	0% CS										
Dataset	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	CSC	DILS	LCC	KMEANS	GT
n500-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	–	0.54	0.54
n500-k5	**0.70**	**0.70**	**0.70**	**0.70**	**0.70**	**0.70**	0.67	0.58	–	0.70	0.61
n500–k10	**0.56**	**0.56**	**0.56**	**0.56**	**0.56**	0.50	0.54	0.10	–	0.56	0.50
n500-k20	**0.39**	**0.39**	**0.39**	**0.39**	**0.39**	0.38	0.36	-0.17	–	0.39	0.21
n500-k50	**0.37**	**0.37**	**0.37**	**0.37**	**0.37**	**0.37**	0.34	-0.45	–	0.37	0.07
n500-k100	**0.37**	**0.37**	**0.37**	**0.37**	**0.37**	0.34	0.33	-0.65	–	0.37	-0.11
n1000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	–	0.54	0.54
n1000-k5	0.66	0.66	0.66	0.66	0.66	**0.69**	0.66	0.38	–	0.66	0.60
n1000-k10	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	0.54	**0.55**	-0.01	–	0.55	0.51
n1000-k20	**0.39**	**0.39**	**0.39**	**0.39**	**0.39**	0.38	0.35	-0.19	–	0.39	0.24
n1000-k50	**0.36**	**0.36**	**0.36**	**0.36**	**0.36**	**0.36**	0.34	-0.36	–	0.36	0.10
n1000-k100	**0.37**	**0.37**	**0.37**	**0.37**	**0.37**	0.35	0.34	-0.47	–	0.37	-0.06
n2000-k2	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	–	0.55	0.55
n2000-k5	**0.70**	**0.70**	**0.70**	**0.70**	**0.70**	0.67	0.67	0.08	–	0.70	0.60
n2000-k10	0.55	0.55	0.55	0.55	0.55	0.54	**0.56**	-0.12	–	0.55	0.51
n2000-k20	**0.39**	**0.39**	**0.39**	**0.39**	**0.39**	0.38	0.35	-0.20	–	0.39	0.24
n2000-k50	0.35	0.35	0.35	0.35	0.35	**0.36**	0.34	-0.29	–	0.35	0.10
n2000-k100	**0.34**	**0.34**	**0.34**	**0.34**	**0.34**	0.33	**0.34**	-0.35	–	0.34	-0.05
n5000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	0.33	–	0.54	0.54
n5000-k5	0.67	0.67	0.67	0.67	0.67	**0.68**	0.67	-0.10	–	0.67	0.60
n5000-k10	**0.57**	**0.57**	**0.57**	**0.57**	**0.57**	0.54	0.55	-0.13	–	0.57	0.51
n5000-k20	**0.38**	**0.38**	**0.38**	**0.38**	**0.38**	**0.38**	0.34	-0.14	–	0.38	0.23
n5000-k50	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	0.34	0.32	-0.16	–	0.35	0.10
n5000-k100	**0.34**	**0.34**	**0.34**	**0.34**	**0.34**	0.33	0.32	-0.23	–	0.34	-0.05
Mean	**0.48**	**0.48**	**0.48**	**0.48**	**0.48**	0.47	0.46	-0.04	-1.00*	0.48	0.32

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

Table W21: 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.

