Dataset	Objects	Features	Clusters	5% CS											
				PCCC-N2-S	PCCC-N3-S	PCCC-N4-S	PCCC-N5-S	PCCC-N6-S	PCCC-N2-S-RD	COPKM	LCC	CSC	DILS	KMEANS	
Banana	5,300	2	2	**0**	**0**	**0**	**0**	**0**	**0**	–	–	**0**	–	–	
Letter	20,000	16	26	10,804	8,297	9,869	7,166	6,656	**4,035**	–	–	–	–	–	
Shuttle	57,999	9	7	3,722	1,922	24	**0**	**0**	**0**	–	–	–	–	–	
CIFAR-10	60,000	3,072	10	629,094	448,349	447,715	357,261	267,887	**0**	–	–	–	–	–	
CIFAR-100	60,000	3,072	100	81,778	122,448	116,611	105,542	95,356	**49,941**	–	–	–	–	–	
MNIST	70,000	784	10	355,448	343,835	343,749	343,727	341,816	**0**	–	–	–	–	–	
Mean				180,141	154,142	152,995	135,616	118,619	8,996	–*	–*	**0***	–*	–	

*Nan values (–) are ignored when computing the sum.

Table W115: Average number of cannot-link constraint violations of the 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 number of cannot-link constraint violations obtained with the k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 1,800 seconds.