A PROOFS

A.1 Proof of Proposition 1

According to the updating formula of the null cell $x_{ij} \in M$ in Line 4 in Algorithm 1, we have

$$\begin{aligned}
&\Theta\left(C'^{(k+1)}, \mathcal{G}\right) \\
=&\Theta\left(C'^{(k+1)}, \mathcal{G} \mid x_{ij}'^{(k+1)} = x_{ij}'^{(k)} - \eta \frac{\partial\Theta(C'^{(k)}, \mathcal{G})}{\partial x_{ij}'^{(k)}}, \forall x_{ij} \in M\right).
\end{aligned}$$

Given $\left\| -\eta \frac{\partial\Theta(C'^{(k)},\mathcal{G})}{\partial x_{ij}'^{(k)}} \right\| \leq \epsilon$, i.e., $\eta \leq \epsilon / \left\| \frac{\partial\Theta(C'^{(k)},\mathcal{G})}{\partial x_{ij}'^{(k)}} \right\|$, referring to the first-order Taylor expansion [47], it follows

$$\begin{aligned}
&\Theta\left(C'^{(k+1)}, \mathcal{G} \mid x_{ij}'^{(k+1)} = x_{ij}'^{(k)} - \eta \frac{\partial\Theta(C'^{(k)}, \mathcal{G})}{\partial x_{ij}'^{(k)}}, \forall x_{ij} \in M\right) \\
=&\Theta\left(C'^{(k+1)}, \mathcal{G} \mid x_{ij}'^{(k+1)} = x_{ij}'^{(k)}, \forall x_{ij} \in M\right) \\
&+ \sum_{x_{ij} \in M} \left\| -\eta \frac{\partial\Theta(C'^{(k)}, \mathcal{G})}{\partial x_{ij}'^{(k)}} \right\| \cdot \frac{\partial\Theta(C'^{(k)}, \mathcal{G})}{\partial x_{ij}'^{(k)}} \\
=&\Theta\left(C'^{(k+1)}, \mathcal{G} \mid x_{ij}'^{(k+1)} = x_{ij}'^{(k)}, \forall x_{ij} \in M\right) - \eta \sum_{x_{ij} \in M} \left\| \frac{\partial\Theta(C'^{(k)}, \mathcal{G})}{\partial x_{ij}'^{(k)}} \right\|^2 \\
\leq&\Theta\left(C'^{(k+1)}, \mathcal{G} \mid x_{ij}'^{(k+1)} = x_{ij}'^{(k)}, \forall x_{ij} \in M\right) \\
=&\Theta\left(C'^{(k)}, \mathcal{G}\right)
\end{aligned}$$

A.2 Proof of Proposition 2

We first consider the left term $\Theta(\mathbf{C}_1', \mathcal{G} \mid \mathbf{C}_2')$ of Formula 7, which holds

$$\begin{aligned}
\Theta(\mathbf{C}_1', \mathcal{G} \mid \mathbf{C}_2') &= \sum_{g_p \in \mathcal{G}} \sum_{C_{ip}' \in \mathbf{C}_1'} \|x_{ip}' - g_p(C_{ip}' \setminus \{x_{ip}\})\|^2 \\
&= \Theta(\mathbf{C}_1', \mathcal{G}).
\end{aligned}$$

According to Definition 5, we have $\mathbf{C}_1 \cap \mathbf{C}_2 \cap M = \emptyset$, which means that the missing cells in $\mathbf{C}_2$ do not appear in $\mathbf{C}_1$. Therefore, we also have

$$\begin{aligned}
\Theta(\mathbf{C}_1', \mathcal{G} \mid \mathbf{C}_2'') &= \sum_{g_p \in \mathcal{G}} \sum_{C_{ip}' \in \mathbf{C}_1'} \|x_{ip}' - g_p(C_{ip}' \setminus \{x_{ip}\})\|^2 \\
&= \Theta(\mathbf{C}_1', \mathcal{G}).
\end{aligned}$$

It completes the proof.

