Appendix A. Supplementary figures

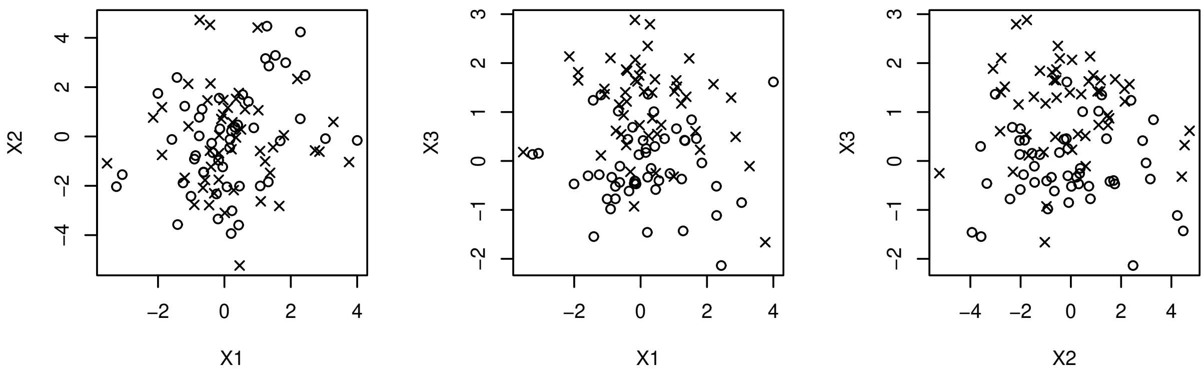

Figure 6: Configuration of simulated data following Equation 5, displayed in two-dimension projected from each axis.

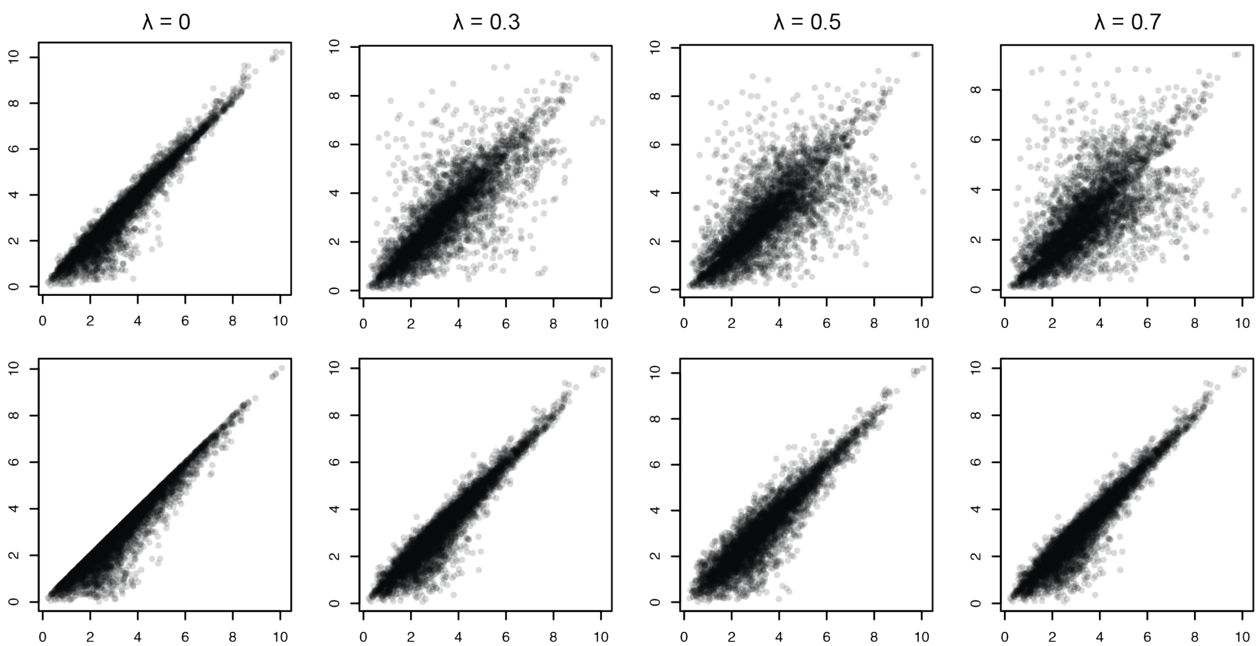

Figure 7: Shepard plot in simulation data using SuperMDS [2] (first row) and proposed MDS (second row). Because a different hyperparameter is used in the two methods, a comparison is made based on a ratio between confirmatory and MDS term in the objective function (Equation 4).

