Appendix

A Reconstruction Weight Calculation

By multiplying K, the main components of Eq.(3) can be reformulated as:

$$
\begin{aligned}
&\|K\mathbf{g}_i - \sum_j \big(\mathbf{g}_j + (w_{ij} + m_{ij})(\mathbf{g}_i - \mathbf{g}_j)\big)\|_2^2 \\
=&\|K\mathbf{g}_i - \sum_j \mathbf{g}_j - \sum_j w_{ij}(\mathbf{g}_i - \mathbf{g}_j) - \sum_j m_{ij}(\mathbf{g}_i - \mathbf{g}_j)\|_2^2 \\
=&\|(K-1)\mathbf{g}_i - \sum_j \mathbf{g}_j - \sum_j m_{ij}(\mathbf{g}_i - \mathbf{g}_j) + \sum_j w_{ij}\mathbf{g}_j\|_2^2 \\
=&\|\mathbf{p}_i + \sum_j w_{ij}\mathbf{g}_j\|_2^2 \\
=&\|\sum_j w_{ij}\mathbf{p}_i + \sum_j w_{ij}\mathbf{g}_j\|_2^2 \\
=&\|\sum_j w_{ij}(\mathbf{p}_i + \mathbf{g}_j)\|_2^2,
\end{aligned} \tag{A.1}
$$

where $\mathbf{p}_i = (K-1)\mathbf{g}_i - \sum_j \mathbf{g}_j - \sum_j m_{ij}(\mathbf{g}_i - \mathbf{g}_j)$ can be numerically computed. For simplicity, we denote $\mathbf{C}_{js} = (\mathbf{p}_i + \mathbf{g}_j) \cdot (\mathbf{p}_i + \mathbf{g}_s)$. Then, above equation is rewritten as:

$$
\|\sum_j w_{ij}(\mathbf{p}_i + \mathbf{g}_j)\|_2^2 = \sum_{js} w_{ij} w_{is} \mathbf{C}_{js}, \tag{A.2}
$$

which can be minimized via Lagrange multiplier method by enforcing $\sum_j w_{ij} = 1$:

$$
\mathcal{L} = \sum_{js} w_{ij} w_{is} \mathbf{C}_{js} + \gamma\|\mathbf{w}_i\|_2 + \lambda(\sum_j w_{ij} - 1), \tag{A.3}
$$

Through solving $\partial\mathcal{L}/\partial w_{ij} = 0$, we can obtain the closed form of the optimal weights:

$$
(\sum_s \mathbf{C}_{js} + \gamma) w'_{ij} = \frac{-\lambda}{2}, \tag{A.4}
$$

$$
w_{ij} = \frac{w'_{ij}}{\sum_j w'_{ij}}. \tag{A.5}
$$

B Computing Embedding Matrix

The analytic form of Eq.(5) can be derived as:

$$
\begin{aligned}
&\sum_i \|K\mathbf{t}_i - \sum_j \big(\mathbf{t}_j + h_{ij}(\mathbf{t}_i - \mathbf{t}_j)\big)\|_2^2 \\
=&\sum_i \|(K-2)\mathbf{t}_i - \sum_j \mathbf{t}_j + \sum_j h_{ij}\mathbf{t}_j\|_2^2 \\
=&\sum_i \|(K-2)(\mathbf{t}_i - \sum_j \frac{1-h_{ij}}{K-2}\mathbf{t}_j)\|_2^2,
\end{aligned} \tag{B.1}
$$

which equals to minimize following

$$
\Phi(\mathbf{T}) = \sum_i \|\mathbf{t}_i - \sum_j h'_{ij}\mathbf{t}_j\|_2^2. \tag{B.2}
$$

where $h'_{ij} = \frac{1-h_{ij}}{K-2}$, and omit the constant coefficient $K-2$. The cost essentially defines a quadratic form:

$$
\Phi(\mathbf{T}) = \sum_{ij} \mathbf{Q}_{ij}(\mathbf{t}_i \mathbf{t}_j), \tag{B.3}
$$

which involves inner products of the embedding vectors and the $N \times N$ matrix $\mathbf{Q}$:

$$
\mathbf{Q}_{ij} = \delta_{ij} - h'_{ij} - h'_{ji} + \sum_s h'_{si} h'_{sj}, \tag{B.4}
$$

where δ_{ij} is 1 if $i = j$ and 0 otherwise. The optimal embedding is found by computing the bottom $d+1$(in our case, $d = 2$) eigenvectors of the matrix $\mathbf{Q}$.

