x_i^k - A_i x_i^{k+1}) + (\rho/2)\|A_i x_i^k - A_i x_i^{k+1}\|_2^2 + \rho(\sum\nolimits_{j=1}^{i} A_j x_j^{k+1} + \sum\nolimits_{j=i+1}^{n} A_j x_j^k - z^k)^T(A_i x_i^k - A_i x_i^{k+1})\\
&= f(\cdots, x_{i-1}^{k+1}, x_i^k, \cdots) - f(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots)\\
&+ (A_i^T y^k + \rho A_i^T(\sum\nolimits_{j=1}^{i} A_j x_j^{k+1} + \sum\nolimits_{j=i+1}^{n} A_j x_j^k - z^k))^T(x_i^k - x_i^{k+1}) + (\rho/2)\|A_i x_i^k - A_i x_i^{k+1}\|_2^2.
\end{aligned}$$

where the second equality follows from the cosine rule: $\|b+c\|^2 - \|a+c\|^2 = \|b-a\|^2 + 2(a+c)^T(b-a)$ with $a = A_i x_i^{k+1}, b = A_i x_i^k$ and $c = \sum_{j=1}^{i-1} A_j x_j^{k+1} + \sum_{j=i+1}^{n} A_j x_j^k - z^k$.
The optimality condition of x_i^{k+1} leads to

$$\begin{aligned}
&0 \in \partial_{x_i} L_\rho(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots)\\
&= \partial_{x_i} f(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots) + A_i^T y^k + \rho A_i^T(\sum\nolimits_{j=1}^{i} A_j x_j^{k+1} + \sum\nolimits_{j=i+1}^{n} A_j x_j^k - z^k)\\
&- A_i^T y^k - \rho A_i^T(\sum\nolimits_{j=1}^{i} A_j x_j^{k+1} + \sum\nolimits_{j=i+1}^{n} A_j x_j^k - z^k) \in \partial_{x_i} f(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots).
\end{aligned}$$

We have the following result according to the definition of subgradient

$$\begin{aligned}
&f(\cdots, x_{i-1}^{k+1}, x_i^k, \cdots)\\
&\geq f(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots) + (-A_i^T y^k - \rho A_i^T(\sum\nolimits_{j=1}^{i} A_j x_j^{k+1} + \sum\nolimits_{j=i+1}^{n} A_j x_j^k - z^k))^T(x_i^{k+1} - x_i^k)\\
&= f(\cdots, x_i^{k+1}, x_{i+1}^k, \cdots) + (A_i^T y^k + \rho A_i^T(\sum\nolimits_{j=1}^{i} A_j x_j^{k+1} + \sum\nolimits_{j=i+1}^{n} A_j x_j^k - z^k))^T(x_i^k - x_i^{k+1}).
\end{aligned}$$

Therefore, the lemma is proved.

Lemma 6. *If $\rho > 2H$ so that $C_1 = \rho/2 - H/2 - H^2/\rho > 0$, then it holds that*

$$L_\rho(\cdots, x_n^{k+1}, z^k, y^k) - L_\rho(\cdots, x_n^{k+1}, z^{k+1}, y^{k+1}) \geq C_1 \|z^{k+1} - z^k\|_2^2. \tag{9}$$

Proof.