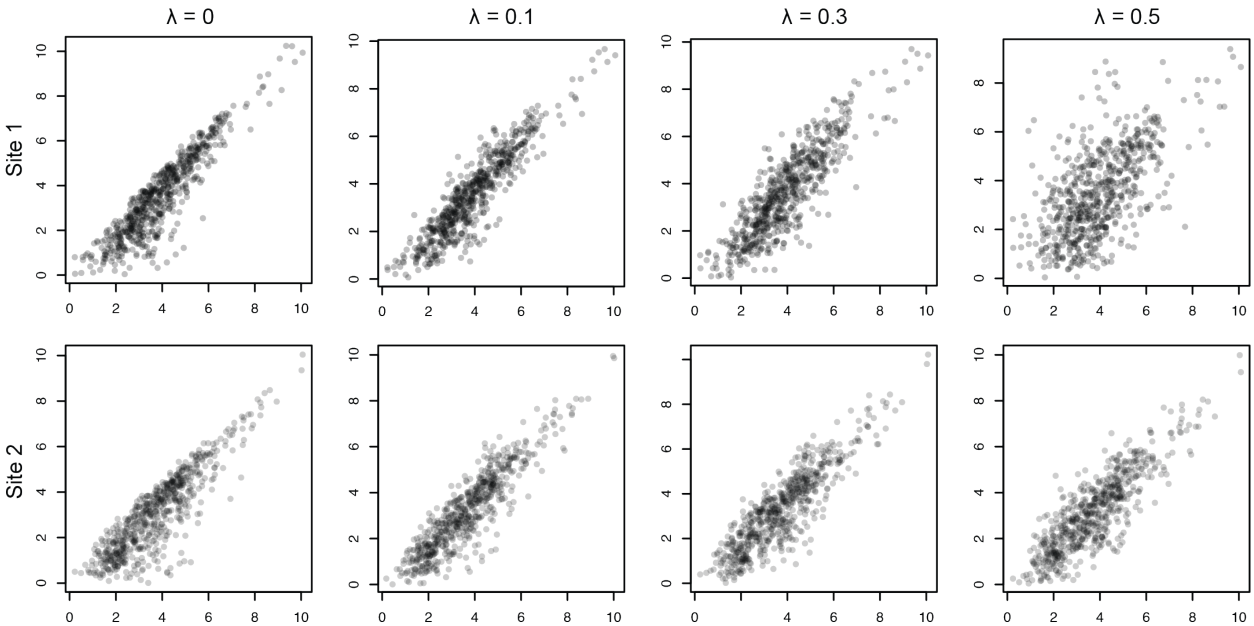

Figure 8: Shepard plot in microbial community data collected from site 1 (first row) and site 2 (second row) for a range of hyperparameter λ.

Appendix B. Derivation of mapping function $f_{\mathbf{z}}(\Phi_o)$

The confirmatory term in Equation (4) originates in an effort to minimize a difference in the distribution of pseudo F of the original and two-dimensional data, represented by the p-values of each pseudo F statistics. Here p-value is obtained by a quantile of pseudo F among its distribution of a set of permuted labels [4].

To estimate the closest pseudo F minimizing the difference in p-values, we first derive the following statistics

$$\Phi_o^\Pi = \frac{\sum_{i,j} \mathbb{I}\{y_i^\Pi \neq y_j^\Pi\} d_{ij}^2}{\sum_{i,j} \mathbb{I}\{y_i^\Pi = y_j^\Pi\} d_{ij}^2}, \quad \Phi_{\mathbf{z}}^\Pi = \frac{\sum_{i,j} \mathbb{I}\{y_i^\Pi \neq y_j^\Pi\} \|\mathbf{z}_i - \mathbf{z}_j\|_2^2}{\sum_{i,j} \mathbb{I}\{y_i^\Pi = y_j^\Pi\} \|\mathbf{z}_i - \mathbf{z}_j\|_2^2}, \tag{6}$$

where the superscript Π denotes a permutation of labels $\{y_i\}$.

Next step is to correlate Φ_o^Π and $\Phi_{\mathbf{z}}^\Pi$ by repeating the permutation and performing a local regression (LOESS) on the pair set. This allows us to obtain an approximation of function $f_{\mathbf{z}}(\cdot)$ that maps Φ_o to $\Phi_{\mathbf{z}}$ (Figure 9).

