We thank the reviewers for their valuable and detailed comments. Before addressing individual concerns, we highlight that our paper is the first functional map-based method that computes correspondences directly on point clouds and uses a learned basis, made possible through our linearly-invariant embedding formulation. We believe that this will encourage further work as highlighted by the reviewers and pointed out in the following.

We will release the source code to reproduce all the results and we will include all the details about our implementation and the hyperparameters in the supplementary material. We will also add all the requested additional analysis and discussions. If we have enough space, we will move a compact version of Fig. 1 from the supplementary materials to the main manuscript.

Necessity of learning the transformation $A_{\mathcal{X}\mathcal{Y}}$ (R4): The goal of our method is to predict unknown correspondences between a pair of shapes at test time. The closed-form expression for $A_{\mathcal{X}\mathcal{Y}}$ assumes the knowledge of ground truth correspondences, which is only available during training. At test time, we use the learned transformation matrix to estimate the correspondences. We will clarify this in the final version.

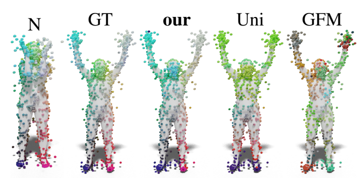
Impact of the transformation matrix (R1): As suggested by **R1**, the inset table here shows the results with a 60 dimensional Universal Embedding (Uni). We consider that an improvement of 20% from the baseline is a promising starting point for the first approach in a new direction. Furthermore, in the partial experiment we show that our basis is very robust, while the Uni suffers from this kind of changes.

	noNoise	Noise
our	**5.4e-2**	**6.6e-2**
Uni20	7.5e-2	8.5e-2
Uni60	6.9e-2	8.1e-2

Alternative losses for Eq. 3: The loss proposed in Eq. 3 is a simple one that is supported by existing work. We tested extensively all of the alternative losses mentioned on lines 237-247 of the main manuscript and will be happy to provide the results and the properties of the resulting bases as an ablation study.

Outliers in the experiments (R1): Our method is very robust. In the inset table, we show quantitative results on **100 pairs** of the test set with 30% outlier points, compared to the best baselines. We also show a qualitative example on one shape pair in the figure below. We will be happy to include experiments with different levels and density of outliers in the final version. Note also that we considered different kind of noise in Figures 2 and 3 of the main manuscript. The training set does not contain this type of data, highlighting the robustness of our method.

	Outlier
our	**2.7e-1**
Uni60	3.5e-1
GFM	3.5e-1



Qualitative result with outliers

Limitations (R2): The main limitations of our method are its global and supervised nature. We believe that extending our method to use a multiscale feature extractor and unsupervised losses are both remarkable future directions.

Baselines (R2): While we do not compare to Deep Functional maps [30] (which relies on having a mesh as input for its feature extractor), we compare to the most recent state-of-the-art method GFM [12] that outperforms [30].

Stability of the basis (R2): We provide visualizations of the estimated basis (resp. descriptors) on some real scans from the FAUST dataset in Fig. 4 and 5 (Fig. 6, 7, 8) of the supplementary. These scans present missing parts, holes and partiality and are represented by different samplings. We will be happy to include additional visualizations, showing the stability and robustness to different sampling.

Additional experiments (R3): We will be happy to give more details about the adopted evaluation as suggested by **R3** (Figure 2 and 3) and on the results obtained through our pipeline adding more in-depth analysis, including adding more noise/outlier results and using different basis sizes as mentioned above.

Descriptors from basis (R2): Although this is not our main goal, the use of our learned embedding as a basis for descriptor computation as proposed by **R2** is an excellent idea. We will be happy to add illustrations to highlight properties of the basis. Note that the additional request for the basis to be orthonormal could be well-suited in this case.

Two stage training (R4): Using two stage training helps to regularize the network. Importantly our embedding network loss is designed to promote the existence of *some* linear transformation across different embeddings, which is exploited by our transformation network. We will clarify this.

Notation and references (R1, R2, R4): We refer to *linear invariance* since the loss we use to train the embedding is invariant to linear transformations. We will clarify this and will move the PointNet sentence mentioned by **R2** to the main body. We will also clarify the notation (including line 231 – $A_{\mathcal{X}\mathcal{Y}}$ should be marked gt), fix the typos and include all the references suggested by the reviewers.