$L_\rho(x_1^{k+1}, \cdots, x_n^{k+1}, z^k, y^k) - L_\rho(x_1^{k+1}, \cdots, x_n^{k+1}, z^{k+1}, y^{k+1})$
$= h(z^k) + (y^k)^T (\sum_{i=1}^n A_i x_i^{k+1} - z^k) + (\rho/2)\| \sum_{i=1}^n A_i x_i^{k+1} - z^k\|_2^2$
$- h(z^{k+1}) - (y^{k+1})^T (\sum_{i=1}^n A_i x_i^{k+1} - z^{k+1}) - (\rho/2)\| \sum_{i=1}^n A_i x_i^{k+1} - z^{k+1}\|_2^2$
$= h(z^k) - h(z^{k+1}) + (y^k - y^{k+1})^T \sum_{i=1}^n A_i x_i^{k+1} + (y^{k+1})^T z^{k+1} - (y^k)^T z^k + (\rho/2)\| \sum_{i=1}^n A_i x_i^{k+1} - z^k\|_2^2$
$- (\rho/2)\| \sum_{i=1}^n A_i x_i^{k+1} - z^{k+1}\|_2^2$
$= h(z^k) - h(z^{k+1}) + (y^k - y^{k+1})^T \sum_{i=1}^n A_i x_i^{k+1} + (y^{k+1})^T z^{k+1} - (y^k)^T z^k$
$+ (\rho/2)\|z^{k+1} - z^k\|_2^2 + \rho(z^{k+1} - \sum_{i=1}^n A_i x_i^{k+1})^T (z^k - z^{k+1})$
(cosine rule $\|a-b\|^2 - \|a-c\|^2 = \|c-b\|^2 + 2(c-a)^T(b-c)$ where $a = \sum_{i=1}^n A_i x_i^{k+1}$, $b = z^k$ and $c = z^{k+1}$.)
$= h(z^k) - h(z^{k+1}) + (y^k - y^{k+1})^T \sum_{i=1}^n A_i x_i^{k+1} + (y^{k+1})^T z^{k+1} - (y^k)^T z^k + (\rho/2)\|z^{k+1} - z^k\|_2^2 + (y^k - y^{k+1})^T (z^k - z^{k+1})$
(Because $y^{k+1} = y^k + \rho(\sum_{i=1}^n A_i x_i^{k+1} - z^{k+1})$.)
$= h(z^k) - h(z^{k+1}) + (y^k - y^{k+1})^T (\sum_{i=1}^n A_i x_i^{k+1} + z^k - z^{k+1}) + (y^{k+1})^T z^{k+1} - (y^k)^T z^k + (\rho/2)\|z^{k+1} - z^k\|_2^2$
$= h(z^k) - h(z^{k+1}) + (y^k - y^{k+1})^T (\sum_{i=1}^n A_i x_i^{k+1} - z^{k+1}) + (y^k - y^{k+1})^T z^k + (y^{k+1})^T z^{k+1} - (y^k)^T z^k + (\rho/2)\|z^{k+1} - z^k\|_2^2$
$= h(z^k) - h(z^{k+1}) - (1/\rho)(y^k - y^{k+1})^T (y^k - y^{k+1}) - (y^{k+1})^T (z^k - z^{k+1}) + (\rho/2)\|z^{k+1} - z^k\|_2^2$
(Because $y^{k+1} = y^k + \rho(\sum_{i=1}^n A_i x_i^{k+1} - z^{k+1})$.)
$= h(z^k) - h(z^{k+1}) - (y^{k+1})^T (z^k - z^{k+1}) + (\rho/2)\|z^{k+1} - z^k\|_2^2 - (1/\rho)\|y^{k+1} - y^k\|_2^2$
$= h(z^k) - h(z^{k+1}) + \nabla h(z^{k+1})^T (z^{k+1} - z^k) + (\rho/2)\|z^{k+1} - z^k\|_2^2 - (1/\rho)\|y^{k+1} - y^k\|_2^2$ (Lemma 4)
$\geq (-H/2)\|z^{k+1} - z^k\|_2^2 + (\rho/2)\|z^{k+1} - z^k\|_2^2 - (1/\rho)\|\nabla h(z^{k+1}) - \nabla h(z^k)\|_2^2$
($-\nabla h(z)$ is Lipschitz differentiable, Lemma 3 and Lemma 4)
$\geq (-H/2)\|z^{k+1} - z^k\|_2^2 + (\rho/2)\|z^{k+1} - z^k\|_2^2 - (H^2/\rho)\|z^{k+1} - z^k\|_2^2$ (Assumption 1)
$= C_1 \|z^{k+1} - z^k\|_2^2$.

We choose $\rho > 2H$ to make $C_1 > 0$.

Lemma 7. $\forall k \in \mathbb{N}$, *we have* $\|y^{k+1} - y^k\| \leq H\|z^{k+1} - z^k\|$.