In addition, we explain why Φ_o, $\Phi_{\mathbf{z}}$ in Equation (6) represent our pseudo F statistics in terms of obtaining the p-values. This is simply by observing the pseudo F of the original and two-dimensional

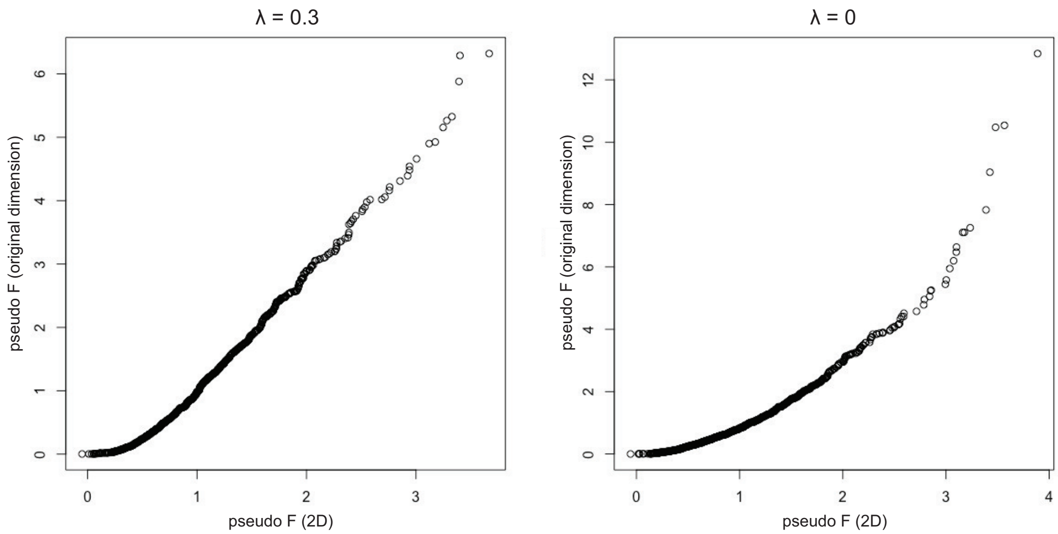

Figure 9: Correlation plot of pseudo F in the orginal and two-dimensional by permuting y labels over 1,000 iteration. Created by setting a hyperparameter λ for MM algorithm (left, 0.3; right, 0).

configuration,

$$F_{\mathbf{z}} = (N-2) \cdot \frac{\sum_{i,j} \|\mathbf{z}_i - \mathbf{z}_j\|_2^2}{\sum_{i,j} \mathbb{I}\{y_i = y_j\} \|\mathbf{z}_i - \mathbf{z}_j\|_2^2} = (N-2)(1+\Phi_{\mathbf{z}}),$$

$$F_o = (N-2) \cdot \frac{\sum_{i,j} d_{ij}^2}{\sum_{i,j} \mathbb{I}\{y_i = y_j\} d_{ij}^2} = (N-2)(1+\Phi_o),$$

so it suggests that the same p-value can be obtained by calculating a quantile of Φ-distribution.

Finally, we derive the confirmatory term in Equation (4) by observing

$$\begin{aligned}
&\arg\min_{\mathbf{z}} \, |\Phi_{\mathbf{z}}(\mathbf{z}) - f_{\mathbf{z}}(\Phi_o)| && (7)\\
&= \arg\min_{\mathbf{z}} \left| \frac{\sum_{i,j} \mathbb{I}\{y_i \neq y_j\} \|\mathbf{z}_i - \mathbf{z}_j\|_2^2}{\sum_{i,j} \mathbb{I}\{y_i = y_j\} \|\mathbf{z}_i - \mathbf{z}_j\|_2^2} - f_{\mathbf{z}}(\Phi_o) \right| && (8)\\
&\approx \arg\min_{\mathbf{z}} \left| \sum_{i,j} \mathbb{I}\{y_i \neq y_j\} \|\mathbf{z}_i - \mathbf{z}_j\|_2^2 - f_{\mathbf{z}}(\Phi_o) \cdot \sum_{i,j} \mathbb{I}\{y_i = y_j\} \|\mathbf{z}_i - \mathbf{z}_j\|_2^2 \right| && (9)\\
&= \arg\min_{\mathbf{z}} \left| \sum_{i,j} \left[1 - (f_{\mathbf{z}}(\Phi_o) + 1)\mathbb{I}\{y_i = y_j\}\right] \|\mathbf{z}_i - \mathbf{z}_j\|_2^2 \right|. && (10)
\end{aligned}$$

Note that the last equation is strictly convex in regards to $\mathbf{z}$; therefore, when it is added to the MDS term we ensure that the MM algorithm will still work.

