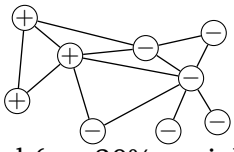
1 Experimental settings

Let's consider a graph with binary labels on its node such as the following one: 

Once I set a training set size (say 30% in that case) and a perturbation level (say 30% again), there are 3 independent dimensions along which I have to make decision uniformly at random before starting to predict node signs:

- A spanning tree, whether random or with minimum weight. Its edges are denoted by thick red lines ——
- Which node are part of the training set, in blue ⊕
- Which node have their sign flipped, with the ± sign ±

Hence prior to the any experiments, I start by choosing random spanning trees

$$\{T_1 = \quad , T_2 = \quad , \ldots, T_{10} = \quad\}$$

10 training sets

$$\{Tr_1 = \quad , Tr_2 = \quad , \ldots, Tr_{10} = \quad\}$$

and 10 perturbations

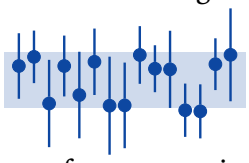
$$\{P_1 = \quad , P_2 = \quad , \ldots, P_{10} = \quad\}$$

Note that the perturbations are independent of the choice of the training set, because they are only defined by the indices (within the training set) of the nodes whose sign is flipped.

The goal of the experiments is to study the variance of our algorithms with respect to those choices, in the single tree case, i.e. when we don't aggregate over several spanning trees. I do that by varying one dimension while keeping the two others fixed. For instance, I change the training set while keeping the first spanning tree T_1 and the first perturbation P_1, resulting in the following experimental instances

$$\{E_1 = \quad , E_2 = \quad , \ldots, E_{10} = \quad\}$$

Running SHAZOO on all those instances produces 10 performance results, which I aggregate and display as |●|, where the circle is the average performance and the error bar represent one standard deviation in each directions. Likewise I plot the performance of SHAZOO when using Minimum

Spanning Tree[1] instead of random one[2]  , WTA+RST  , WTA+MST and LAB. PROP  (which doesn't use tree). In the case of RTA, we need to keep track of another source of randomness, namely the order in which the training set is presented to the algorithm. This also adds a question regarding the influence of the number of such presentations and therefore requires a different graphical representation. For each experimental instance E_i, I run RTA with 117 different presentation orders and record the resulting 117 predictions of the test set. Then I choose a number of orders $n_o \in [11, 33, 57, 79, 101]$ and pick n_o orders out of the 117 possible. I aggregate these n_o predictions with majority vote and as before, I represent the results over the varying dimension by  , and  when using MST. In addition this time, I repeat the choice of n_o orders 15 times, giving 15 such error bars representing the variance along the random choice of training set presentation order. The overall performance is summarized by a lighter box of the same color, which is centered at the average of the 15 results and whose height is twice the average of the standard deviation 

