

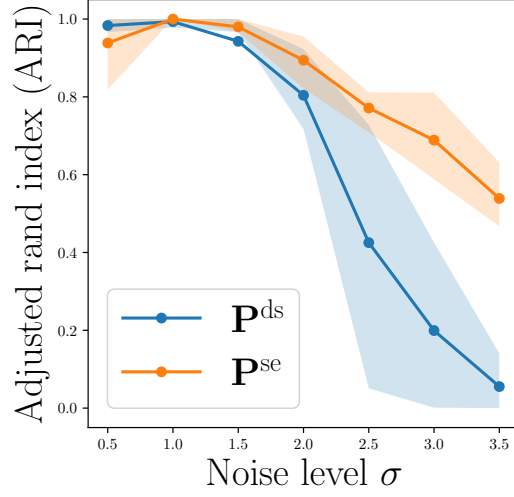


Figure 1: ARI spectral clustering score as a function of the noise on the example of three Gaussian clusters with varying variance : σ^2 , $2\sigma^2$ and $3\sigma^2$ (as in figure 2 in the main paper).

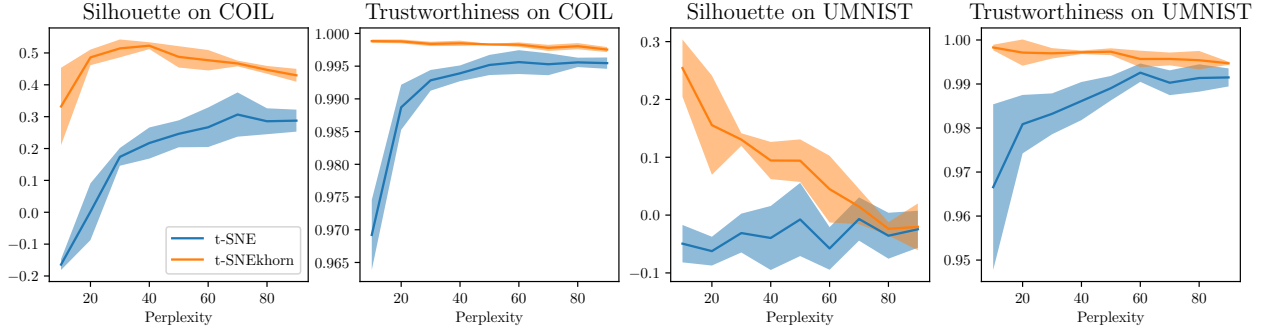


Figure 2: Dimensionality reduction scores as a function of the perplexity parameter.

Table 1: Timings (in seconds) for t-SNE and t-SNEkhorn.

	t-SNE			t-SNEkhorn		
	Affinity	Embedding	Total	Affinity	Embedding	Total
COIL	0.38 ± 0.01	5.28 ± 0.17	5.66 ± 0.17	13.02 ± 0.34	13.39 ± 0.07	26.41 ± 0.35
OLIVETTI	0.07 ± 0.02	0.96 ± 0.02	1.03 ± 0.03	0.26 ± 0.02	2.42 ± 0.08	2.68 ± 0.08
UMNIST	0.12 ± 0.01	4.27 ± 0.05	4.39 ± 0.05	1.77 ± 0.03	8.50 ± 0.16	10.27 ± 0.16

Table 2: Silhouette scores with varying dimensions for the pre-processing PCA step.

Dim. PCA	t-SNE			t-SNEkhorn		
	30	50	70	30	50	70
COIL	31.2 ± 3.1	30.7 ± 6.9	30.2 ± 4.2	52.6 ± 1.3	52.3 ± 1.1	51.8 ± 0.9
OLIVETTI	3.9 ± 1.2	4.5 ± 3.1	4.4 ± 2.7	15.2 ± 1.8	15.7 ± 2.2	15.5 ± 1.2
UMNIST	-0.1 ± 1.9	-0.2 ± 1.5	-0.2 ± 1.5	26.1 ± 4.6	25.4 ± 4.9	23.1 ± 3.8