Proof.

$$\|y^{k+1} - y^k\| = \|\nabla h(z^{k+1}) - \nabla h(z^k)\| \text{ (Lemma 4) } \leq H\|z^{k+1} - z^k\| \text{ (Assumption 1).}$$

B Proofs of Lemmas 1- 2

Proof (Proof of Lemma 1). This follows directly from Lemmas 5 and 6.

Proof (Proof of Lemma 2). There exists z' such that $\sum_{i=1}^n A_i x_i^k - z' = 0$. Therefore, we have

$$f(x_1^k, \cdots, x_n^k) + h(z') \geq \min S > -\infty.$$

where $S = \{f(x_1, \cdots, x_n) + h(z) : \sum_{i=1}^n A_i x_i - z = 0\}$, which is the objective value of Problem 1, and therefore bounded from below. Then we have

$$\begin{aligned}
&L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k) \\
&= f(x_1^k, \cdots, x_n^k) + h(z^k) + (y^k)^T(\textstyle\sum_{i=1}^n A_i x_i^k - z^k) + (\rho/2)\|\textstyle\sum_{i=1}^n A_i x_i^k - z^k\|^2 \\
&= f(x_1^k, \cdots, x_n^k) + h(z^k) + (y^k)^T(z' - z^k) + (\rho/2)\|\textstyle\sum_{i=1}^n A_i x_i^k - z^k\|^2 \; (\textstyle\sum_{i=1}^n A_i x_i^k - z' = 0) \\
&= f(x_1^k, \cdots, x_n^k) + h(z^k) + (\nabla h(z^k))^T(z' - z^k) + (\rho/2)\|\textstyle\sum_{i=1}^n A_i x_i^k - z^k\|^2 \quad \text{(Lemma 4)} \\
&\geq f(x_1^k, \cdots, x_n^k) + h(z') + (\rho - H)/2\|\textstyle\sum_{i=1}^n A_i x_i^k - z^k\|_2^2 \quad \text{(Lemmas 3 and 4 ,}h(z)\text{ is Lipschitz differentiable)} \\
&\geq \min S + (\rho - H)/2\|\textstyle\sum_{i=1}^n A_i x_i^k - z^k\|_2^2 \geq \min S > -\infty.
\end{aligned}$$

Therefore, $L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k)$ is bounded from below.

C Proofs of Theorems 1-3

Proof (Proof of Theorem 1). We show residual convergence and objective convergence based on Lemmas 1 and 2.
From Lemma 1, $L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k)$ decreases monotonically, and $L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k)$ is lower bounded by Lemma 2. Therefore, $L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k)$ is convergent because a monotone bounded sequence converges (Monotone Convergence Theorem). According to the continuity of L_ρ, we take $k \to \infty$ on the both sides of Inequality (4) to obtain

$$\begin{aligned}
&\lim_{k\to\infty}(L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k) - L_\rho(x_1^{k+1}, \cdots, x_n^{k+1}, z^{k+1}, y^{k+1})) \\
&\geq \lim_{k\to\infty} C_2(\|z^{k+1} - z^k\|_2^2 + \textstyle\sum_{i=1}^n \|A_i(x_i^{k+1} - x_i^k)\|_2^2).
\end{aligned}$$

On one hand, $L_\rho(x_1, \cdots, x_n, z, y)$ is convergent, so we have

$$\lim_{k\to\infty} C_2(\|z^{k+1} - z^k\|_2^2 + \textstyle\sum_{i=1}^n \|A_i(x_i^{k+1} - x_i^k)\|_2^2) \leq 0.$$

On the other hand, $C_2(\|z^{k+1} - z^k\|_2^2 + \sum_{i=1}^n \|A_i(x_i^{k+1} - x_i^k)\|_2^2$ is nonnegative, so we get

$$\lim_{k\to\infty} C_2(\|z^{k+1} - z^k\|_2^2 + \textstyle\sum_{i=1}^n \|A_i(x_i^{k+1} - x_i^k)\|_2^2) = 0.$$

This suggests that $\lim_{k\to\infty}(z^{k+1} - z^k) = 0$ and $\lim_{k\to\infty} A_i(x_i^{k+1} - x_i^k) = 0 (i = 1, \cdots, n)$. Moreover, by Lemma 7, $\lim_{k\to\infty}\|y^{k+1} - y^k\| \leq H \lim_{k\to\infty}\|z^{k+1} - z^k\| = 0$. So we have $\lim_{k\to\infty}(y^{k+1} - y^k) = 0$.