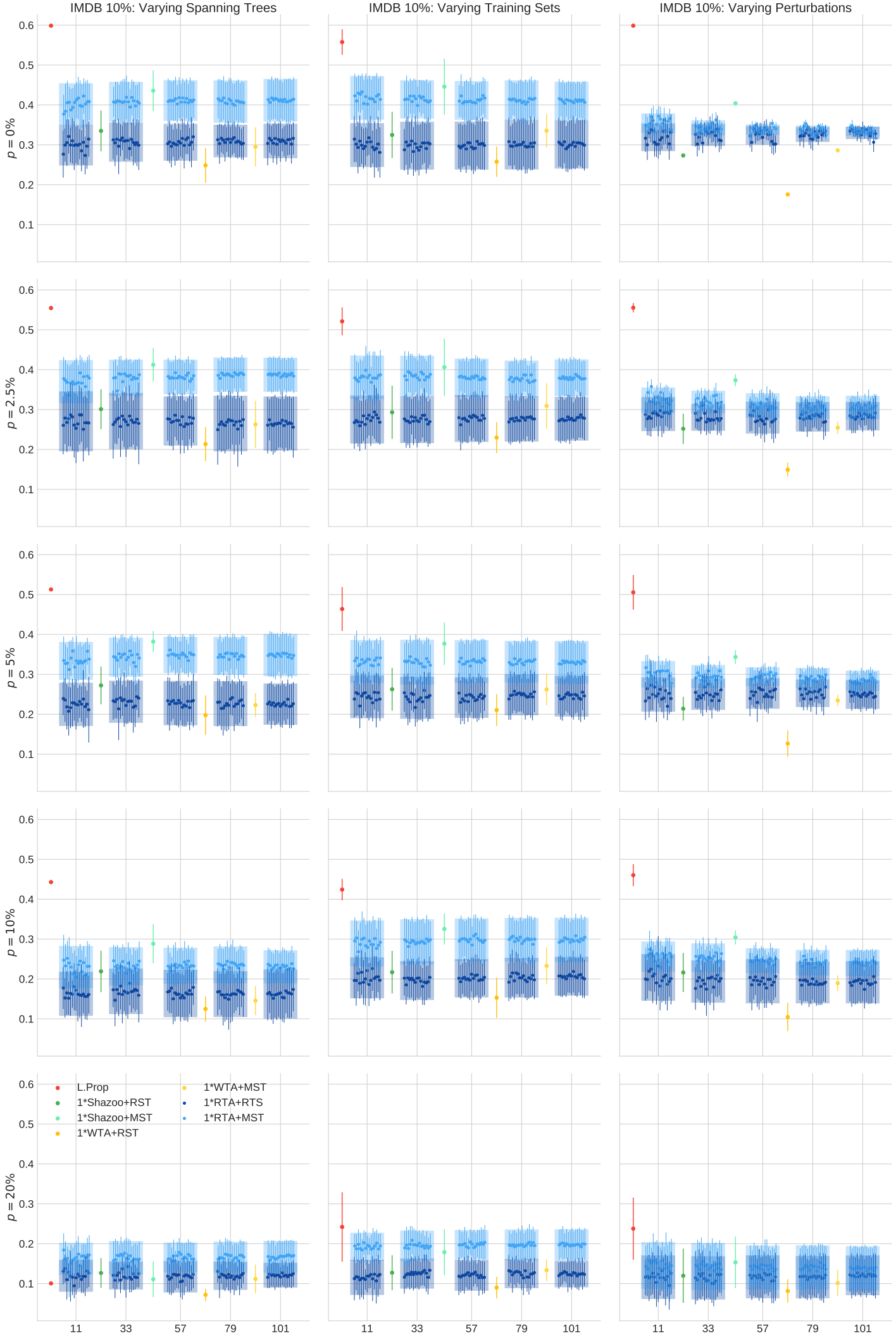
As showed in Table 1, not all datasets are balanced and therefore I measure performance using Matthews Correlation Coefficient (MCC) ranging from -1 (when all predictions are incorrect) to $+1$ (when all predictions are correct).