A.3 Proof of Proposition 3

Considering the initialization of $\mathcal{C}_m = \{\mathbf{C}_1, \dots, \mathbf{C}_u\}$ in Line 2 in Algorithm 2, it follows

$$C' = \mathbf{C}_1' \cup \mathbf{C}_2' \cup \dots \cup \mathbf{C}_u' \cup \{C' \setminus C_m'\}$$

According to Definition 2 for the imputation cost, we can obtain

$$\Theta(C', \mathcal{G}) = \Theta(\mathbf{C}_1', \mathcal{G}) + \Theta(\mathbf{C}_2', \mathcal{G}) + \dots + \Theta(\mathbf{C}_u', \mathcal{G}) + \Theta(C' \setminus C_m', \mathcal{G})$$

Combing with Proposition 2, for any $\mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}_m$, they always have

$$\Theta(\mathbf{C}_i', \mathcal{G} \mid \mathbf{C}_j') = \Theta(\mathbf{C}_i', \mathcal{G} \mid \mathbf{C}_j'') = \Theta(\mathbf{C}_i', \mathcal{G}),$$

where $\mathbf{C}_j'$ and $\mathbf{C}_j''$ are two different fillings of $\mathbf{C}_j$.

Moreover, for any $\mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}_m$ at the k-th round update in Algorithm 2, they always hold

$$\begin{aligned}
&\Theta(C_i'^{(k+1)}, \mathcal{G} \mid C_j'^{(k+1)}) \\
=&\Theta(C_i'^{(k+1)}, \mathcal{G} \mid C_j'^{(k)}) \\
=&\Theta(C_i'^{(k+1)}, \mathcal{G}) \\
=&\Theta\left(\mathbf{C}_i'^{(k+1)}, \mathcal{G} \mid x_{lq}'^{(k+1)} = x_{lq}'^{(k)} - \eta \frac{\partial\Theta(\mathbf{C}_i'^{(k)}, \mathcal{G})}{\partial x_{lq}'^{(k)}}, \forall x_{lq} \in \mathbf{C}_i \cap M\right).
\end{aligned}$$

Therefore, for the k-th round update in Algorithm 2, referring to Line 4 in Algorithm 1, it follows

$$\begin{aligned}
&\Theta(C'^{(k+1)}, \mathcal{G}) \\
=&\Theta\left(C'^{(k+1)}, \mathcal{G} \mid x_{ij}'^{(k+1)} = x_{ij}'^{(k)} - \eta \frac{\partial\Theta(C'^{(k)}, \mathcal{G})}{\partial x_{ij}'^{(k)}}, \forall x_{ij} \in M\right) \\
=&\Theta\left(\mathbf{C}_1'^{(k+1)}, \mathcal{G} \mid x_{ij}'^{(k+1)} = x_{ij}'^{(k)} - \eta \frac{\partial\Theta(\mathbf{C}_1'^{(k)}, \mathcal{G})}{\partial x_{ij}'^{(k)}}, \forall x_{ij} \in \mathbf{C}_1 \cap M\right) \\
&+ \dots \\
&+ \Theta\left(\mathbf{C}_u'^{(k+1)}, \mathcal{G} \mid x_{ij}'^{(k+1)} = x_{ij}'^{(k)} - \eta \frac{\partial\Theta(\mathbf{C}_u'^{(k)}, \mathcal{G})}{\partial x_{ij}'^{(k)}}, \forall x_{ij} \in \mathbf{C}_u \cap M\right) \\
&+ \Theta(C' \setminus C_m', \mathcal{G}).
\end{aligned}$$

That is, PCIMD Algorithm 2 returns the same result C' with CIMD Algorithm 1, for fixed updates.

A.4 Proof of Proposition 4

Given $\left\| -\eta \frac{\partial\Theta(C'^{(k)},\mathcal{G}'^{(k)})}{\partial x_{ij}'^{(k)}} \right\|, \left\| -\eta \frac{\partial\Theta(C'^{(k)},\mathcal{G}'^{(k)})}{\partial \Phi_{\mathcal{G}'}^{(k)}} \right\| \leq \epsilon$, which is equivalent to

$$\eta \leq min \left\{ \epsilon / \left\| \frac{\partial\Theta(C'^{(k)}, \mathcal{G}'^{(k)})}{\partial x_{ij}'^{(k)}} \right\|, \epsilon / \left\| \frac{\partial\Theta(C'^{(k)}, \mathcal{G}'^{(k)})}{\partial \Phi_{\mathcal{G}'}^{(k)}} \right\| \right\},$$

according to the proof of Proposition 1, it leads to

$$\Theta\left(C'^{(k+1)}, \mathcal{G}'^{(k)}\right) \leq \Theta\left(C'^{(k)}, \mathcal{G}'^{(k)}\right).$$

