

A Overview of Supplemental Material

We organize supplementary material as follows. In Section B, we present our results compared to baselines on the *full-to-full* matching task between two complete shapes. In Section C, we show a small set of qualitative results for animal shapes. In Section D, we explain the implementation details of reweighting module, which is an important component of our transformation regularization module. In Section E, we detail the back-propagation rule for the regression submodule that estimates local transformations, as described in Equation (2). Finally, in Section F we prove Theorem 1 on the exact recovery conditions for our transformation synchronization formulation.

B Full-to-full Matching Results

In this section, we compare our method with the four baselines, namely, DeepGeoFunc [11], DHBC [51], 3D-CODED [13], and three variants of LES [10]: point translation and patch deformation 3D (LES-PTD3), point translation 10D (LES-PT10), patch deformation 10D (LES-PD10). All correspondences are refined by designated algorithms: 3D-CODED and LES use deformation during prediction, DeepGeoFunc uses ZoomOut [35], and DHBC and our method uses non-rigid ICP for correspondence refinement. All statistics are reported in Table 2 and qualitative results are provided in Figure 4.

Overall, our method achieved competitive results when compared against these state-of-the-art baselines. Specifically, we achieved 1.74cm/1.72cm/1.42cm/3.50cm average correspondence error on the SURREAL/FAUST-inter/FAUST-intra/SHREC19 datasets, where the top-performing baseline achieved 2.54cm/1.64cm/1.45cm/5.02cm. We achieved the most improvement on SHREC19, which is recognized as the most difficult dataset due to large domain gap. This result demonstrates the power of our regularization module.

It should be noted that all the baseline approaches involve isometry-invariance constraints (e.g., spectral techniques and those preserve geodesic distances) possess strong regularizations in the full-2-full setting. Still, our approach, which leverages the implicit articulated deformation structure, outperform them by a considerable margin. This shows the advantages of modeling approximated piece-wise rigidity constraints for establishing dense correspondences between 3D models of humans.

Method	SURREAL			FAUST-inter			FAUST-intra			SHREC19		
	AE(cm)	5cm-recall	10cm-recall	AE(cm)	5cm-recall	10cm-recall	AE(cm)	5cm-recall	10cm-recall	AE(cm)	5cm-recall	10cm-recall
DHBC	/	/	/	2.35	0.900	0.972	2.00	0.911	0.975	/	/	/
3D-CODED	2.89	0.958	0.971	2.08	0.956	0.983	1.97	0.964	0.986	6.69	0.740	0.898
LES-PTD3	2.74	0.962	0.971	1.64	0.969	0.992	1.49	0.976	0.992	6.75	0.742	0.894
LES-PT10	2.54	0.962	0.971	1.69	0.971	0.991	1.45	0.976	0.992	6.95	0.733	0.895
LES-PD10	2.75	0.962	0.972	2.05	0.949	0.978	2.32	0.949	0.971	5.02	0.766	0.948
DeepGeoFunc	3.45	0.905	0.971	2.92	0.828	0.972	2.11	0.886	0.975	17.02	0.440	0.727
Ours-Refine-ICP	1.74	0.980	0.990	1.72	0.962	0.990	1.42	0.969	0.992	3.50	0.824	0.968

Table 2: Evaluation of all correspondence computation methods for the full-to-full correspondence task. We report the average correspondence error, 5cm-recall, and 10cm-recall for all correspondence computation methods on each dataset. The unit of error is cm. All reported correspondences results are evaluated after designated refinement methods have been applied. All the results are reported the same set of datasets reported in this paper by running the code from the authors of the baseline approaches.

C Qualitative Results on Animal Shapes

As an additional experiment, we train our model on a synthetic dataset utilizing SMAL [61]. We generate 20K triangular meshes using latent parameters generated from a gaussian distribution. For each mesh, we rendered 100 depth scans from random viewpoints. The model was trained in a similar fashion as the human shape and evaluated on the horse class of TOSCA dataset [7] (see Figure 5). The results indicates the usefulness of our transformation synchronization module. For more results, please refer to our code base.

D Implementation of Reweighting Module

In this section, we demonstrate the implementation details of reweighting module.