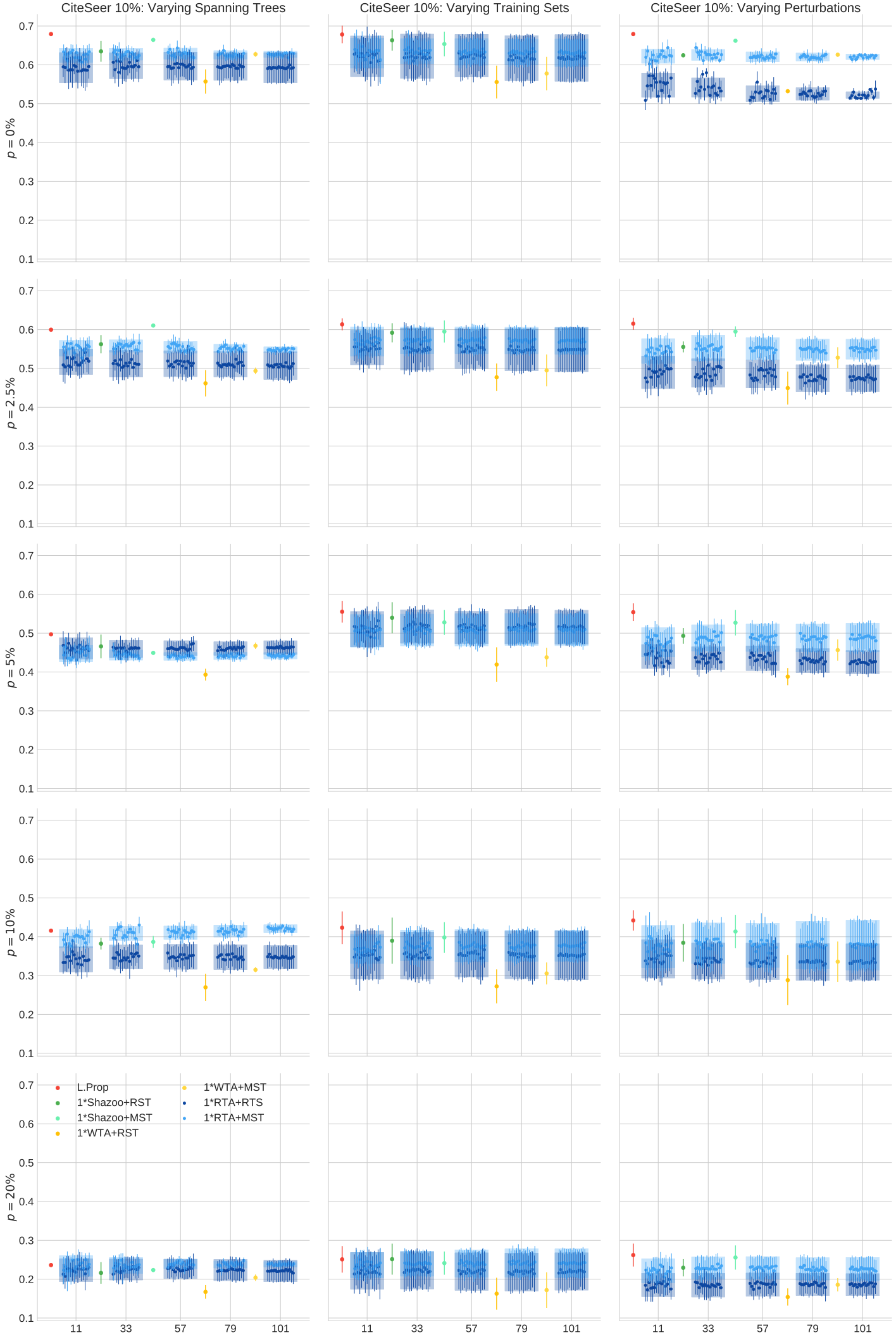
Table 1: Datasets statistics

Dataset	$\lvert V\rvert$	$\lvert E\rvert$	number of classes	chosen class	$+1$ fraction
IMDB	1,126	20,282	2	0	50.2%
CITESEER	2,110	3,668	6	1	21.9%
CORA	2,484	5,068	7	2	29.2%
PUBMED	4,201	21,042	3	1	45.5%
USPS-10	4,500	33,121	10	4	9.2%
RCV1-10	4,500	32,715	4	2	46.4%