According to Line 6 in Algorithm 3, we have

$$\begin{aligned}
&\Theta\left(C'^{(k+1)}, \mathcal{G}'^{(k+1)}\right) \\
=&\Theta\left(C'^{(k+1)}, \mathcal{G}'^{(k+1)} \mid \Phi_{\mathcal{G}'}^{(k+1)} = \Phi_{\mathcal{G}'}^{(k)} - \eta \frac{\partial\Theta(C'^{(k)}, \mathcal{G}'^{(k)})}{\partial \Phi_{\mathcal{G}'}^{(k)}}\right).
\end{aligned}$$

Moreover, according to the first-order Taylor expansion, it has

$$
\begin{aligned}
&\Theta\left(C'^{(k+1)}, \mathcal{G}'^{(k+1)} \mid \Phi_{\mathcal{G}'}^{(k+1)} = \Phi_{\mathcal{G}'}^{(k)} - \eta\frac{\partial\Theta(C'^{(k)}, \mathcal{G}'^{(k)})}{\partial\Phi_{\mathcal{G}'}^{(k)}}\right) \\
=&\Theta\left(C'^{(k+1)}, \mathcal{G}'^{(k+1)} \mid \Phi_{\mathcal{G}'}^{(k+1)} = \Phi_{\mathcal{G}'}^{(k)}\right) \\
&+\left\|-\eta\frac{\partial\Theta(C'^{(k)}, \mathcal{G}'^{(k)})}{\partial\Phi_{\mathcal{G}'}^{(k)}}\right\| \cdot \frac{\partial\Theta(C'^{(k)}, \mathcal{G}'^{(k)})}{\partial\Phi_{\mathcal{G}'}^{(k)}} \\
=&\Theta\left(C'^{(k+1)}, \mathcal{G}'^{(k+1)} \mid \Phi_{\mathcal{G}'}^{(k+1)} = \Phi_{\mathcal{G}'}^{(k)}\right) - \eta\left\|\frac{\partial\Theta(C'^{(k)}, \mathcal{G}'^{(k)})}{\partial\Phi_{\mathcal{G}'}^{(k)}}\right\|^2 \\
\leq&\Theta\left(C'^{(k+1)}, \mathcal{G}'^{(k+1)} \mid \Phi_{\mathcal{G}'}^{(k+1)} = \Phi_{\mathcal{G}'}^{(k)}\right) \\
=&\Theta\left(C'^{(k+1)}, \mathcal{G}'^{(k)}\right).
\end{aligned}
$$

A.5 Proof of Lemma 5

According to Formula 11, we know

$$
\begin{aligned}
&\mathbb{E}\|\Phi_{\mathcal{G}'}^{(\kappa)} - \Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}\|^2 \\
=&\mathbb{E}_{P_j\sim\mathbf{P}}\left\|\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)} - \eta\sum_{t=0}^{\tau_\kappa-1}\frac{\partial\Theta(C'_j, \mathcal{G}'^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})}} - \Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}\right\|^2.
\end{aligned}
$$

Combining with Assumption 1.1, it further leads to

$$
\begin{aligned}
&\mathbb{E}\|\Phi_{\mathcal{G}'}^{(\kappa)} - \Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}\|^2 \\
=&\eta^2\mathbb{E}_{P_j\sim\mathbf{P}}\left\|\sum_{t=0}^{\tau_\kappa-1}\frac{\partial\Theta(C'_j, \mathcal{G}'^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})}}\right\|^2 \\
\overset{Assumption\ 1.1}{=}&\eta^2\mathbb{E}\left\|\sum_{t=0}^{\tau_\kappa-1}\frac{\partial\Theta(C', \mathcal{G}'^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})}}\right\|^2 \\
\leqslant&\tau_k\eta^2\sum_{t=0}^{\tau_\kappa-1}\mathbb{E}\left\|\frac{\partial\Theta(C', \mathcal{G}'^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})}}\right\|^2.
\end{aligned}
$$

