

Dataset	10% CS					
	PCCC	PCCC-N2-S-RD	COPKM	CSC	DILS	LCC
n1000-k10	0.976	0.976	–	0.011	0.086	0.771
n1000-k100	0.270	0.271	0.251	0.000	0.007	0.252
n1000-k2	1.000	1.000	1.000	1.000	1.000	0.667
n1000-k20	0.700	0.755	–	0.003	0.036	0.499
n1000-k5	0.993	0.993	0.964	0.201	0.640	0.877
n1000-k50	0.444	0.444	0.399	0.001	0.018	0.409
n2000-k10	0.986	0.986	–	0.027	0.012	0.644
n2000-k100	0.014	0.269	0.253	0.002	0.002	0.273
n2000-k2	1.000	1.000	1.000	1.000	1.000	1.000
n2000-k20	0.902	0.902	–	0.010	0.009	0.515
n2000-k5	0.999	0.999	0.999	0.708	0.202	0.938
n2000-k50	0.479	0.513	0.415	0.005	0.006	0.459
n5000-k10	0.999	0.999	–	0.325	0.001	–
n5000-k100	0.297	0.335	0.269	0.005	0.000	0.294
n5000-k2	1.000	1.000	–	1.000	0.401	–
n5000-k20	0.982	0.982	–	0.032	0.001	–
n5000-k5	1.000	1.000	–	0.731	0.003	–
n5000-k50	0.553	0.777	–	0.009	0.001	0.411
n500-k10	0.948	0.948	0.813	0.006	0.242	0.766
n500-k100	0.212	0.212	0.215	0.004	0.008	0.223
n500-k2	1.000	1.000	1.000	1.000	1.000	0.533
n500-k20	0.570	0.642	0.542	0.001	0.080	0.518
n500-k5	0.971	0.971	–	0.034	0.772	0.922
n500-k50	0.386	0.395	0.381	0.004	0.030	0.397
Mean	0.737	0.765	0.354*	0.255	0.231	0.474*

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

Table W77: Average Adjusted Rand Index (ARI) values 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 10% 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.