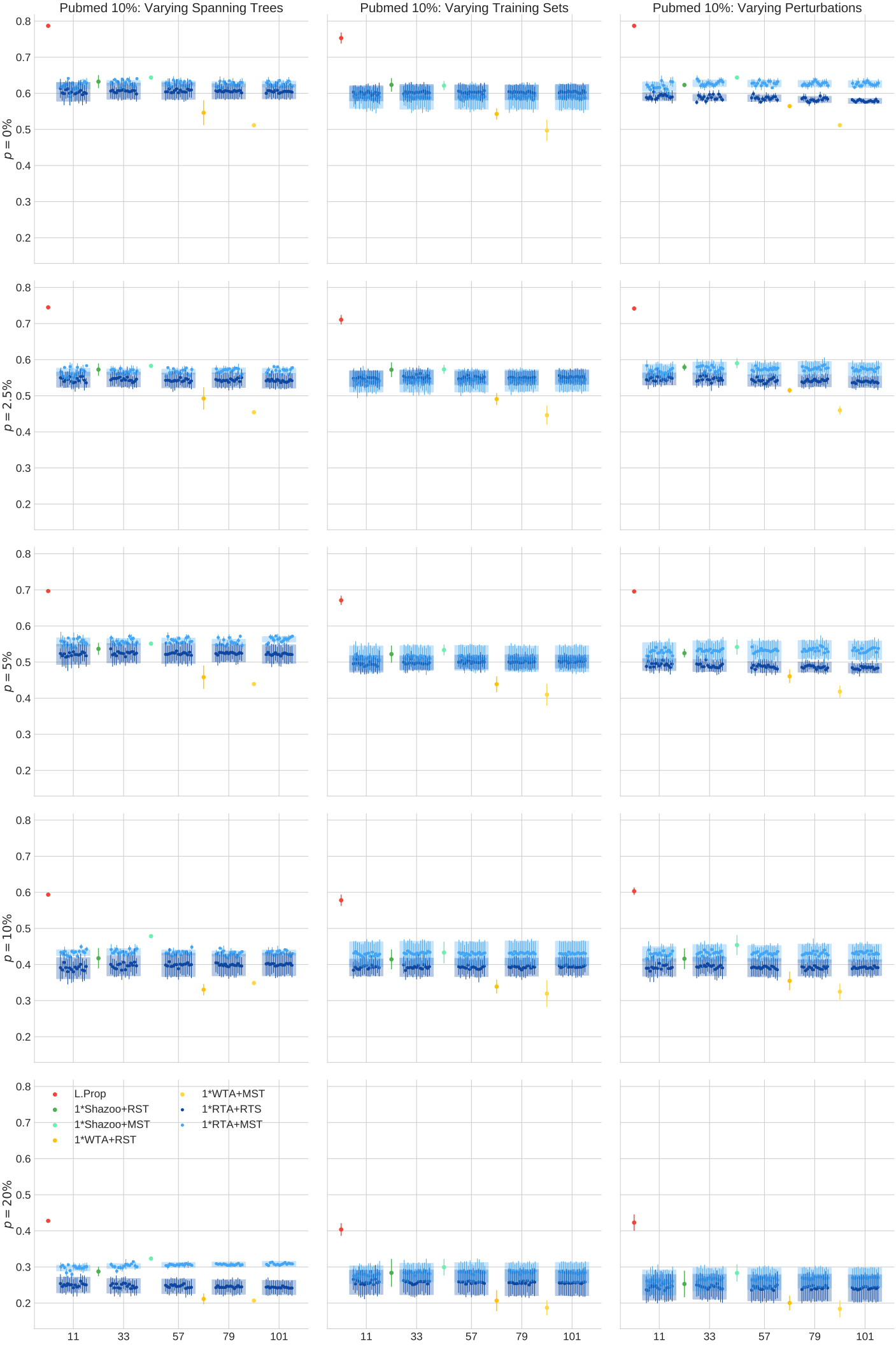
In the following plots, the three columns correspond to varying spanning trees, training sets and perturbations from left to right respectively, while the rows are ordered in increasing level perturbations, from 0 to 20%. I set the training size to 10% in the plots from pages 3 to 8 (and to 5% in the plots from pages 16 to 22).