a). For residual convergence, by the Line 8 of Algorithm 1, we have

$$\lim_{k\to\infty} r^k = \lim_{k\to\infty}(y^k - y^{k-1})/\rho = 0.$$

b). For objective convergence, since

$$L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k) = F(x_1^k, \cdots, x_n^k, z^k, y^k) + (y^k)^T r^k + (\rho/2)\|r^k\|_2^2$$

and $L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k)$ is convergent, r^k converges to 0 and y^k is bounded, then $F(x_1^k, \cdots, x_n^k, z^k, y^k)$ is also convergent.

Proof (Proof of Theorem 2). Obviously $\lim_{k\to\infty}(z^{k+1} - z^k) = 0$ and $\lim_{k\to\infty}(y^{k+1} - y^k) = 0$ from the proof of Theorem 1. In order to prove this theorem, we firstly prove that $\lim_{k\to\infty}(x_i^{k+1} - x_i^k) = 0 (i = 1, \cdots, n)$ if either of two assumptions holds, then prove that any limit point $(x_1^*, \cdots, x_n^*, z^*)$ is a feasible Nash point of Problem 1.

(a). Suppose $A_i (i = 1, \cdots, n)$ have full rank. Because $\lim_{k\to\infty} A_i(x_i^{k+1} - x_i^k) = 0$ from the proof of Theorem 1, then obviously $\lim_{k\to\infty}(x_i^{k+1} - x_i^k) = 0$ [23].

(b). Suppose F is strongly convex with regard to x_i. Because $L_\rho(x_1, \cdots, x_n, z, y) = F(x_1, \cdots, x_n, z) + y^T(\sum A_i x_i - z) + (\rho/2)\|\sum A_i x_i - z\|_2^2$, $F(x_1, \cdots, x_n, z)$, and $y^T(\sum A_i x_i - z) + (\rho/2)\|\sum A_i x_i - z\|_2^2$ are strongly convex, L_ρ is also strongly convex regard to x_i [24] with the assumed constant $D_i > 0$. We have

$$\begin{aligned}
L_\rho(x_1^{k+1}, \cdots, x_{i-1}^{k+1}, x_i^k, x_{i+1}^k, \cdots, x_n^k, z^k, y^k) \geq\; & L_\rho(x_1^{k+1}, \cdots, x_{i-1}^{k+1}, x_i^{k+1}, x_{i+1}^k, \cdots, x_n^k, z^k, y^k) + (v_i^{k+1})^T(x_i^k - x_i^{k+1}) \\
&+ (D_i/2)\|x_i^{k+1} - x_i^k\|_2^2
\end{aligned}$$

where $\forall v_i^{k+1} \in \partial_{x_i} L_\rho(x_1^{k+1}, \cdots, x_{i-1}^{k+1}, x_i^{k+1}, x_{i+1}^k, \cdots, x_n^k, z^k, y^k)$. The optimality condition of x_i^{k+1} leads to
$0 \in \partial_{x_i} L_\rho(x_1^{k+1}, \cdots, x_{i-1}^{k+1}, x_i^{k+1}, x_{i+1}^k, \cdots, x_n^k, z^k, y^k)$. Therefore, we have

$$L_\rho(x_1^{k+1}, \cdots, x_{i-1}^{k+1}, x_i^k, x_{i+1}^k, \cdots, x_n^k, z^k, y^k) \geq L_\rho(x_1^{k+1}, \cdots, x_{i-1}^{k+1}, x_i^{k+1}, x_{i+1}^k, \cdots, x_n^k, z^k, y^k) + (D_i/2)\|x_i^{k+1} - x_i^k\|_2^2 \tag{10}$$

