	0% CS									
Dataset	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS
Banana	0.019	0.019	0.019	0.019	0.019	0.018	−	**0.030**	0.008	0.017
Letter	0.149	0.149	0.149	0.149	0.149	**0.156**	−	−	0.000	0.149
Shuttle	0.193	0.193	0.193	0.193	0.193	**0.429**	−	−	−	0.411
Cifar 10	**0.040**	0.039	0.039	**0.040**	**0.040**	−	−	−	−	0.040
Cifar 100	**0.021**	**0.021**	**0.021**	**0.021**	**0.021**	−	−	−	−	0.021
Mnist	**0.306**	0.304	0.304	**0.306**	**0.306**	0.193	−	−	−	0.312
Mean	0.121	0.121	0.121	0.121	0.121	**0.133**[*]	0.000[*]	0.005[*]	0.001[*]	0.159

[*]Nan values (−) are replaced with 0 before computing the mean.

Table W31: Average Adjusted Rand Index (ARI) values of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) obtained with constraint sets of size 0% CS. Higher values indicate more overlap with the ground truth assignment. The highest values are stated in bold. The column KMEANS reports the average ARI values that were obtained with the unconstrained k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 3,600 seconds. We noticed that the LCC algorithm stops with a runtime error when the constraint set is empty. This is why the LCC algorithm did not return any solutions for the constraint sets of size 0% CS. The different PCCC versions sometimes devise slightly different assignments for the same instance. This is because the Gurobi solver terminates when the relative MIP optimality gap is less than 0.0001. For some instances the different PCCC versions satisfy this stopping criterion with slightly different assignments. The differences between the results of the PCCC versions and the results of the COPKM and the KMEANS algorithms can be explained as follows. The COPKM algorithm uses a different implementation of the kmeans++ algorithm, and the PCCC versions use the Euclidean distance for parameter d_{il} and not the squared Euclidean distance which can lead to different assignments for some instances.

Dataset	0.5% CS									
	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS
Banana	**0.018**	**0.018**	**0.018**	**0.018**	**0.018**	–	0.006	0.000	0.003	0.017
Letter	0.155	0.155	0.154	0.155	0.155	**0.156**	0.148	–	0.000	0.149
Shuttle	0.502	0.502	**0.503**	0.501	0.502	–	0.312	–	–	0.411
Cifar 10	**0.044**	**0.044**	0.044	0.044	0.044	–	–	–	–	0.040
Cifar 100	0.021	**0.022**	0.022	0.022	0.022	–	–	–	–	0.021
Mnist	**0.391**	**0.391**	0.387	**0.391**	**0.391**	0.217	–	–	–	0.312
Mean	0.188	0.188	0.188	**0.189**	**0.189**	0.062*	0.078*	0.000*	0.000*	0.159

*Nan values (–) are replaced with 0 before computing the mean.

Table W32: Average Adjusted Rand Index (ARI) values of five 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 more overlap with the ground truth assignment. The highest values are stated in bold. The column KMEANS reports the average ARI values that were obtained with the unconstrained k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 3,600 seconds.

Dataset	1% CS									
	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS
Banana	**0.014**	**0.014**	0.013	0.013	0.013	–	0.011	-0.000	0.004	0.017
Letter	0.150	0.150	0.150	0.151	0.150	0.154	**0.162**	–	0.000	0.149
Shuttle	**0.975**	**0.975**	0.974	0.974	**0.975**	–	–	–	–	0.411
Cifar 10	0.022	0.022	**0.060**	0.047	0.038	–	–	–	–	0.040
Cifar 100	0.000	0.014	**0.021**	**0.021**	**0.021**	–	–	–	–	0.021
Mnist	0.119	0.119	**0.594**	0.374	0.203	–	–	–	–	0.312
Mean	0.213	0.216	**0.302**	0.263	0.233	0.026*	0.029*	-0.000*	0.001*	0.159

*Nan values (–) are replaced with 0 before computing the mean.

Table W33: Average Adjusted Rand Index (ARI) values of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) obtained with constraint sets 1%. Higher values indicate more overlap with the ground truth assignment. The highest values are stated in bold. The column KMEANS reports the average ARI values that were obtained with the unconstrained k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 3,600 seconds.

| | 5% CS | | | | | | | | | |
Dataset	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS
Banana	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	–	–	0.996	0.100	0.017
Letter	0.651	0.651	0.588	0.658	**0.667**	–	–	–	0.000	0.149
Shuttle	**1.000**	**1.000**	0.985	**1.000**	**1.000**	–	–	–	–	0.411
Cifar 10	**1.000**	**1.000**	0.570	0.570	0.672	–	–	–	–	0.040
Cifar 100	–	–	**0.303**	0.225	0.229	–	–	–	–	0.021
Mnist	**1.000**	**1.000**	0.745	0.749	0.749	–	–	–	–	0.312
Mean	**0.775***	**0.775***	0.699	0.701	0.719	0.000*	0.000*	0.166*	0.017*	0.159

*Nan values (–) are replaced with 0 before computing the mean.