Moreover, referring to Assumption 1.3 and Assumption 1.4, we can derive the conclusion that

$$
\begin{aligned}
&\mathbb{E}\|\Phi_{\mathcal{G}'}^{(\kappa)} - \Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}\|^2 \\
\overset{Assumption\ 1.3}{\leqslant}&T\eta^2\sum_{t=0}^{\tau_\kappa-1}\mathbb{E}\left\|\frac{\partial\Theta(C', \mathcal{G}'^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa+t-\tau_{\kappa-\tau_\kappa+t})}}\right\|^2 \\
\overset{Assumption\ 1.4}{\leqslant}&T\eta^2\sum_{t=0}^{\tau_\kappa-1}V^2 \\
\overset{Assumption\ 1.3}{\leqslant}&T^2\eta^2V^2.
\end{aligned}
$$

A.6 Proof of Proposition 6

We start from Formula 11, combining with Line 9 in Algorithm 4 and Proposition 1, it has

$$
\begin{aligned}
&\mathbb{E}_{P_j\sim\mathbf{P}}\Theta\left(C'_j, \mathcal{G}'^{(\kappa+1)}\right) \\
=&\mathbb{E}_{P_j\sim\mathbf{P}}\Theta\left(C'_j, \mathcal{G}'^{(\kappa+1)} \mid \Phi_{\mathcal{G}'}^{(\kappa+1)} = \Phi_{\mathcal{G}'}^{(\kappa)} - \eta\frac{\partial\Theta(C'_j, \mathcal{G}'^{(\kappa-\tau_\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}}\right).
\end{aligned}
$$

Then, accordinig to Assumption 1.2, we have

$$
\begin{aligned}
&\mathbb{E}_{P_j\sim\mathbf{P}}\Theta\left(C'_j, \mathcal{G}'^{(\kappa+1)}\right) \\
\overset{Assumption 1.2}{\leqslant}&\mathbb{E}_{P_j\sim\mathbf{P}}\Theta\left(C'_j, \mathcal{G}'^{(\kappa)}\right) \\
&-\eta\mathbb{E}_{P_j\sim\mathbf{P}}\left\langle\frac{\partial\Theta(C'_j, \mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}, \frac{\partial\Theta(C'_j, \mathcal{G}'^{(\kappa-\tau_\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}}\right\rangle \\
&+\frac{L\eta^2}{2}\mathbb{E}_{P_j\sim\mathbf{P}}\left\|\frac{\partial\Theta(C'_j, \mathcal{G}'^{(\kappa-\tau_\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}}\right\|^2.
\end{aligned}
$$

Combining with Assumption 1.1, we can obtain

$$
\begin{aligned}
&\mathbb{E}\Theta\left(C', \mathcal{G}'^{(\kappa+1)}\right) \\
\overset{Assumption 1.1}{\leqslant}&\mathbb{E}\Theta\left(C', \mathcal{G}'^{(\kappa)}\right) \\
&-\eta\mathbb{E}\left\langle\frac{\partial\Theta(C', \mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}, \frac{\partial\Theta(C', \mathcal{G}'^{(\kappa-\tau_\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}}\right\rangle \\
&+\frac{L\eta^2}{2}\mathbb{E}\left\|\frac{\partial\Theta(C', \mathcal{G}'^{(\kappa-\tau_\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}}\right\|^2 \\
=&\mathbb{E}\Theta\left(C', \mathcal{G}'^{(\kappa)}\right) \\
&+\frac{\eta}{2}\mathbb{E}\left\|\frac{\partial\Theta(C', \mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}} - \frac{\partial\Theta(C', \mathcal{G}'^{(\kappa-\tau_\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}}\right\|^2 \\
&-\frac{\eta}{2}\mathbb{E}\left\|\frac{\partial\Theta(C', \mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}\right\|^2 - \frac{\eta}{2}\mathbb{E}\left\|\frac{\partial\Theta(C', \mathcal{G}'^{(\kappa-\tau_\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}}\right\|^2 \\
&+\frac{L\eta^2}{2}\mathbb{E}\left\|\frac{\partial\Theta(C', \mathcal{G}'^{(\kappa-\tau_\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}}\right\|^2.
\end{aligned}
$$