We sum up Inequality (10) from $i = 1, \cdots, n$ and Inequality (9) to obtain

$$L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k) - L_\rho(x_1^{k+1}, \cdots, x_n^{k+1}, z^{k+1}, y^{k+1}) \geq \sum\nolimits_{i=1}^n (D_i/2)\|x_i^{k+1} - x_i^k\|_2^2 + C_1\|z^{k+1} - z^k\|_2^2 \tag{11}$$

where $C_1 > 0$ by Lemma 6 if $\rho > 2H$. According to the continuity of L_ρ, we take $k \to \infty$ on the both sides of Inequality (11) to obtain

$$\lim_{k\to\infty}(L_\rho(x_1^k, \cdots, x_n^k, z^k, y^k) - L_\rho(x_1^{k+1}, \cdots, x_n^{k+1}, z^{k+1}, y^{k+1})) \geq \lim_{k\to\infty}(\sum\nolimits_{i=1}^n (D_i/2)\|x_i^{k+1} - x_i^k\|_2^2 + C_1\|z^{k+1} - z^k\|_2^2)$$

On one hand, $L_\rho(x_1, \cdots, x_n, z, y)$ is convergent, so we have

$$\lim_{k\to\infty}(\sum\nolimits_{i=1}^n (D_i/2)\|x_i^{k+1} - x_i^k\|_2^2 + C_1\|z^{k+1} - z^k\|_2^2) \leq 0$$

On the other hand, $\sum_{i=1}^n (D_i/2)\|x_i^{k+1} - x_i^k\|_2^2 + C_1\|z^{k+1} - z^k\|_2^2$ is nonnegative, so we get

$$\lim_{k\to\infty}(\sum\nolimits_{i=1}^n (D_i/2)\|x_i^{k+1} - x_i^k\|_2^2 + C_1\|z^{k+1} - z^k\|_2^2) = 0$$

This suggests that $\lim_{k\to\infty}(x_i^{k+1} - x_i^k) = 0 (i = 1, \cdots, n)$ and $\lim_{k\to\infty}(z^{k+1} - z^k) = 0$.

Therefore, $\lim_{k\to\infty}(x_i^{k+1} - x_i^k) = 0 (i = 1, \cdots, n)$ if either of two assumptions holds. Because $(x_1^k, \cdots, x_n^k, z^k, y^k)$ is bounded, there exists a subsequence $(x_1^s, \cdots, x_n^s, z^s, y^s)$ such that $(x_1^s, \cdots, x_n^s, z^s, y^s) \to (x_1^*, \cdots, x_n^*, z^*, y^*)$ where $(x_1^*, \cdots, x_n^*, z^*, y^*)$ is a limit point. Because $\lim_{s\to\infty}(x_i^{s+1} - x_i^s) = 0 (i = 1, \cdots, n)$, $\lim_{s\to\infty}(z^{s+1} - z^s) = 0$ and $\lim_{s\to\infty}(y^{s+1} - y^s) = 0$, we have $(x_1^{s+1}, \cdots, x_n^{s+1}, z^{s+1}, y^{s+1}) \to (x_1^*, \cdots, x_n^*, z^*, y^*)$. Now we prove that the limit point $(x_1^*, \cdots, x_n^*, z^*)$ is a feasible Nash point of Problem 1.

For feasibility, since $\lim_{k\to\infty} r^k = \lim_{k\to\infty} \sum_{i=1}^n A_i x_i^k - z^k = 0$, so for the subsequence $(x_1^s, \cdots, x_n^s, z^s, y^s) \to (x_1^*, \cdots, x_n^*, z^*, y^*)$, we have $\lim_{s\to\infty} r^s = \lim_{s\to\infty}(\sum_{i=1}^n A_i x_i^s - z^s) = 0$ then $\sum_{i=1}^n A_i x_i^* - z^* = 0$.

For the Nash point, we obtain the following according to the optimality conditions of $x_i^{s+1} (i = 1, \cdots, n)$ and z^{s+1} in Equations (2) and (3), respectively.

$$\begin{aligned}
&L_\rho(x_1^{s+1}, \cdots, x_{i-1}^{s+1}, x_i^{s+1}, x_{i+1}^s, \cdots, x_n^s, z^s, y^s) \leq L_\rho(x_1^{s+1}, \cdots, x_{i-1}^{s+1}, x_i, x_{i+1}^s, \cdots, x_n^s, z^s, y^s), \\
&\forall (x_1^{s+1}, \cdots, x_{i-1}^{s+1}, x_i, x_{i+1}^s, \cdots, x_n^s, z^s) \in dom(F) \\
&L_\rho(x_1^{s+1}, \cdots, x_n^{s+1}, z^{s+1}, y^s) \leq L_\rho(x_1^{s+1}, \cdots, x_n^{s+1}, z, y^s),\ \forall (x_1^{s+1}, \cdots, x_n^{s+1}, z) \in dom(F)
\end{aligned}$$