Appendix C. Derivation MM algorithm for proposed MDS

For each $k = 1, \cdots N$, we want to find $\mathbf{z}^*_{\mathbf{k}}$ such that

$$\begin{aligned}
\mathbf{z}_k^* &= \arg\min_{\mathbf{z}_k} O(\mathbf{z}) && (11)\\
&= \arg\min_{\mathbf{z}_k} \sum_{i,j} (d_{ij} - \|\mathbf{z}_i - \mathbf{z}_j\|_2)^2 + \lambda \left| \sum_{i,j} [1 - (f_{\mathbf{z}}(\Phi) + 1)\epsilon_{ij}] \|\mathbf{z}_i - \mathbf{z}_j\|_2^2 \right| && (12)\\
&= \arg\min_{\mathbf{z}_k} \sum_{j=1}^{N} (d_{jk} - \|\mathbf{z}_j - \mathbf{z}_k\|_2)^2 + \lambda\delta(\mathbf{z}) \sum_{i,j} [1 - (f_{\mathbf{z}}(\Phi) + 1)\epsilon_{ij}] \|\mathbf{z}_i - \mathbf{z}_j\|_2^2 && (13)\\
&= \arg\min_{\mathbf{z}_k} \sum_{j=1}^{N} \|\mathbf{z}_k - \mathbf{z}_j\|_2^2 - 2\sum_{j=1}^{N} d_{jk} \|\mathbf{z}_k - \mathbf{z}_j\|_2 + \lambda\delta(\mathbf{z}) \sum_{j=1}^{N} [1 - (f_{\mathbf{z}}(\Phi) + 1)\epsilon_{jk}] \|\mathbf{z}_k - \mathbf{z}_j\|_2^2 && (14)\\
&= \arg\min_{\mathbf{z}_k} (1 + \lambda\delta(\mathbf{z})) \sum_{\substack{j=1 \\ \epsilon_{jk}=0}}^{N} \|\mathbf{z}_k - \mathbf{z}_j\|_2^2 + (1 - \lambda f_{\mathbf{z}}(\Phi)\delta(\mathbf{z})) \sum_{\substack{j=1 \\ \epsilon_{kj}=1}}^{N} \|\mathbf{z}_k - \mathbf{z}_j\|_2^2 - 2\sum_{j=1}^{N} d_{jk} \|\mathbf{z}_k - \mathbf{z}_j\|_2 && (15)
\end{aligned}$$

where we have defined

$$\epsilon_{ij} = \mathbb{1}\{y_i = y_j\}, \quad \delta_i(\mathbf{z}) = \text{sign} \sum_{j=1}^{N} [1 - (f_{\mathbf{z}}(\Phi) + 1)\epsilon_{ij}] \|\mathbf{z}_i - \mathbf{z}_j\|_2^2$$

for simplicity. Next we majorize Equation 15 by writing

$$(1 + \lambda\delta(\mathbf{z})) \sum_{\substack{j=1 \\ \epsilon_{jk}=0}}^{N} \|\mathbf{z}_k - \mathbf{z}_j\|_2^2 + (1 - \lambda f_{\mathbf{z}}(\Phi)\delta(\mathbf{z})) \sum_{\substack{j=1 \\ \epsilon_{jk}=1}}^{N} \|\mathbf{z}_k - \mathbf{z}_j\|_2^2 - 2\sum_{j=1}^{N} d_{jk} \frac{\sum_{s=1}^{2} (z_{ks} - z_{js})(\tilde{z}_{ks} - z_{js})}{\|\tilde{\mathbf{z}}_k - \mathbf{z}_j\|_2}. \tag{16}$$

Knowing the above is convex and quadratic, we take a derivative with respect to z_{ks} to find its minimum at $\mathbf{z}_k = \mathbf{z}_k^\dagger$. Taking derivative in Equation 16 and setting it zero we have

$$2(1 + \lambda\delta(\mathbf{z})) \sum_{\substack{j=1 \\ \epsilon_{jk}=0}}^{N} (z_{ks}^\dagger - z_{js}) + 2(1 - \lambda f_{\mathbf{z}}(\Phi)\delta(\mathbf{z})) \sum_{\substack{j=1 \\ \epsilon_{kj}=1}}^{N} (z_{ks}^\dagger - z_{js}) - 2\sum_{j=1}^{N} d_{jk} \frac{\tilde{z}_{ks} - z_{js}}{\|\tilde{\mathbf{z}}_k - \mathbf{z}_j\|_2} = 0$$

$$
\therefore z_{ks}^{\dagger} = \frac{2}{2(N-1) + \lambda\delta(\mathbf{z})(N - (N-2) f_{\mathbf{z}}(\Phi))} \cdot \\
\left[(1 + \lambda\delta(\mathbf{z})) \sum_{\substack{j=1 \\ \epsilon_{jk}=0}}^{N} z_{js} + (1 - \lambda f_{\mathbf{z}}(\Phi)\delta(\mathbf{z})) \sum_{\substack{j=1 \\ \epsilon_{jk}=1}}^{N} z_{js} + \sum_{j=1}^{N} d_{jk} \frac{\tilde{z}_{ks} - z_{js}}{\|\tilde{\mathbf{z}}_k - \mathbf{z}_j\|_2} \right], \quad (17)
$$

giving the update rule as written in Algorithm 1. □