Table W34: Average Adjusted Rand Index (ARI) values of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) obtained with constraint sets of size 5% CS. Higher values indicate more overlap with the ground truth assignment. The highest values are stated in bold. The column KMEANS reports the average ARI values that were obtained with the unconstrained k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 3,600 seconds.

	0% CS										
Dataset	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS	GT
Banana	**0.389**	**0.389**	**0.389**	**0.389**	**0.389**	**0.389**	–	0.388	0.185	0.389	0.032
Letter	**0.145**	**0.145**	**0.145**	**0.145**	**0.145**	0.141	–	–	-0.020	0.145	0.010
Shuttle	0.356	0.356	0.356	0.356	0.356	**0.459**	–	–	–	0.463	0.300
Cifar 10	**0.051**	0.048	0.048	**0.051**	**0.051**	–	–	–	–	0.051	-0.053
Cifar 100	0.015	0.015	0.015	0.015	**0.016**	–	–	–	–	0.015	-0.114
Mnist	0.004	0.003	0.003	0.004	0.004	**0.012**	–	–	–	0.007	-0.043
Mean	**0.160**	0.159	0.159	**0.160**	**0.160**	-0.167[*]	-1.000[*]	-0.769[*]	-0.639[*]	0.178	0.022

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

Table W35: Average Silhouette coefficients of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) obtained with constraint sets of size 0% 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 3,600 seconds. We noticed that the LCC algorithm stops with a runtime error when the constraint set is empty. This is why the LCC algorithm did not return any solutions for the constraint sets of size 0% CS. The different PCCC versions sometimes devise slightly different assignments for the same instance. This is because the Gurobi solver terminates when the relative MIP optimality gap is less than 0.0001. For some instances the different PCCC versions satisfy this stopping criterion with slightly different assignments. The differences between the results of the PCCC versions and the results of the COPKM and the KMEANS algorithms can be explained as follows. The COPKM algorithm uses a different implementation of the kmeans++ algorithm, and the PCCC versions use the Euclidean distance for parameter d_{il} and not the squared Euclidean distance which can lead to different assignments for some instances.

Dataset				0.5% CS							
	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS	GT
Banana	0.353	0.353	**0.358**	**0.358**	**0.358**	–	0.347	0.284	0.143	0.389	0.032
Letter	0.134	0.134	**0.135**	**0.135**	0.134	0.134	0.117	–	-0.020	0.145	0.010
Shuttle	**0.145**	**0.145**	0.145	0.145	0.145	–	0.005	–	–	0.463	0.300
Cifar 10	**0.030**	**0.030**	0.030	0.030	0.030	–	–	–	–	0.051	-0.053
Cifar 100	0.012	**0.014**	**0.014**	**0.014**	**0.014**	–	–	–	–	0.015	-0.114
Mnist	**-0.014**	**-0.014**	-0.015	**-0.014**	**-0.014**	-0.018	–	–	–	0.007	-0.043
Mean	0.110	**0.111**	**0.111**	**0.111**	**0.111**	-0.647*	-0.422*	-0.786*	-0.646*	0.178	0.022

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

Table W36: Average Silhouette coefficients of five 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 3,600 seconds.

Dataset				1% CS							
	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS	GT
Banana	0.261	0.261	**0.280**	**0.280**	**0.280**	–	0.261	-0.118	0.103	0.389	0.032
Letter	0.113	0.113	0.113	**0.114**	0.113	0.111	0.110	–	-0.020	0.145	0.010
Shuttle	0.317	0.317	**0.320**	0.318	0.317	–	–	–	–	0.463	0.300
Cifar 10	-0.029	-0.029	**-0.010**	**-0.010**	-0.015	–	–	–	–	0.051	-0.053
Cifar 100	-0.021	-0.058	0.008	**0.009**	**0.009**	–	–	–	–	0.015	-0.114
Mnist	-0.040	-0.040	**-0.039**	-0.048	-0.048	–	–	–	–	0.007	-0.043
Mean	0.100	0.094	**0.112**	0.110	0.109	-0.815*	-0.605*	-0.853*	-0.653*	0.178	0.022

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

Table W37: Average Silhouette coefficients of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) obtained with constraint sets of size 1% 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 3,600 seconds.

Dataset	5% CS										
	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS	GT
Banana	0.032	0.032	0.032	0.032	0.032	–	–	**0.033**	0.004	0.389	0.032
Letter	-0.027	-0.027	-0.049	-0.033	-0.033	–	–	–	**-0.017**	0.145	0.010
Shuttle	0.309	0.309	**0.350**	0.318	0.309	–	–	–	–	0.463	0.300
Cifar 10	**-0.053**	**-0.053**	-0.203	-0.203	-0.145	–	–	–	–	0.051	-0.053
Cifar 100	–	–	**-0.127**	-0.136	-0.135	–	–	–	–	0.015	-0.114
Mnist	**-0.043**	**-0.043**	-0.116	-0.113	-0.116	–	–	–	–	0.007	-0.043
Mean	-0.130*	-0.130*	-0.019	-0.022	**-0.015**	-1.000*	-1.000*	-0.828*	-0.669*	0.178	0.022

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

Table W38: Average Silhouette coefficients of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) obtained with constraint sets of size 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 3,600 seconds.