According to the continuity of L_ρ, we take $s \to \infty$ on the both sides of two inequalities. Because $(x_1^s, \cdots, x_n^s, z^s, y^s) \to (x_1^*, \cdots, x_n^*, z^*, y^*)$ and $(x_1^{s+1}, \cdots, x_n^{s+1}, z^{s+1}, y^{s+1}) \to (x_1^*, \cdots, x_n^*, z^*, y^*)$, we have

$$\begin{aligned}
&L_\rho(x_1^*, \cdots, x_n^*, z^*, y^*) \leq L_\rho(x_1^*, \cdots, x_{i-1}^*, x_i, x_{i+1}^*, \cdots, x_n^*, z^*, y^*),\ \forall (x_1^*, \cdots, x_{i-1}^*, x_i, x_{i+1}^*, \cdots, x_n^*, z^*) \in dom(F) \\
&L_\rho(x_1^*, \cdots, x_n^*, z, y^*) \leq L_\rho(x_1^*, \cdots, x_n^*, z^*, y^*),\ \forall (x_1^*, \cdots, x_n^*, z) \in dom(F)
\end{aligned}$$

Here $\forall (x_1^*, \cdots, x_{i-1}^*, x_i, x_{i+1}^*, \cdots, x_n^*, z^*) \in dom(F)$ and $\forall (x_1^*, \cdots, x_n^*, z) \in dom(F)$ mean $\forall x_i\ s.t.\ \sum_{j=1, j\neq i}^n A_j x_j^* + A_i x_i - z^* = 0$ and $\forall z\ s.t.\ \sum_{j=1}^n A_j x_j^* - z = 0$, respectively. Using the fact that $(x_1^*, \cdots, x_n^*, z^*)$ is feasible in Problem 1, we obtain $L_\rho(x_1^*, \cdots, z^*, y^*) = F(x_1^*, \cdots, z^*)$, $L_\rho(x_1^*, \cdots, x_{i-1}^*, x_i, x_{i+1}^*, \cdots, x_n^*, z^*, y^*) = F(x_1^*, \cdots, x_{i-1}^*, x_i, x_{i+1}^*, \cdots, x_n^*, z^*)$ and $L_\rho(x_1^*, \cdots, x_n^*, z, y^*) = F(x_1^*, \cdots, x_n^*, z)$. Therefore, we prove that $(x_1^*, \cdots, x_n^*, z^*)$ is a feasible Nash point of F defined in Problem 1.

Proof (Proof of Theorem 3). To prove this theorem, we will first show that u_k satisfies two conditions: (1). $u_k \geq u_{k+1}$. (2). $\sum_{k=0}^\infty u_k$ is bounded. We then conclude the convergence rate of $o(1/k)$ based on these two conditions. Specifically, first, we have

$$\begin{aligned}
u_k &= \min\nolimits_{0\leq l\leq k}(\|z^{l+1} - z^l\|_2^2 + \sum\nolimits_{i=1}^n \|A_i(x_i^{l+1} - x_i^l)\|_2^2) \\
&\geq \min\nolimits_{0\leq l\leq k+1}(\|z^{l+1} - z^l\|_2^2 + \sum\nolimits_{i=1}^n \|A_i(x_i^{l+1} - x_i^l)\|_2^2) \\
&= u_{k+1}
\end{aligned}$$