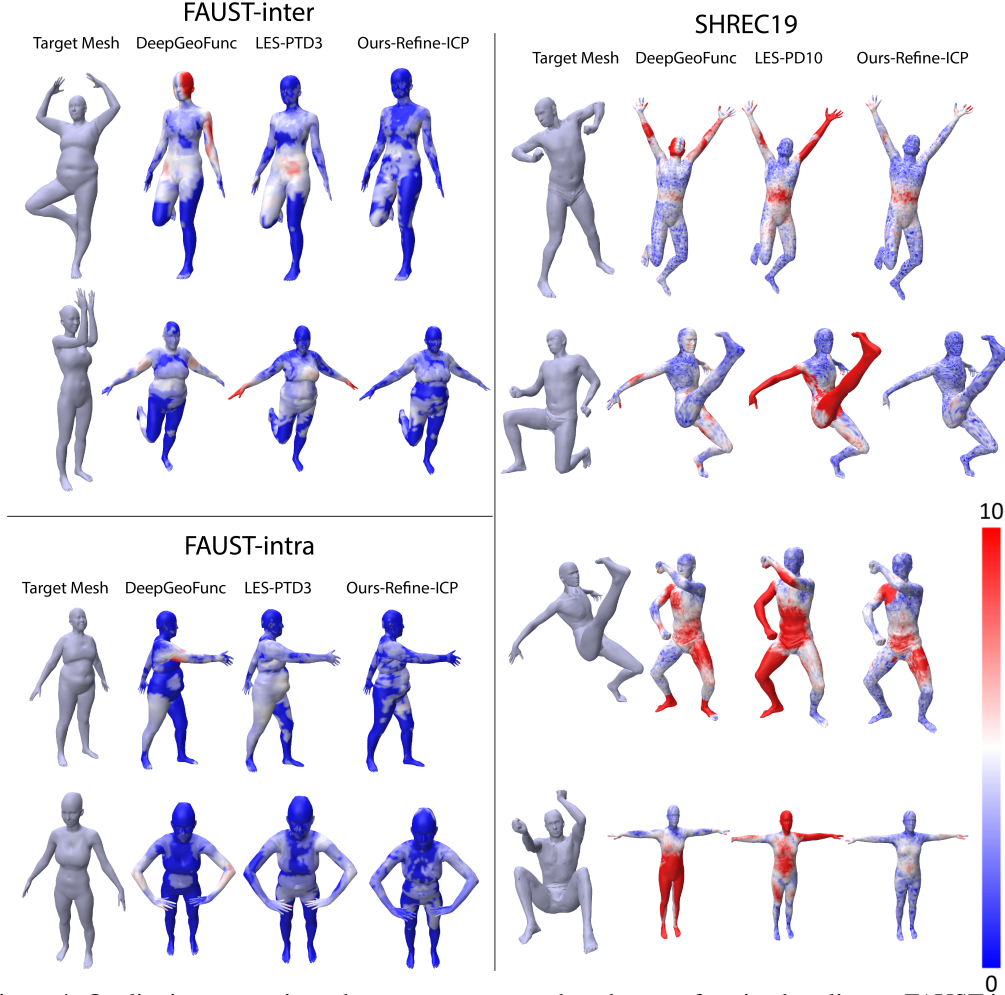


Figure 4: Qualitative comparisons between our approach and top performing baseline on FAUST-inter, FAUST-intra and SHREC19. For each dataset, we report from left to right: target mesh, error maps of two top-performing baselines and our method after refinement. For each method, the error (cm) is visualized on the source mesh.

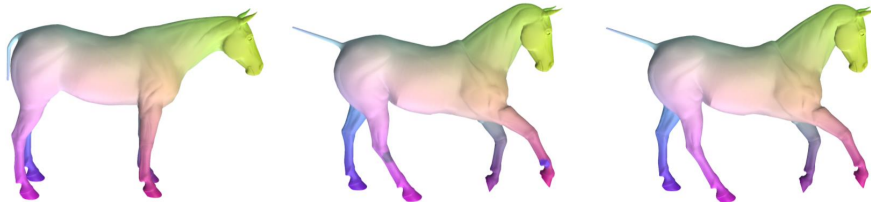


Figure 5: Qualitative Results evaluated on the horse class of TOSCA dataset [7]. From left to right: 1) Ground Truth Correspondence on source shape. 2) Initial Correspondence before transformation regularization. 3) Correspondence after transformation regularization.

The input to this module is a point cloud with vertex-wise correspondences and transformations. The goal is to update vertex weights and edge weights to prepare for transformation propagation by (6). The design of reweighting module is illustrated in figure 6. Specifically, we propagate messages $m_i^{(k)}$ among vertices and aggregate them by element-wise maximum $MAX(\cdot)$ as in the message passing framework [12]. This is repeated by k_{\max} iterations, where in the last layer $m_i^{(k_{\max})}$ is a scalar indicating the vertex-wise weights.