Moreover, according Assumption 1.2 and Assumption 1.4, it has

$$
\begin{aligned}
&\mathbb{E}\Theta\left(C',\mathcal{G}'^{(\kappa+1)}\right)\\
\overset{Assumption 1.2}{\leqslant}&\mathbb{E}\Theta\left(C',\mathcal{G}'^{(\kappa)}\right)+\frac{\eta L^2}{2}\mathbb{E}\|\Phi_{\mathcal{G}'}^{(\kappa)}-\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}\|^2\\
&-\frac{\eta}{2}\mathbb{E}\left\|\frac{\partial\Theta(C',\mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}\right\|^2+\frac{L\eta^2}{2}\mathbb{E}\left\|\frac{\partial\Theta(C',\mathcal{G}'^{(\kappa-\tau_\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}}\right\|^2\\
\overset{Assumption 1.4}{\leqslant}&\mathbb{E}\Theta\left(C',\mathcal{G}'^{(\kappa)}\right)+\frac{\eta L^2}{2}\mathbb{E}\|\Phi_{\mathcal{G}'}^{(\kappa)}-\Phi_{\mathcal{G}'}^{(\kappa-\tau_\kappa)}\|^2\\
&-\frac{\eta}{2}\mathbb{E}\left\|\frac{\partial\Theta(C',\mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}\right\|^2+\frac{L\eta^2V^2}{2}.
\end{aligned}
$$

Referring to Lemma 5, it follows

$$
\begin{aligned}
&\mathbb{E}\Theta\left(C',\mathcal{G}'^{(\kappa+1)}\right)\\
\overset{Lemma 5}{\leqslant}&\mathbb{E}\Theta\left(C',\mathcal{G}'^{(\kappa)}\right)-\frac{\eta}{2}\mathbb{E}\left\|\frac{\partial\Theta(C',\mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}\right\|^2\\
&+\frac{\eta^3L^2T^2V^2}{2}+\frac{L\eta^2V^2}{2}.
\end{aligned}
$$

Summing from $\kappa=0$ to $\kappa=K-1$, we have

$$
\begin{aligned}
&\mathbb{E}\Theta\left(C',\mathcal{G}'^{(K)}\right)\\
\leqslant&\mathbb{E}\Theta\left(C',\mathcal{G}'^{(0)}\right)-\frac{\eta}{2}\sum_{\kappa=0}^{K-1}\mathbb{E}\left\|\frac{\partial\Theta(C',\mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}\right\|^2\\
&+\frac{\eta^3L^2T^2V^2K}{2}+\frac{L\eta^2V^2K}{2}.
\end{aligned}
$$

Considering Definition 2 for the imputation cost, it leads to

$$\mathbb{E}\Theta\left(C',\mathcal{G}'^{(K)}\right)\geqslant 0.$$

Finally, we can obtain the conclusion

$$
\begin{aligned}
\sum_{\kappa=0}^{K-1}\mathbb{E}\left\|\frac{\partial\Theta(C',\mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}\right\|^2\leqslant&\frac{2\Theta\left(C',\mathcal{G}'^{(0)}\right)}{\eta}\\
&+\eta LKV^2(T^2L\eta+1).
\end{aligned}
$$

A.7 Proof of Proposition 7

From Proposition 6, we have

$$\sum_{\kappa=0}^{K-1}\mathbb{E}\left\|\frac{\partial\Theta(C',\mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}\right\|^2\leqslant\frac{\alpha}{\eta}+\eta K\beta.$$

When multiplying the term $\frac{1}{K}$ on both sides of the inequation, we can obtain

$$\frac{1}{K}\sum_{\kappa=0}^{K-1}\mathbb{E}\left\|\frac{\partial\Theta(C',\mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}}^{(\kappa)}}\right\|^2\leqslant\frac{\alpha}{\eta K}+\eta\beta.$$

In the end, if we further take $\eta=\sqrt{\frac{\alpha}{K\beta}}$ for it, which leads to the conclusion

$$\frac{1}{K}\sum_{\kappa=0}^{K-1}\mathbb{E}\left\|\frac{\partial\Theta(C',\mathcal{G}'^{(\kappa)})}{\partial\Phi_{\mathcal{G}'}^{(\kappa)}}\right\|^2\leqslant\sqrt{\frac{\alpha\beta}{K}}+\sqrt{\frac{\alpha\beta}{K}}=2\sqrt{\frac{\alpha\beta}{K}}.$$