

0.5% CS															KMEANS	GT
Objects	Features	Clusters	PCCC-N2-S	PCCC-N3-S	PCCC-N4-S	PCCC-N5-S	PCCC-N6-S	PCCC-N2-S-RD	COPKM	LCC	CSC	DILS				
Dataset																
Banana5,300	2	2	0.354	0.354	0.354	0.354	0.354	0.354	–	0.347	0.284	0.143		0.389	0.0	
Letter20,000	16	26	0.134	0.134	0.134	0.134	0.134	0.134	0.134	0.117	–	-0.020		0.145	0.0	
Shuttle57,999	9	7	0.274	0.304	0.304	0.304	0.305	0.210	–	0.005	–	–		0.463	0.3	
CIFAR100,000	3,072	10	0.031	0.031	0.032	0.032	0.032	0.031	–	–	–	–		0.051	-0.0	
CIFAR100	3,072	100	0.013	0.013	0.013	0.013	0.013	0.013	–	–	–	–		0.015	-0.1	
MNIST10,000	784	10	-0.010	-0.009	-0.009	-0.009	-0.009	-0.010	-0.018	–	–	–		0.007	-0.0	
Mean			0.133	0.138	0.138	0.138	0.138	0.122	-0.647*	-0.422*	-0.786*	-0.646*		0.178	0.0	

*Nan values (–) are replaced with -1 before computing the mean.

Table W109: Average Silhouette coefficients of the versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) obtained with constraint sets of size 0.5% CS. Higher values indicate better separated clusters. The highest values are stated in bold. The column KMEANS reports the average Silhouette coefficients that were obtained with the unconstrained k-means algorithm. The column GT reports the Silhouette coefficients of the ground truth assignment. The hyphen indicates that the respective algorithm returned no solution within the time limit of 1,800 seconds.