Input Point Cloud: \mathbf{x}_i
Initial Correspondence on template: \mathbf{y}_i
Initial Transformations: R_i, \mathbf{t}_i
Initial Embedding: $m_i^{(0)} = \text{MLP}(\mathbf{x}_i)$

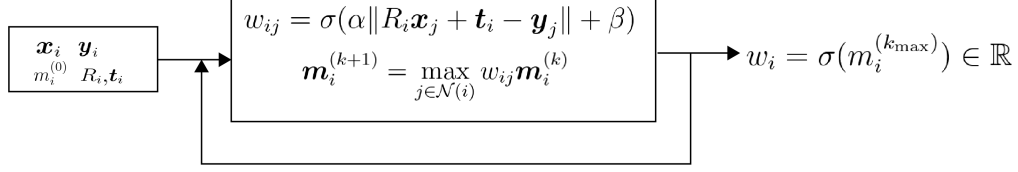


Figure 6: Illustration of Reweighting Module. Note that $\mathbf{y}_i \in \mathbb{R}^3$ are point locations on template mesh.

E Back-Propagation Rule for (2)

We consider a general objective function for pose regression:

$$R_i^{in}, \mathbf{t}_i^{in} = \underset{R, \mathbf{t}}{\operatorname{argmin}} \sum_{j \in \mathcal{N}(i)} w_j \|R \mathbf{p}_j + \mathbf{t} - \mathbf{q}_{c_j}\|^2 \quad (12)$$

where $w_j > 0$. Our goal is derive the partial derivatives of R_i^{in} and \mathbf{t}_i^{in} with respect to w_j .

The following proposition, which is due to [18], characterizes a closed-form expression for R_i^{in} and \mathbf{t}_i^{in} .

Proposition 1. *Let*

$$\mathbf{c}_p = \sum_{j \in \mathcal{N}(i)} w_j \mathbf{p}_j / \sum_{j \in \mathcal{N}(i)} w_j, \quad \mathbf{c}_q = \sum_{j \in \mathcal{N}(i)} w_j \mathbf{q}_{c_j} / \sum_{j \in \mathcal{N}(i)} w_j.$$

Denote

$$S = \sum_{j \in \mathcal{N}(i)} w_j (\mathbf{p}_j - \mathbf{c}_p)(\mathbf{q}_{c_j} - \mathbf{c}_q)^T.$$

Let $S = U \Sigma V^T$ be the singular-value decomposition of S . Then

$$R_i^{in} = V \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \operatorname{sign}(\det(S)) \end{pmatrix} U^T, \quad \mathbf{t}_i^{in} = \mathbf{c}_q - R \mathbf{c}_p. \quad (13)$$

The following proposition characterizes the derivatives of R_i^{in} with respect to elements of S , which then determines the derivatives with respect to w_j .

Lemma 1. *The derivatives of R with respect to S is given by*

$$\begin{aligned} dR = & \sum_{1 \leq i \neq j \leq 3} \mathbf{v}_i \cdot \mathbf{u}_j^T \cdot \frac{\mathbf{u}_j^T \cdot dS \cdot \mathbf{v}_i - \mathbf{u}_i^T \cdot dS \cdot \mathbf{v}_j}{\sigma_i + \sigma_j} \\ & + (\operatorname{sign}(\det(S)) - 1) \sum_{j=1}^2 \left(\mathbf{u}_j^T dS \mathbf{v}_3 \cdot \frac{\sigma_3 \mathbf{v}_3 \mathbf{u}_j^T + \sigma_j \mathbf{v}_j \mathbf{u}_3^T}{\sigma_3^2 - \sigma_j^2} + \mathbf{u}_3^T dS \mathbf{v}_j \cdot \frac{\sigma_3 \mathbf{v}_j \mathbf{u}_3^T + \sigma_j \mathbf{v}_3 \mathbf{u}_j^T}{\sigma_3^2 - \sigma_j^2} \right) \end{aligned} \quad (14)$$

where $U = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$, $V = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, and $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \sigma_3)$.

