## APPENDIX A
## ADDITIONAL EXPERIMENTAL RESULTS

The RF-GAP proximities perfectly preserve the random-forest predictive ability as was shown in [3]. In Figure A.1, we show that the predictive ability is well-preserved in low dimensions. In the figure, we see the diffusion-based methods as well as RFTSNE and RFUMAP methods nearly perfectly preserve the OOB score when using a $k$-NN classifier on the embeddings.

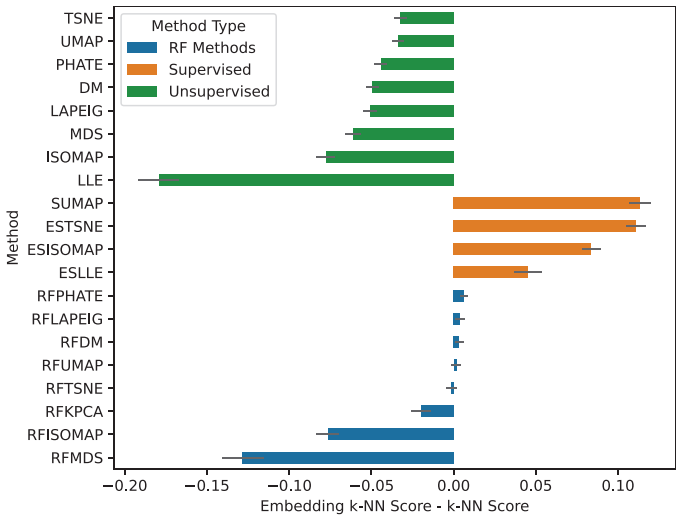


Fig. A.1. The difference between the OOB score from a trained random forest and the $k$-NN cross-validated score on the low-dimensional embedding. $k$-NN models trained on the RF-GAP diffusion-based embeddings as well as RFUMAP and RFTSNE embeddings nearly perfectly match OOB accuracies.

The following figures provide examples to compare unsupervised, supervised, and RF-GAP-based methods. In most examples, the supervised models create near-linear separation between classes, while RF-GAP methods preserve relationships meaningful to the supervised task without global structure disruption.

## APPENDIX B
## DATASETS USED

TABLE B.1
THIS TABLE CONTAINS THE DESCRIPTIONS DATASETS USED IN THE QUANTITATIVE EXPERIMENTS IN FIGURES 3, 4, AND A.1. OBSERVATIONS WITH MISSING VALUES WERE REMOVED, AND UNIQUELY-IDENTIFYING VARIABLES WERE ALSO REMOVED. EACH DATASET WAS NORMALIZED BEFORE APPLYING DIMENSIONALITY REDUCTION.

| Data | Observations | Variables | Source |
|---|---|---|---|
| Banknote | 1372 | 5 | UCI |
| Breast Cancer | 699 | 16 | UCI |
| Car | 1728 | 6 | UCI |
| Diabetes | 678 | 8 | UCI |
| E. Coli | 336 | 8 | UCI |
| Glass | 214 | 10 | UCI |
| Heart Disease | 303 | 13 | UCI |
| Hill Valley | 606 | 101 | UCI |
| Ionosphere | 351 | 34 | UCI |
| Iris | 150 | 4 | UCI |
| Liver | 345 | 7 | UCI |
| Lymphography | 148 | 18 | UCI |
| Optical Digits | 3823 | 64 | UCI |
| Parkinson's | 197 | 23 | UCI |
| Seeds | 210 | 7 | UCI |
| Sonar | 208 | 60 | UCI |
| Tic-Tac-Toe | 958 | 9 | UCI |
| Waveform | 5000 | 21 | UCI |
| Wine | 178 | 13 | UCI |