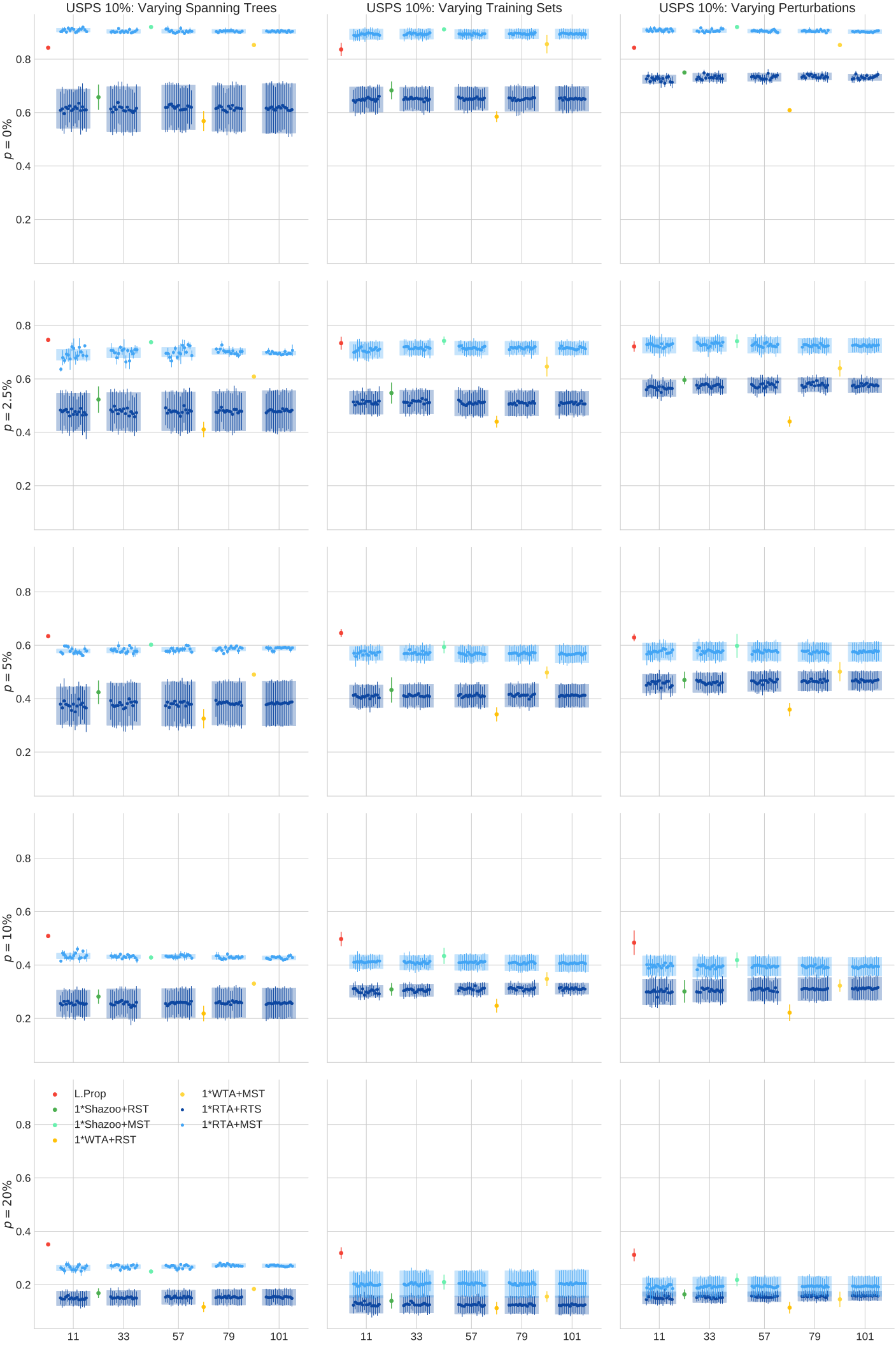
[1] in the sense of resistance distance.
[2] I obtain different MST by permuting the edges with equal weight.