	5% CS										
Dataset	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	CSC	DILS	LCC	KMEANS	GT
n500-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	–	-0.15	0.51	**0.54**	0.54	0.54
n500-k5	0.65	0.65	**0.66**	**0.66**	0.64	0.60	-0.17	0.32	0.49	0.70	0.61
n500–k10	0.53	0.53	**0.56**	**0.56**	0.55	0.45	-0.41	-0.07	0.52	0.56	0.50
n500-k20	**0.39**	**0.39**	**0.39**	**0.39**	**0.39**	0.36	-0.53	-0.21	0.35	0.39	0.21
n500-k50	**0.37**	**0.37**	**0.37**	**0.37**	**0.37**	0.36	-0.72	-0.45	0.35	0.37	0.07
n500-k100	**0.34**	**0.34**	**0.34**	**0.34**	**0.34**	0.32	-0.73	-0.64	0.31	0.37	-0.11
n1000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	0.53	**0.54**	0.54	0.54
n1000-k5	**0.63**	**0.63**	**0.63**	**0.63**	**0.63**	–	-0.45	0.13	0.59	0.66	0.60
n1000-k10	0.53	0.53	**0.55**	**0.55**	0.54	0.50	-0.34	-0.11	0.44	0.55	0.51
n1000-k20	**0.36**	**0.36**	**0.36**	**0.36**	**0.36**	0.35	-0.37	-0.23	0.33	0.39	0.24
n1000-k50	**0.34**	**0.34**	**0.34**	**0.34**	**0.34**	0.32	-0.70	-0.33	**0.34**	0.36	0.10
n1000-k100	**0.36**	**0.36**	**0.36**	**0.36**	**0.36**	0.34	-0.77	-0.47	–	0.37	-0.06
n2000-k2	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	0.29	0.55	0.55
n2000-k5	**0.62**	**0.62**	**0.62**	**0.62**	**0.62**	**0.62**	-0.23	-0.09	0.51	0.70	0.60
n2000-k10	0.48	0.48	**0.56**	0.55	0.53	0.49	-0.31	-0.16	0.48	0.55	0.51
n2000-k20	**0.32**	**0.32**	**0.32**	**0.32**	**0.32**	0.29	-0.30	-0.19	**0.32**	0.39	0.24
n2000-k50	0.31	0.31	0.31	0.31	0.31	0.30	-0.57	-0.26	**0.32**	0.35	0.10
n2000-k100	**0.31**	**0.31**	**0.31**	0.30	**0.31**	0.30	-0.78	-0.33	0.24	0.34	-0.05
n5000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	–	**0.54**	0.25	0.51	0.54	0.54
n5000-k5	**0.61**	**0.61**	**0.61**	**0.61**	**0.61**	0.59	0.14	-0.06	0.54	0.67	0.60
n5000-k10	**0.51**	**0.51**	**0.51**	**0.51**	**0.51**	–	-0.10	-0.06	0.07	0.57	0.51
n5000-k20	0.27	0.27	**0.28**	0.26	0.27	–	-0.13	-0.10	0.18	0.38	0.23
n5000-k50	0.26	0.26	**0.27**	0.26	0.26	0.22	-0.22	-0.14	0.25	0.35	0.10
n5000-k100	**0.26**	**0.26**	**0.26**	**0.26**	**0.26**	0.24	-0.54	-0.20	0.25	0.34	-0.05
Mean	0.44	0.44	**0.45**	**0.45**	0.44	0.11*	-0.28	-0.08	0.32*	0.48	0.32

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

Table W22: 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.

	10% CS										
Dataset	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	CSC	DILS	LCC	KMEANS	GT
n500-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	0.30	0.54	0.54
n500-k5	**0.62**	**0.62**	**0.62**	**0.62**	**0.62**	–	-0.24	0.42	**0.62**	0.70	0.61
n500–k10	0.50	0.50	**0.53**	**0.53**	**0.53**	0.44	-0.34	-0.06	0.40	0.56	0.50
n500-k20	0.33	0.33	**0.34**	0.33	0.33	0.30	-0.43	-0.26	0.29	0.39	0.21
n500-k50	**0.33**	**0.33**	**0.33**	**0.33**	**0.33**	0.29	-0.67	-0.42	0.27	0.37	0.07
n500-k100	**0.34**	**0.34**	**0.34**	**0.34**	**0.34**	0.30	-0.74	-0.65	0.27	0.37	-0.11
n1000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	0.33	0.54	0.54
n1000-k5	0.60	0.60	**0.62**	0.60	0.60	0.58	-0.16	0.29	0.53	0.66	0.60
n1000-k10	0.52	0.52	**0.53**	0.52	0.52	–	-0.16	-0.16	0.38	0.55	0.51
n1000-k20	0.28	0.28	**0.29**	**0.29**	0.28	–	-0.33	-0.22	0.20	0.39	0.24
n1000-k50	0.28	0.28	**0.29**	0.28	0.28	0.26	-0.64	-0.33	0.26	0.36	0.10
n1000-k100	0.29	0.29	**0.30**	0.29	0.29	0.27	-0.73	-0.44	0.26	0.37	-0.06
n2000-k2	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	0.55	0.55
n2000-k5	0.61	0.61	0.61	0.61	0.61	0.61	**0.62**	-0.03	0.57	0.70	0.60
n2000-k10	0.51	0.51	**0.52**	**0.52**	0.51	–	-0.22	-0.12	0.18	0.55	0.51
n2000-k20	**0.27**	**0.27**	0.26	0.25	0.24	–	-0.17	-0.18	0.10	0.39	0.24
n2000-k50	0.21	0.21	**0.24**	0.22	0.22	0.17	-0.32	-0.24	0.21	0.35	0.10
n2000-k100	0.23	0.23	**0.24**	0.23	0.23	0.19	-0.60	-0.32	0.22	0.34	-0.05
n5000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	–	**0.54**	0.24	–	0.54	0.54
n5000-k5	0.60	0.60	0.60	0.60	0.60	–	**0.63**	-0.04	–	0.67	0.60
n5000-k10	**0.51**	**0.51**	0.50	**0.51**	**0.51**	–	-0.18	-0.05	–	0.57	0.51
n5000-k20	**0.24**	**0.24**	0.23	0.23	0.22	–	-0.18	-0.07	–	0.38	0.23
n5000-k50	0.03	0.03	**0.15**	0.14	0.14	–	-0.24	-0.12	0.02	0.35	0.10
n5000-k100	0.03	0.03	**0.12**	0.10	0.10	0.03	-0.44	-0.18	0.08	0.34	-0.05
Mean	0.40	0.40	**0.41**	0.40	0.40	-0.21*	-0.14	-0.05	0.08*	0.48	0.32