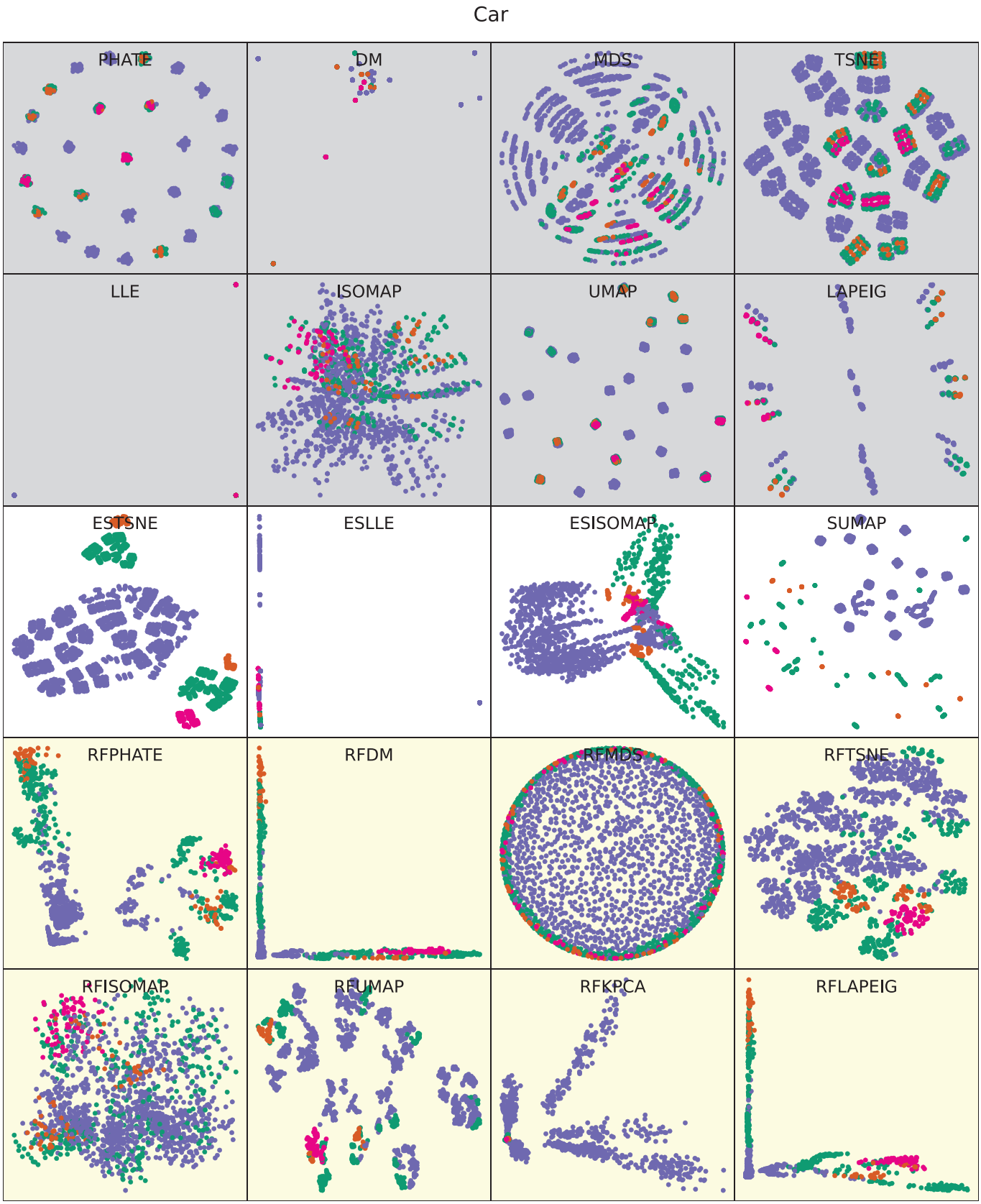


Fig. A.2. 20 compared embeddings on the cars dataset [14]. Embedding types are depicted by background color: grey for unsupervised, white for class-conditional supervised, and yellow for RF-GAP methods.