Dataset	0% CS									
	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS
Banana	6.8	6.6	6.7	6.5	6.4	**1.0**	–	5,208.4	3,849.2	0.1
Letter	684.8	74.3	**74.1**	111.0	150.3	1,007.8	–	–	4,450.1	0.5
Shuttle	41.6	18.4	**18.3**	24.3	31.7	227.7	–	–	–	0.3
Cifar 10	1,956.3	380.8	**375.6**	1,267.4	1,478.0	–	–	–	–	18.0
Cifar 100	3,658.8	3,547.0	**3,546.9**	3,612.4	3,611.2	–	–	–	–	84.5
Mnist	732.6	**197.5**	199.1	372.9	465.5	3,655.5	–	–	–	4.2
Sum	7,080.9	4,224.6	**4,220.9**	5,394.6	5,743.2	12,092.0*	21,600.0*	23,208.4*	22,699.4*	107.6

*Nan values (−) are replaced with 3,600 before computing the sum.

Table W39: Average running times (in seconds) of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) for the constraint sets of size 0% CS. The lowest values are stated in bold. The column KMEANS reports the average running time of the unconstrained k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 3,600 seconds. We noticed that the LCC algorithm stops with a runtime error when the constraint set is empty. This is why the LCC algorithm did not return any solutions for the constraint sets of size 0% CS.

Dataset	0.5% CS									
	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS
Banana	**3.5**	3.6	4.1	4.2	4.1	–	94.1	4,415.0	3,622.2	0.1
Letter	908.5	190.1	**70.9**	103.6	160.8	846.9	3,642.4	–	4,459.9	0.5
Shuttle	148.5	146.7	**25.8**	31.9	50.2	–	4,926.2	–	–	0.3
Cifar 10	1,756.1	1,712.5	**382.5**	587.6	1,057.1	–	–	–	–	18.5
Cifar 100	3,810.7	3,631.2	**3,576.8**	3,624.1	3,612.5	–	–	–	–	82.3
Mnist	1,099.9	1,078.4	**241.9**	321.9	535.5	3,884.6	–	–	–	4.3
Sum	7,727.0	6,762.5	**4,302.1**	4,673.2	5,420.2	19,131.5[*]	19,462.7[*]	22,415.0[*]	22,482.1[*]	106.1

[*]Nan values (–) are replaced with 3,600 before computing the sum.

Table W40: Average running times (in seconds) of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) for the constraint sets of size 0.5% CS. The lowest values are stated in bold. The column KMEANS reports the average running time of the unconstrained k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 3,600 seconds.

Dataset	1% CS									
	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS
Banana	6.7	**6.6**	7.6	7.5	7.4	–	114.5	3,774.8	3,629.6	0.1
Letter	2,038.9	1,051.5	**115.6**	166.6	133.4	841.7	3,785.3	–	4,568.9	0.5
Shuttle	43.8	43.3	**10.0**	12.3	19.7	–	–	–	–	0.4
Cifar 10	3,935.5	3,957.2	**3,085.1**	3,742.0	3,803.6	–	–	–	–	16.7
Cifar 100	3,820.8	3,838.7	3,575.2	**1,837.6**	3,628.8	–	–	–	–	87.1
Mnist	3,785.2	3,775.2	**3,342.5**	3,728.4	3,666.3	–	–	–	–	4.3
Sum	13,630.9	12,672.4	10,136.0	**9,494.4**	11,259.2	18,841.7[*]	18,299.8[*]	21,774.8[*]	22,598.5[*]	109.0

[*]Nan values (–) are replaced with 3,600 before computing the sum.

Table W41: Average running times (in seconds) of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) for the constraint sets of size 1% CS. The lowest values are stated in bold. The column KMEANS reports the average running time of the unconstrained k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 3,600 seconds.

	5% CS									
Dataset	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	LCC	CSC	DILS	KMEANS
Banana	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	–	–	4,386.9	4,283.3	0.1
Letter	1,820.1	1,813.6	**229.5**	2,365.4	2,874.9	–	–	–	4,109.1	0.5
Shuttle	21.2	21.4	**21.0**	21.4	22.4	–	–	–	–	0.3
Cifar 10	**18.1**	**18.1**	21.8	22.3	51.6	–	–	–	–	17.5
Cifar 100	–	–	3,759.6	**3,631.8**	3,700.6	–	–	–	–	85.2
Mnist	18.2	**18.1**	20.7	21.2	22.1	–	–	–	–	4.3
Sum	5,477.9*	5,471.6*	**4,053.1**	6,062.4	6,672.0	21,600.0*	21,600.0*	22,386.9*	22,792.4*	108.0

*Nan values (–) are replaced with 3,600 before computing the sum.

Table W42: Average running times (in seconds) of five versions of the PCCC algorithm and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) for the constraint sets of size 5% CS. The lowest values are stated in bold. The column KMEANS reports the average running time of the unconstrained k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 3,600 seconds.