Dataset	15% CS						KMEANS
	PCCC	PCCC-R	COPKM	CSC	DILS	LCC	
Appendicitis	**1.000**	**1.000**	–	0.971	0.732	0.045	0.330
Breast Cancer	**1.000**	**1.000**	**1.000**	0.005	**1.000**	**1.000**	0.675
Bupa	**1.000**	**1.000**	**1.000**	-0.007	**1.000**	–	-0.005
Circles	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	-0.003
Ecoli	**0.929**	**0.929**	–	0.239	0.627	0.799	0.387
Glass	0.764	**0.791**	–	0.006	0.375	0.446	0.239
Haberman	**1.000**	**1.000**	**1.000**	0.012	**1.000**	0.959	0.094
Hayesroth	**0.978**	**0.978**	–	0.107	0.514	0.439	0.062
Heart	**1.000**	**1.000**	**1.000**	0.646	**1.000**	0.835	0.407
Ionosphere	**1.000**	**1.000**	**1.000**	0.078	**1.000**	0.504	0.170
Iris	**0.625**	0.613	0.585	0.564	0.566	0.618	0.626
Led7Digit	0.823	**0.884**	–	0.007	0.091	–	0.429
Monk2	**1.000**	**1.000**	**1.000**	0.972	**1.000**	**1.000**	0.051
Moons	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	0.838	0.470
Movement Libras	0.324	**0.336**	0.299	0.017	0.067	0.321	0.307
Newthyroid	**1.000**	**1.000**	**1.000**	0.560	0.820	0.681	0.593
Saheart	**1.000**	**1.000**	**1.000**	0.784	**1.000**	–	0.072
Sonar	**1.000**	**1.000**	–	-0.002	**1.000**	0.782	0.005
Soybean	**1.000**	**1.000**	0.816	0.031	0.552	0.774	0.896
Spectfheart	**1.000**	**1.000**	0.983	-0.043	**1.000**	**1.000**	-0.103
Spiral	**1.000**	**1.000**	–	0.973	**1.000**	0.926	0.027
Tae	**0.713**	**0.713**	–	0.203	0.592	0.164	0.035
Vehicle	**1.000**	**1.000**	**1.000**	0.635	0.813	–	0.084
Wine	**0.933**	**0.933**	**0.933**	0.543	0.584	0.783	0.852
Zoo	0.888	0.889	0.754	0.333	0.707	**0.923**	0.794
Mean	0.919	**0.923**	0.615*	0.385	0.762	0.593*	0.300

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

Table W7: Average Adjusted Rand Index (ARI) values of the PCCC and the PCCC-R algorithms and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) obtained with constraint sets of size 15% 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 1,800 seconds.