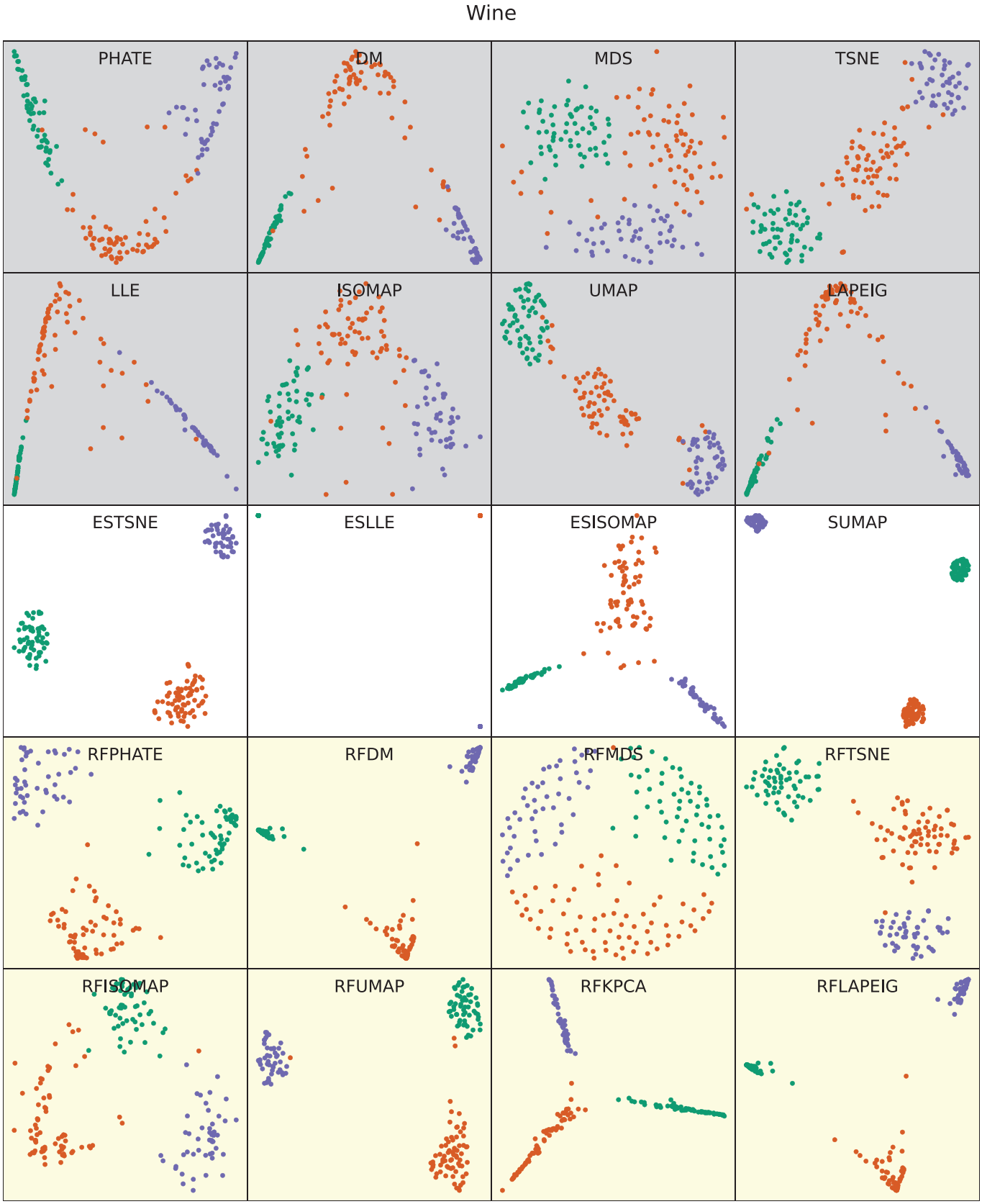


Fig. A.3. The wine dataset [22] is simple enough that using labeled information does not add much to the embedding. At the same time, the RF-GAP methods are not diminished by the gained information from the random forest.


Fig. A.4. The sonar dataset [23] reduced to two dimensions. The unsupervised models fail to find meaningful patterns, while the non-random forest supervised creates linear separation between classes. The RF-GAP models meaningfully retain the random forest model's learning.

