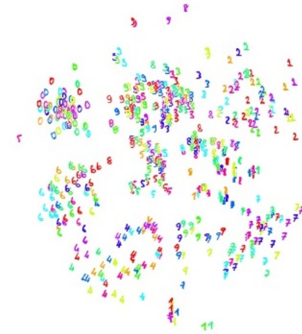
We warmly thank the four reviewers for their work and constructive feed-backs. Due to the one page limit, we briefly address here the most crucial comments in groups (R1, R2, R3 and R4 denote concerns raised by the corresponding reviewers). In the revised paper, we will of course do our best to address all reviewers' comments using the additional page allowed for the camera-ready.

ClassNeRV use cases (R1, R2) ClassNeRV is designed for exploratory analysis of labeled data, as unsupervised techniques, but steers unavoidable distortions to minimize their impact on classes. Thus, it shows the global structure of classes (class segmentation) rather than extracting features for classification (class separation). It is useful to detect if classes are well-separated or not given a data feature space, to question the labels (meaning of the classes) or the features (feature engineering). For the Isolet data, global structure of ClassNeRV map could help a domain expert discover that, in this feature space, letters are strongly grouped by vowel sounds (BDEGPTV, FS or MN), with a secondary effect of the consonant. This will be clarified in introduction and Isolet interpretation.

Advantages of Neighbourhood Embedding (NE) family (R3, R4) NE methods benefit from interesting practical properties such as shift-invariance, making them robust to the curse of dimensionality, which justifies better performances of ClassNeRV compared to Classimap on high dimensional datasets, as observed on Isolet data (Figure 4). On the theoretical side, NE has been explained through a probabilistic framework as a tool for performing a neighbourhood retrieval task in the map [Venna *et al.* 2010], leading to more interpretability for the visual exploration process. **Choice of NeRV among that family (R1)** NeRV is better-suited than other more popular methods such as tSNE due to its divergence penalizing both false and missed neighbours through two independent terms that may be balanced, providing a built-in tunability. We may note that JSE also satisfies those properties, and that we plan to extend the approach to ClassJSE in a longer paper. The reasons of that choice will be incorporated in the section concerning the NE family.

Hyper-parameters (R2) Figure 2 illustrates the flexibility of the method by showing its sensitivity to several values of $\tau^{\in}$ and $\tau^{\notin}$. Based on that, we then restrict the number of parameters by fixing $\tau^* = 0.5$ and $\epsilon = 0.5$ for the supervised ClassNeRV. An ablation study will be added in supplemental to show individual impact of each of the four components of ClassNeRV stress. **Equivalence of unsupervised ClassNeRV ($\tau^{\in} = \tau^{\notin}$) and NeRV (R2)** We detail here (and will add in supplemental) why ClassNeRV with parameters $\tau^{\in} = \tau^{\notin} = \tau$, where $\tau \in [0, 1]$, is unsupervised and corresponds to NeRV with trade-off parameter τ. In that case, ClassNeRV stress (Equation 3) may be factored by τ and $(1-\tau)$, so that the sums of within class terms (*i.e.* $\sum_{i,j\in\mathcal{S}_i^{\in}} ...$) and between class terms (*i.e.* $\sum_{i,j\in\mathcal{S}_i^{\notin}} ...$) collapse in a sum of all terms that does not take into account the class-information (*i.e.* $\sum_{i,j\in\mathcal{S}_i^{\in}\cup\mathcal{S}_i^{\notin}} ... = \sum_{i,j\neq i} ...$), leading to: $\zeta_{\text{ClassNeRV}} = \tau \sum_{i,j\neq i} \left(\beta_{ij} \log\left(\frac{\beta_{ij}}{b_{ij}}\right) + b_{ij} - \beta_{ij}\right) + (1-\tau) \sum_{i,j\neq i} \left(b_{ij} \log\left(\frac{b_{ij}}{\beta_{ij}}\right) + \beta_{ij} - b_{ij}\right)$. Knowing that $\sum_{j\neq i} \beta_{ij} = \sum_{j\neq i} b_{ij} = 1$ (due to the normalization in Equation 1), $\sum_{j\neq i} \beta_{ij}$ and $\sum_{j\neq i} b_{ij}$ cancel each other out, so that $b_{ij} - \beta_{ij}$ and $\beta_{ij} - b_{ij}$ terms may be removed from the above equation. As a result, the Bregman divergence becomes a Kullback-Leibler divergence and ClassNeRV stress equals the stress of NeRV (Equation 3).

Supplementary datasets and quality indicators (R1, R2, R3, R4) In the submitted paper, we chose to focus on a few datasets with detailed interpretation and evaluation. We especially resorted to several toy datasets to illustrate limitations of existing unsupervised and supervised DR techniques, as well as the premises of our methodology, which seems to be successful, since all the reviewers well understood our approach. Yet, we fully agree that the confirmatory results obtained with Isolet data are not sufficient, and we would like to benefit from the supplementary page of the camera-ready paper to add results on other high dimensional data. The choice of a well-known image dataset such as the SculptFaces appears very relevant, since it allows representations that intuitively show the true similarities between data points. As this specific dataset does not contain class-information, we propose to use the similar handwritten digits dataset, both with its true labels and with randomly selected labels. The latter provides a case of high dimensional data with conflicting neighbourhood and class structures (as observed in Figure 1d). Example 1 shows a preliminary ClassNeRV map for this dataset, with digits shapes coloured based on the random classes provided to the algorithm. We see that the preservation of neighbourhoods prevails over the preservation of classes, with the fake classes remaining mixed. Some of the many supervised indicators of the literature will be added. Yet, most of them being based on k-NN performances in the embedding space [Maaten, Postma, Herrick 2009, Venna *et al.* 2010, de Bodt *et al.* 2019] , they should show the same trends as the k-NN accuracy presented in the paper.



Ex. 1: Random labels digits

Other issues (R1, R3, R4) The suggested references will be incorporated. The normalizing terms $\mathcal{T}_{\max}(\kappa) = \mathcal{C}_{\max}(\kappa)$ are given by $\kappa N(2N - 3\kappa - 1)/2$ if $\kappa \leq N/2$ and $N(N - \kappa - 1)/2$ if $\kappa > N/2$ [Venna PhD Thesis]. The leave one-out-classifier (in k-NN accuracy) attributes to each point i the majority label (winner takes all strategy) of it k nearest neighbours (among all points except i) and the equality case is decided randomly.