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

Table W23: 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 10% 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.

	15% CS										
Dataset	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	CSC	DILS	LCC	KMEANS	GT
n500-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	0.54	0.54
n500-k5	0.62	0.62	**0.64**	0.62	0.62	0.62	0.05	0.50	0.56	0.70	0.61
n500–k10	0.50	0.50	**0.53**	**0.53**	0.50	–	-0.22	-0.04	0.49	0.56	0.50
n500-k20	**0.26**	**0.26**	**0.26**	**0.26**	**0.26**	0.18	-0.29	-0.24	0.22	0.39	0.21
n500-k50	0.27	0.27	**0.28**	**0.28**	**0.28**	0.25	-0.64	-0.45	0.25	0.37	0.07
n500-k100	0.28	0.28	**0.29**	0.28	0.28	0.25	-0.72	-0.66	0.25	0.37	-0.11
n1000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	–	0.54	0.54
n1000-k5	0.60	0.60	0.62	0.60	0.60	0.60	**0.63**	0.31	0.59	0.66	0.60
n1000-k10	**0.51**	**0.51**	**0.51**	**0.51**	**0.51**	**0.51**	-0.23	-0.16	0.29	0.55	0.51
n1000-k20	**0.26**	**0.26**	0.25	0.25	0.24	–	-0.23	-0.21	0.06	0.39	0.24
n1000-k50	0.20	0.20	**0.22**	0.21	0.21	0.12	-0.50	-0.33	0.19	0.36	0.10
n1000-k100	0.24	0.24	**0.25**	0.24	0.24	0.19	-0.69	-0.49	0.22	0.37	-0.06
n2000-k2	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	–	**0.55**	**0.55**	–	0.55	0.55
n2000-k5	0.60	0.60	0.60	0.60	0.60	0.60	**0.64**	-0.01	0.60	0.70	0.60
n2000-k10	**0.51**	**0.51**	**0.51**	**0.51**	**0.51**	**0.51**	0.02	-0.13	–	0.55	0.51
n2000-k20	**0.25**	**0.25**	**0.25**	0.24	**0.25**	0.23	-0.19	-0.14	0.02	0.39	0.24
n2000-k50	-0.05	-0.05	**0.15**	0.14	0.14	–	-0.34	-0.20	0.06	0.35	0.10
n2000-k100	-0.30	-0.30	**0.14**	**0.14**	0.13	0.07	-0.55	-0.32	0.11	0.34	-0.05
n5000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	–	**0.54**	0.24	–	0.54	0.54
n5000-k5	0.60	0.60	0.60	0.60	0.60	–	**0.63**	-0.06	–	0.67	0.60
n5000-k10	**0.51**	**0.51**	**0.51**	**0.51**	**0.51**	–	0.25	-0.06	–	0.57	0.51
n5000-k20	**0.23**	**0.23**	0.21	0.21	**0.23**	**0.23**	-0.37	-0.06	–	0.38	0.23
n5000-k50	-0.06	-0.06	**0.11**	0.10	0.10	–	-0.29	-0.12	–	0.35	0.10
n5000-k100	-0.19	-0.19	**0.01**	-0.00	-0.01	–	-0.45	-0.19	-0.08	0.34	-0.05
Mean	0.33	0.33	**0.38**	**0.38**	0.37	-0.15*	-0.06	-0.05	-0.15*	0.48	0.32