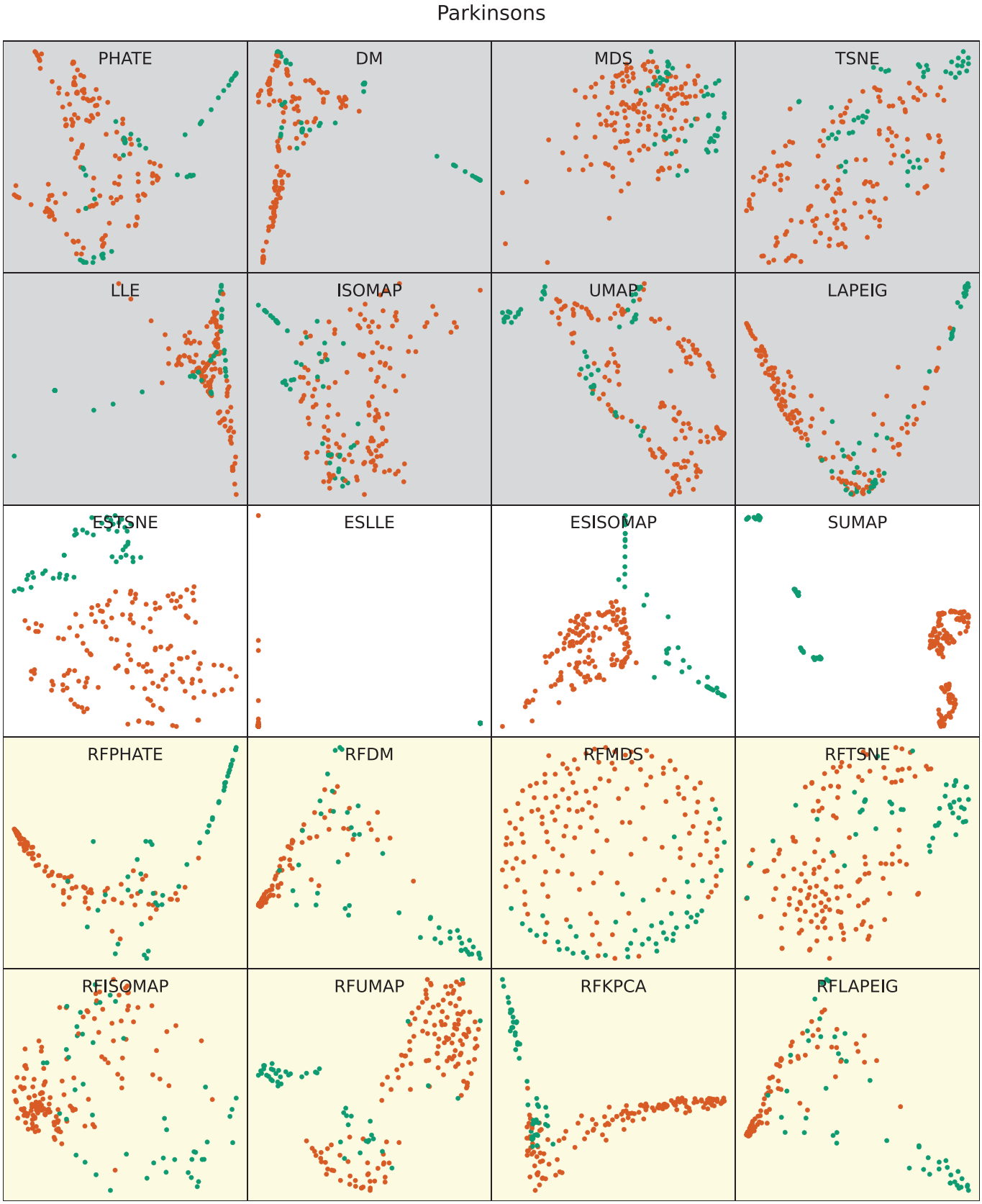


Fig. A.5. This example of the Parkinson's dataset [22] shows that the RF-GAP embeddings provide a somewhat clearer relationship between the two classes, while the other supervised methods provide perfect separation and the unsupervised methods contain extra noise.