Proof. It is clear that

$$S \mathbf{v}_i = \sigma_i \mathbf{u}_i, \mathbf{u}_i^T S = \sigma_i \mathbf{v}_i^T, \quad 1 \leq i \leq 3.$$

Taking the differential on both sides, we obtain

$$dS \cdot \mathbf{v}_i + S \cdot d\mathbf{v}_i = d\sigma_i \cdot \mathbf{u}_i + \sigma_i \cdot d\mathbf{u}_i \quad (15)$$

$$d\mathbf{u}_i^T S + \mathbf{u}_i^T dS = d\sigma_i \mathbf{v}_i^T + \sigma_i d\mathbf{v}_i^T \quad (16)$$

Left multiplying both sides of (15) by \mathbf{u}_j with $j \neq i$ and observing that $\mathbf{u}_j^T \mathbf{u}_i = 0$, we obtain

$$\mathbf{u}_j^T dS\mathbf{v}_i + \mathbf{u}_j^T S d\mathbf{v}_i = \sigma_i \mathbf{u}_j^T d\mathbf{u}_i. \quad (17)$$

Similarly right multiplying both sides of (16) by \mathbf{v}_j^T with $j \neq i$ gives

$$d\mathbf{u}_i^T S\mathbf{v}_j + \mathbf{u}_i^T dS\mathbf{v}_j = \sigma_i d\mathbf{v}_i^T \mathbf{v}_j. \quad (18)$$

It follows that

$$\mathbf{u}_j^T dS\mathbf{v}_i + \sigma_j \mathbf{v}_j^T d\mathbf{v}_i = \sigma_i \mathbf{u}_j^T d\mathbf{u}_i \quad (19)$$

$$\sigma_j \cdot d\mathbf{u}_i^T \mathbf{u}_j + \mathbf{u}_i^T dS\mathbf{v}_j = \sigma_i d\mathbf{v}_i^T \mathbf{v}_j \quad (20)$$

Observe that $\mathbf{u}_j^T d\mathbf{u}_i + d\mathbf{u}_i^T \mathbf{u}_j = 0$ and $\mathbf{v}_j^T d\mathbf{v}_i + d\mathbf{v}_i^T \mathbf{v}_j = 0$. Combining (19) and (20) to solve for $\mathbf{u}_j^T d\mathbf{u}_i$ and $\mathbf{v}_j^T d\mathbf{v}_i$, we obtain

$$\mathbf{u}_j^T d\mathbf{u}_i = \frac{\sigma_i \mathbf{u}_j^T dS\mathbf{v}_i + \sigma_j \mathbf{u}_i^T dS\mathbf{v}_j}{\sigma_i^2 - \sigma_j^2} \quad (21)$$

$$\mathbf{v}_k^T d\mathbf{v}_i = \frac{\sigma_i \mathbf{u}_i^T dS\mathbf{v}_j + \sigma_j \mathbf{u}_j^T dS\mathbf{v}_i}{\sigma_i^2 - \sigma_j^2} \quad (22)$$

Since $\mathbf{u}_i d\mathbf{u}_i = 0$, we have

$$d\mathbf{u}_i = \sum_{j \neq i} \frac{\sigma_i \mathbf{u}_j^T dS\mathbf{v}_i + \sigma_j \mathbf{u}_i^T dS\mathbf{v}_j}{\sigma_i^2 - \sigma_j^2} \mathbf{u}_j$$

$$d\mathbf{v}_i = \sum_{j \neq i} \frac{\sigma_i \mathbf{u}_i^T dS\mathbf{v}_j + \sigma_j \mathbf{u}_j^T dS\mathbf{v}_i}{\sigma_i^2 - \sigma_j^2} \mathbf{v}_j$$

In the case $\det(S) > 0$, we have

$$dR = \sum (\mathbf{v}_i d\mathbf{u}_i^T + d\mathbf{v}_i \mathbf{u}_i^T) \\ \sum_{1 \leq i \neq j \leq 3} \mathbf{v}_i \cdot \mathbf{u}_j^T \cdot \frac{\mathbf{u}_j^T \cdot dS \cdot \mathbf{v}_i - \mathbf{u}_i^T \cdot dS \cdot \mathbf{v}_j}{\sigma_i + \sigma_j}.$$

The proof for the case $\det(S) < 0$ is similar and it omitted for brevity. \square

For the network training, at each iteration and for each instance, we fix the nearest neighbors and apply the expression described above to back-propagation the gradients from the initial transformations to the descriptor tower.

F Proof of Robust Recovery Condition

This section presents the proof of Theorem. We begin with a few useful lemmas in Section F.1. We then complete the proof in Section F.2.