Therefore u_k satisfies the first condition. Second,

$$\begin{aligned}
&\sum\nolimits_{k=0}^{\infty} u_k \\
&= \sum\nolimits_{k=0}^{\infty} \min\nolimits_{0\leq l\leq k}(\|z^{l+1}-z^{l}\|_2^2 + \sum\nolimits_{i=1}^{n} \|A_i(x_i^{l+1}-x_i^{l})\|_2^2) \\
&\leq \sum\nolimits_{k=0}^{\infty} (\|z^{k+1}-z^{k}\|_2^2 + \sum\nolimits_{i=1}^{n} \|A_i(x_i^{k+1}-x_i^{k})\|_2^2) \\
&\leq (L_\rho(x_1^0,\cdots,x_n^0,z^0,y^0) - L_\rho^*)/C_2 \text{ (Lemma 1)}
\end{aligned}$$

where $L_\rho^* = \lim_{k\rightarrow\infty} L_\rho(x_1^k,\cdots,x_n^k,z^k,y^k)$. So $\sum_{k=0}^{\infty} u_k$ is bounded and u_k satisfies the second condition. Finally, it has been proved that the sufficient conditions of convergence rate $o(1/k)$ are: (1) $u_k \geq u_{k+1}$, and (2) $\sum_{k=0}^{\infty} u_k$ is bounded, and (3) $u_k \geq 0$ (Lemma 1.2 in [12]). Since we have proved the first two conditions and the third one $u_k \geq 0$ is obvious, the convergence rate of $o(1/k)$ is proven.

D Algorithms for Applications

D.1 Weakly-constrained Multi-task Learning

Applying the proposed ADMM to solve the problem in Equation (6), we get Algorithm 2. Specifically, Lines 4-9 update primal variables $w_i(i=1,\cdots,n)$ and $z_i(i=1,\cdots,n)$, Line 10 updates the dual variable $y_i(i=1,\cdots,n)$.

Algorithm 2 The Proposed ADMM to Solve Equation (6).

1: Denote $z=[z_1;\cdots;z_n]$,$y=[y_1;\cdots;y_n]$.
2: Initialize $\rho, k=0$.
3: **repeat**
4: Update w_1^{k+1} by Equation (12).
5: **for** i=2 to n-1 **do**
6: Update w_i^{k+1} by Equation (13).
7: **end for**
8: Update w_n^{k+1} by Equation (14).
9: Update z_i^{k+1} by Equation (15) in parallel.
10: $y_i^{k+1} \leftarrow y_i^k + \rho(w_i^{k+1} - z_i^{k+1})$ $(i=1,\cdots,n)$ in parallel.
11: $k \leftarrow k+1$.
12: **until** convergence.
13: Output $w_i(i=1,\cdots,n)$, z.

All subproblems are detailed as follows:

1. Update w^{k+1}

The $w_i^{k+1}(i=1,\cdots,n)$ are updated as follows:

$$w_1^{k+1} \leftarrow \arg\min\nolimits_{w_1} Loss_1(w_1) + \lambda_1 \sum\nolimits_{j=1}^{m} c_1(w_{1,j} w_{2,j}^k) + (\rho/2)\|w_1 - z_1^k + y_1^k/\rho\|_2^2. \tag{12}$$

$$w_i^{k+1} \leftarrow \arg\min\nolimits_{w_i} Loss_i(w_i) + \lambda_1 \sum\nolimits_{j=1}^{m} c_1(w_{i,j} w_{i+1,j}^k) + \lambda_1 \sum\nolimits_{j=1}^{m} c_1(w_{i-1,j}^{k+1} w_{i,j}) + (\rho/2)\|w_i - z_i^k + y_i^k/\rho\|_2^2. \tag{13}$$

$$w_n^{k+1} \leftarrow \arg\min\nolimits_{w_n} Loss_n(w_n) + \lambda_1 \sum\nolimits_{j=1}^{m} c_1(w_{n-1,j}^{k+1} w_{n,j}) + (\rho/2)\|w_n - z_n^k + y_n^k/\rho\|_2^2. \tag{14}$$

They can be solved by the Iterative Soft Thresholding Algorithm (ISTA) [4]. Take w_1^{k+1} as an example, The ISTA leads to

$$w_1^{t+1} \leftarrow \lambda_1 \sum\nolimits_{j=1}^{m} c_1(w_{1,j} w_{2,j}^k) + 1/(2\eta)\|w_1 - (w_1^t - \eta\nabla\phi(w_1^t))\|_2^2.$$

where w_1^t is the t-th iteration in the ISTA, $\eta > 0$ is a learning rate, $\phi(w_1^t) = Loss_1(w_1^t) + \rho/2\|w_1^t - z_1^k + y_1^k/\rho\|_2^2$. For each entry of $w_{1,j}^{t+1}(j=1,2,\cdots,m)$, we have the following closed-form solution as follows:

1). If $w_{1,j}^{t+1} w_{2,j}^{k} \leq 0$, then $w_{1,j}^{t+1} \leftarrow (w_{1,j}^{t} - \eta \nabla_j \phi(w_1^t))/(2\eta\lambda_1 w_{2,j}^2 + 1)$.
2). If $w_{1,j}^{t+1} w_{2,j}^{k} \geq 0$, then $w_{1,j}^{t+1} \leftarrow w_{1,j}^{t} - \eta \nabla_j \phi(w_1^t)$. where $\nabla_j \phi(w_1^t)$ is the j-th entry of $\nabla \phi(w_1^t)$.

2. Update z^{k+1}

The $z_i^{k+1}(i = 1, \cdots, n)$ are updated as follows:

$$z_i^{k+1} \leftarrow \arg\min_{z_i} \varOmega_i(z_i) + (\rho/2)\|w_i^{k+1} - z_i + y_i^k/\rho\|_2^2 (i = 1, \cdots, n). \tag{15}$$

For ℓ_1 or ℓ_2 regularization, they have closed-form solutions.

D.2 Learning with Signed-Network Constraints

Applying proposed ADMM to solve the problem in Equation (7), we get Algorithm 3. Specifically, Lines 4-7 update primal variables $\beta_i(i = 1, \cdots, n)$ and z, Line 8 updates the dual variable y. All subproblems are detailed as follows:

Algorithm 3 The Proposed ADMM to Solve Equation (7).

1: Denote $z = [z_1; \cdots; z_n]$,$y = [y_1; \cdots; y_n]$.
2: Initialize ρ, $k = 0$.
3: **repeat**
4:   **for** i=1 to n **do**
5:     Update β_i^{k+1} by Equation (16).
6:   **end for**
7:   Update $z_i^{k+1}(i = 1, \cdots, n)$ by Equation (17) in parallel.
8:   $y_i^{k+1} \leftarrow y_i^k + \rho(\beta_i^{k+1} - z_i^{k+1})$ $(i = 1, \cdots, n)$ in parallel.
9:   $k \leftarrow k + 1$.
10: **until** convergence.
11: Output $\beta_i(i = 1, \cdots, n), z$.

1. Update β^{k+1}

The $\beta_i^{k+1}(i = 1, \cdots, n)$ are updated as follows:

$$\begin{aligned}\beta_i^{k+1} \leftarrow \arg\min_{\beta_i} Loss(\cdots, \beta_{i-1}^{k+1}, \beta_i, \beta_{i+1}^{k}, \cdots) + \lambda_2 \Big(\sum\nolimits_{(\beta_i, \beta_j^{k+1}) \in E_s, j<i} c_2(\beta_i, \beta_j^{k+1}) + \sum\nolimits_{(\beta_i, \beta_j^{k}) \in E_s, j>i} c_2(\beta_i, \beta_j^{k}) \\ + \sum\nolimits_{(\beta_i, \beta_q^{k+1}) \in E_d, q<i} c_3(\beta_i, \beta_q^{k+1}) + \sum\nolimits_{(\beta_i, \beta_q^{k}) \in E_d, q>i} c_3(\beta_i, \beta_q^{k}) + (\rho/2)\|\beta_i - z_i^k + y_i^k/\rho\|_2^2\Big).\end{aligned} \tag{16}$$

Similar to updating w_i^{k+1} in Algorithm 2, they can be solved efficiently by the ISTA [4].

2. Update z^{k+1}

The $z_i^{k+1}(i = 1, \cdots, n)$ are updated as follows:

$$z_i^{k+1} \leftarrow \arg\min_{z_i} \omega_i(z_i) + (\rho/2)\|\beta_i^{k+1} - z_i + y_i^k/\rho\|_2^2 (i = 1, \cdots, n). \tag{17}$$

Similar to updating z_i^{k+1} in Algorithm 2, they usually have closed-form solutions.