We can calculate $\mathbf{Q}$ following (Bai et al. 2000) without performing a full matrix diagonalization:

$$
\begin{aligned}
\Phi(\mathbf{T}) &= \sum_i \|\mathbf{t}_i - \sum_j h'_{ij}\mathbf{t}_j\|_2^2 \\
&= \sum_i \|\mathbf{T}\mathbf{I}_i - \mathbf{T}\mathbf{h}_i'^{\intercal}\|_2^2 \\
&= \mathrm{Tr}(\mathbf{T}(\mathbf{I} - \mathbf{H}'^{\intercal})(\mathbf{I} - \mathbf{H}'^{\intercal})^{\intercal}\mathbf{T}^{\intercal}) \\
&= \mathrm{Tr}(\mathbf{T}\mathbf{Q}\mathbf{T}^{\intercal}),
\end{aligned} \tag{B.5}
$$

where $\mathbf{Q} = (\mathbf{I}-\mathbf{H}'^{\intercal})(\mathbf{I}-\mathbf{H}'^{\intercal})^{\intercal}$, and $\mathbf{I}$ is the identity matrix. The optimal embedding is found by computing the eigenvectors corresponding to the the smallest $d+1$ eigenvalues of the matrix $\mathbf{Q}$, which is well defined in the Rayleitz-Ritz theorem (Horn and Johnson 2012). And we discard the bottom eigenvector because it is usually the unit vector with all equal components. The eigenvectors of shape $N \times d$ form the embedding vectors $\mathbf{T} = \{\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_N\}$ in d-dimensional space.

C Evaluation Protocols

We use following metrics to measure the prediction performance:

1. Root Mean Squared Error (RMSE):

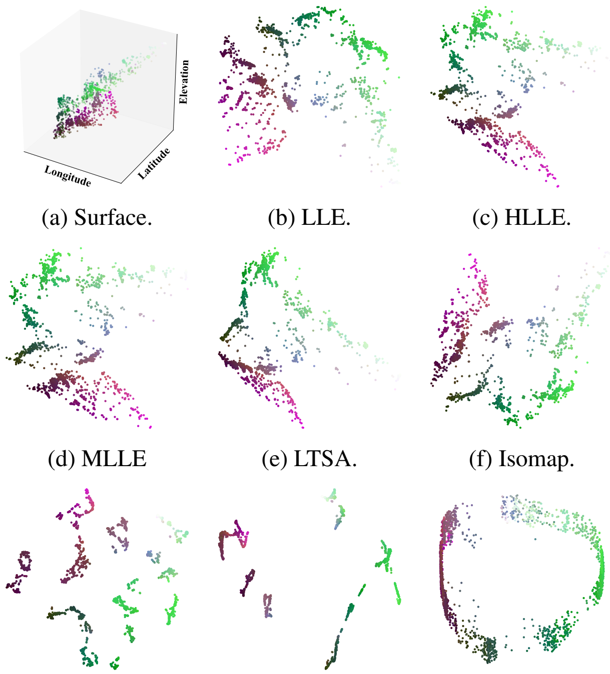
$$
RMSE = \sqrt{\frac{1}{T'N}\sum_{j=1}^{T'}\sum_{i=1}^{N}\left(\boldsymbol{m}_i^{T+j} - \hat{\boldsymbol{m}}_i^{T+j}\right)^2}
$$

2. Mean Absolute Error (MAE):

$$
MAE = \frac{1}{T'N}\sum_{j=1}^{T'}\sum_{i=1}^{N}\left|\boldsymbol{m}_i^{T+j} - \hat{\boldsymbol{m}}_i^{T+j}\right|
$$

3. Accuracy:

$$
Accuracy = 1 - \frac{\|\mathbf{M} - \widehat{\mathbf{M}}'\|_F}{\|\mathbf{M}\|_F}
$$



(a) Surface. (b) LLE. (c) HLLE.
(d) MLLE (e) LTSA. (f) Isomap.
(g) t-SNE. (h) UMAP. (i) WLLE.

Figure 1: The embedded space of all methods on east side.

4. Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\sum_{j=1}^{T'} \sum_{i=1}^{N} \left(\boldsymbol{m}_i^{T+j} - \hat{\boldsymbol{m}}_i^{T+j} \right)^2}{\sum_{j=1}^{T'} \sum_{i=1}^{N} \left(\boldsymbol{m}_i^{T+j} - \overline{\mathbf{M}} \right)^2}$$

5. Explained Variance Score (var):

$$var = 1 - \frac{var\{\mathbf{M} - \widehat{\mathbf{M}}'\}}{Var\{\mathbf{M}\}}$$

where $\boldsymbol{m}_i^{T+j}$ and $\hat{\boldsymbol{m}}_i^{T+j}$ represent the true and predicted measurement at time $T+j$. The range of predictions is from T to $T+T'$ on all N nodes. $\mathbf{M}$, $\widehat{\mathbf{M}}$ and $\overline{\mathbf{M}}$ are the real deformations, predicted deformations and the average deformations, respectively.

D Additional Visualizations

Figure 1 shows the embedding results of all methods on the east side. Similar to the results on west side, manifold learning methods (e.g., LLE, HLLE, IsoMap) are very close on the final node embeddings, while dimension reduction methods t-SNE and UMAP obtain intra-clustered but inter-distanced embeddings. Our WLLE, which is specifically designed for learning relative positions, preserves both the distance and the slope information in the latent space.

References

Bai, Z.; Demmel, J.; Dongarra, J.; Ruhe, A.; and van der Vorst, H. 2000. *Templates for the solution of algebraic eigenvalue problems: a practical guide*. SIAM.

Horn, R. A.; and Johnson, C. R. 2012. *Matrix analysis*. Cambridge university press.