F.1 Useful Lemmas

The first lemma concerns the relation between different norms of a symmetric matrix A .

Lemma 2. *Given a symmetric matrix A , we have*

$$\|A\| \leq \|A\|_1. \quad (23)$$

Proof. Consider any eigenvalue λ of A . Denote $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ as its corresponding eigenvector. Let

$$i^* = \operatorname{argmax}_{1 \leq i \leq n} |x_i|.$$

Then

$$\sum_{j=1}^n a_{ij}x_j = \lambda x_{i^*}.$$

It follows that

$$|\lambda| |x_{i^*}| \leq \left| \sum_{j=1}^n a_{ij}x_j \right| \leq \sum_{j=1}^n |a_{ij}| |x_j| \leq \sum_{j=1}^n |a_{ij}| |x_{i^*}| \leq \|A\|_1 \cdot |x_{i^*}|$$

It follows that

$$|\lambda| \leq \|A\|_1,$$

which ends the proof. \square

The second lemma concerns the L_1 norm of $(A - B)^{-1}$.

Lemma 3. *Suppose $\|A^{-\frac{1}{2}}\|_1^2 \|B\|_1 < 1$. Then*

$$\|(A - B)^{-1}\|_1 \leq \frac{\|A^{-\frac{1}{2}}\|_1^2}{1 - \|A^{-\frac{1}{2}}\|_1^2 \|B\|_1}. \quad (24)$$

Proof. First of all, since $\|A^{-\frac{1}{2}}\|_1^2 \|B\|_1 < 1$. It follows that

$$\|A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\| < 1.$$

In other words, we have

$$\begin{aligned} (A - B)^{-1} &= A^{-\frac{1}{2}}(I - A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{-\frac{1}{2}} \\ &= A^{-\frac{1}{2}} \cdot \sum_{i=0}^{\infty} (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^i \cdot A^{-\frac{1}{2}}. \end{aligned}$$

Therefore

$$\begin{aligned} \|(A - B)^{-1}\|_1 &\leq \|A^{-\frac{1}{2}}\|_1^2 \cdot \sum_{i=0}^{\infty} \|(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^i\|_1 \\ &\leq \|A^{-\frac{1}{2}}\|_1^2 \cdot \sum_{i=0}^{\infty} (\|A^{-\frac{1}{2}}\|_1^2 \|B\|_1)^i \\ &\leq \frac{\|A^{-\frac{1}{2}}\|_1^2}{1 - \|A^{-\frac{1}{2}}\|_1^2 \|B\|_1}. \end{aligned}$$

\square

E.2 Completing the Proof

Note that our reweighted least square formulation solves the following linear system at each iteration:

$$\min_{\mathbf{v}_i, 1 \leq i \leq n} \sum_{i=1}^n w_i^{(t)} \|\mathbf{v}_i - \mathbf{v}_i^{gt}\|^2 + \lambda \sum_{(i,j) \in \mathcal{E}} w_{ij}^{(t)} \|\mathbf{v}_i - \mathbf{v}_j\|^2 \quad (25)$$

where $w_i^{(t)}$, $1 \leq i \leq n$ and $w_{ij}^{(t)}$, $(i, j) \in \mathcal{E}$ are weights associated with vertices and edges at iteration t . Define

$$W_V^{(t)} := \text{Diag}(w_i^{(t)} |_{1 \leq i \leq n}) \in \mathbb{R}^{n \times n}, \quad W_{\mathcal{E}}^{(t)} := \text{Diag}(w_{ij}^{(t)} |_{(i,j) \in \mathcal{E}}) \in \mathbb{R}^{|\mathcal{E}| \times |\mathcal{E}|}.$$

Proposition 2. *Suppose $\mathbf{v}_i = \mathbf{v}_i^{gt} + \delta_i$, $1 \leq i \leq n$ and $\delta_{ij} = \mathbf{v}_i^{gt} - \mathbf{v}_j^{gt}$, $(i, j) \in \mathcal{E}$. Then the solution $\mathbf{x}^{(t)}$ to (25) satisfies*

$$\mathbf{v}^{(t)} - \mathbf{v}^{gt} := L^{(t)-1} \mathbf{r}^{(t)}.$$

where

$$L^{(t)} := W_{\mathcal{V}}^{(t)} + \lambda J^T W_{\mathcal{E}}^{(t)} J \quad (26)$$

$$\mathbf{r}^{(t)} := W_{\mathcal{V}}^{(t)} \text{mat}(\delta_i) + \lambda J^T W_{\mathcal{E}}^{(t)} \text{mat}(\delta_{ij}). \quad (27)$$

where $J \in \mathbb{R}^{|\mathcal{E}| \times n}$ is the vertex and edge indicator matrix; $\text{mat}(\delta_i) \in \mathbb{R}^{n \times 12}$ collects the vertex indicators; $\text{mat}(\delta_{ij}) \in \mathbb{R}^{|\mathcal{E}| \times 12}$ collects the edge indicators.