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

Table W24: 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 15% 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.

	20% CS										
Dataset	PCCC	PCCC-N2	PCCC-N2-S	PCCC-N3-S	PCCC-N5-S	COPKM	CSC	DILS	LCC	KMEANS	GT
n500-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	0.54	0.54
n500-k5	**0.61**	**0.61**	**0.61**	**0.61**	**0.61**	**0.61**	0.58	0.51	**0.61**	0.70	0.61
n500–k10	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	-0.21	0.04	0.29	0.56	0.50
n500-k20	**0.26**	**0.26**	**0.26**	**0.26**	0.25	–	-0.34	-0.26	0.15	0.39	0.21
n500-k50	0.22	0.22	**0.23**	**0.23**	0.22	0.16	-0.59	-0.43	0.19	0.37	0.07
n500-k100	0.23	0.23	**0.24**	0.23	0.23	0.18	-0.70	-0.63	0.19	0.37	-0.11
n1000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	–	0.54	0.54
n1000-k5	0.60	0.60	0.60	0.60	0.60	0.60	**0.64**	0.30	0.60	0.66	0.60
n1000-k10	**0.51**	**0.51**	**0.51**	**0.51**	**0.51**	**0.51**	-0.06	-0.11	0.38	0.55	0.51
n1000-k20	**0.25**	**0.25**	0.24	0.23	0.24	–	-0.24	-0.20	0.08	0.39	0.24
n1000-k50	0.15	0.15	**0.17**	0.16	0.15	–	-0.45	-0.30	0.10	0.36	0.10
n1000-k100	0.16	0.16	**0.18**	0.17	0.16	0.11	-0.64	-0.45	0.13	0.37	-0.06
n2000-k2	**0.55**	**0.55**	**0.55**	**0.55**	**0.55**	–	**0.55**	**0.55**	–	0.55	0.55
n2000-k5	0.60	0.60	0.60	0.60	0.60	0.60	**0.64**	-0.04	–	0.70	0.60
n2000-k10	0.51	0.51	**0.55**	0.51	0.51	0.51	0.33	-0.15	0.44	0.55	0.51
n2000-k20	**0.24**	**0.24**	0.23	0.23	**0.24**	**0.24**	-0.27	-0.13	0.04	0.39	0.24
n2000-k50	-0.18	-0.18	**0.11**	0.10	0.09	–	-0.30	-0.21	-0.01	0.35	0.10
n2000-k100	-0.09	-0.09	**0.06**	0.05	0.04	-0.04	-0.57	-0.30	0.01	0.34	-0.05
n5000-k2	**0.54**	**0.54**	**0.54**	**0.54**	**0.54**	–	**0.54**	0.25	–	0.54	0.54
n5000-k5	0.60	0.60	0.60	0.60	0.60	–	**0.62**	-0.07	–	0.67	0.60
n5000-k10	**0.51**	**0.51**	**0.51**	**0.51**	**0.51**	–	0.26	-0.06	–	0.57	0.51
n5000-k20	**0.23**	**0.23**	**0.23**	**0.23**	**0.23**	–	-0.28	-0.07	–	0.38	0.23
n5000-k50	0.10	0.10	0.10	0.10	**0.11**	0.10	-0.30	-0.12	–	0.35	0.10
n5000-k100	–	–	**-0.03**	-0.04	-0.04	–	-0.43	-0.16	–	0.34	-0.05
Mean	0.30*	0.30*	**0.36**	**0.36**	**0.36**	-0.20*	-0.01	-0.04	-0.22*	0.48	0.32

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

Table W25: 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 20% 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.