	20% CS								
Dataset	PCCC	PCCC-N2-S	PCCC-N5-S	PCCC-N2-S-RD	COPKM	CSC	DILS	LCC	KMEANS
n500-k2	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	0.968
n500-k5	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	0.682	0.896	0.992	0.773
n500–k10	**0.982**	**0.982**	**0.982**	**0.982**	0.973	0.066	0.440	0.725	0.874
n500-k20	0.828	0.777	0.832	**0.846**	–	0.024	0.083	0.606	0.563
n500-k50	0.499	0.494	0.499	**0.513**	0.447	0.007	0.036	0.432	0.376
n500-k100	0.272	**0.278**	0.272	0.272	0.239	0.005	0.021	0.263	0.221
n1000-k2	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	–	0.984
n1000-k5	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	0.727	0.667	**1.000**	0.860
n1000-k10	**0.996**	**0.996**	**0.996**	**0.996**	**0.996**	0.356	0.147	0.870	0.801
n1000-k20	**0.958**	0.891	**0.958**	0.957	–	0.026	0.018	0.631	0.574
n1000-k50	0.558	0.636	0.679	**0.685**	–	0.011	0.011	0.463	0.398
n1000-k100	0.064	0.317	**0.324**	0.320	0.296	0.007	0.007	0.299	0.238
n2000-k2	**1.000**	**1.000**	**1.000**	**1.000**	–	**1.000**	**1.000**	–	0.963
n2000-k5	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	0.727	0.212	–	0.770
n2000-k10	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	0.758	0.023	0.930	0.835
n2000-k20	**0.998**	0.893	**0.998**	**0.998**	0.996	0.045	0.005	0.716	0.589
n2000-k50	**0.891**	0.734	0.857	0.864	–	0.020	0.003	0.531	0.406
n2000-k100	0.414	0.446	**0.505**	0.498	0.382	0.008	0.002	0.352	0.246
n5000-k2	**1.000**	**1.000**	**1.000**	**1.000**	–	**1.000**	0.406	–	0.985
n5000-k5	**1.000**	**1.000**	**1.000**	**1.000**	–	0.821	0.011	–	0.866
n5000-k10	**1.000**	**1.000**	**1.000**	**1.000**	–	0.753	0.002	–	0.855
n5000-k20	**1.000**	0.934	**1.000**	**1.000**	–	0.241	0.001	–	0.593
n5000-k50	**0.993**	0.915	**0.993**	0.992	0.988	0.025	0.000	–	0.421
n5000-k100	0.795	0.731	0.824	**0.834**	–	0.018	0.000	–	0.248
Mean	0.844	0.834	0.863	**0.865**	0.472*	0.388	0.250	0.409*	0.642

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

Table W79: 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 20% 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.