Proof. Let $V = (\mathbf{v}_1, \dots, \mathbf{v}_n)^T \in \mathbb{R}^{n \times 12}$. It is clear that (25) is identical to

$$\min_V \text{Trace}(V^T L^{(t)} V) - 2\text{Trace}(V^T \mathbf{r}^{(t)})$$

which ends the proof. \square

Let $\mathcal{V}^{(t)} \subset [n]$ and $\mathcal{E}^{(t)} \subset \mathcal{E}$ be the remaining edges at iteration t . By the default setting, we assume $\mathcal{V}^{(t)}$ and $\mathcal{E}^{(t)}$ satisfy the recursion properties:

$$\begin{aligned} \mathcal{V}^g &\subset \mathcal{V}^{(t)} \subset [n] \\ \mathcal{E}^g &\subset \mathcal{E}^{(t)} \subset \mathcal{E} \end{aligned}$$

Moreover,

$$\begin{aligned} \|\delta_i\| &\leq c_0 c^t, \quad \forall 1 \leq i \leq n \\ \|\delta_{ij}\| &\leq c_0 c^t, \quad \forall (i, j) \in \mathcal{E} \end{aligned}$$

It follows that

$$\|\mathbf{r}^{(t)}\|_{\infty} \leq \max \left(\epsilon_0 \|D_W^g\|_1 + c_0 c^t \|L_W^{g, \text{ol}}\|_1, c_0 c^t (\|D_W^b\|_1 + \|L_W^{b, \text{ol}}\|_1) \right) \quad (28)$$

Note that the solution at the next iteration satisfies

$$\|V^{(t+1)} - V^{gt}\|_{\infty} \leq \|L^{(t)}\|_1^{-1} \|\mathbf{r}^{(t)}\|_{\infty}. \quad (29)$$

Applying Lemma 3, we have

$$\begin{aligned} \|L^{(t)}\|_1^{-1} &\leq \frac{\|L^{-\frac{1}{2}}\|_1^2}{1 - \|L^{-\frac{1}{2}}\|_1^2 \|L_{\text{ol}}\|_1} \\ &\leq \frac{\|L^{-\frac{1}{2}}\|_1^2}{1 - \|L^{-\frac{1}{2}}\|_1^2 \max \left(\|L_W^{g, \text{ol}}\|_1, \|D_W^b\|_1 + \|L_W^{b, \text{ol}}\|_1 \right)}. \end{aligned} \quad (30)$$

where $L = D_W + L_W$, and L_{ol} collects all the elements of L that are attached to outlier observations.

Combing (28)-(30), we have

$$\|V^{(t+1)} - V^{gt}\|_{\infty} \leq \frac{\|L^{-\frac{1}{2}}\|_1^2 \cdot \max \left(\epsilon_0 \|D_W^g\|_1 + c_0 c^t \|L_W^{g, \text{ol}}\|_1, c_0 c^t (\|D_W^b\|_1 + \|L_W^{b, \text{ol}}\|_1) \right)}{1 - \|L^{-\frac{1}{2}}\|_1^2 \max \left(\|L_W^{g, \text{ol}}\|_1, \|D_W^b\|_1 + \|L_W^{b, \text{ol}}\|_1 \right)} \quad (31)$$

Therefore, under the assumptions of the theorem, i.e.,

$$\|(D_W + L_W)^{-\frac{1}{2}}\|_1^2 \cdot \max \left(\frac{1}{c_1} \|D_W^g\|_1 + 2\|L_W^{g, \text{ol}}\|_1, \|D_W^b\|_1 + 2\|L_W^{b, \text{ol}}\|_1 \right) \leq \frac{c}{2+c}$$

we have

$$\|\mathbf{r}^{(t)}\|_{\infty} \leq \frac{1}{2} \max(c_1 \epsilon_0, c_0 c^{t+1})$$

This means the iterative procedure only prunes the outliers, and now we end the proof.