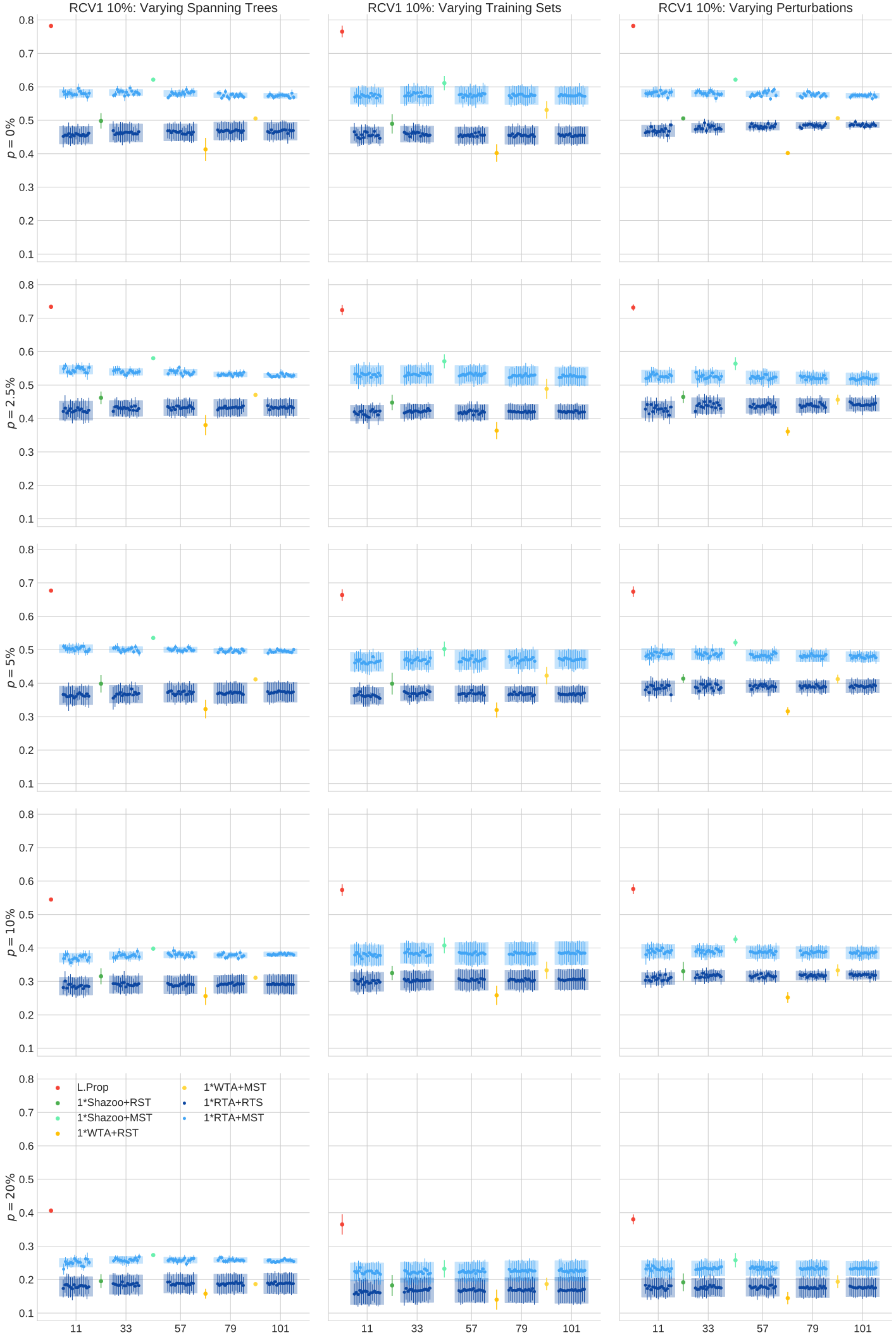

IMDB 10%: Varying Spanning Trees
IMDB 10%: Varying Training Sets
IMDB 10%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.6
0.5
0.4
0.3
0.2
0.1
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


CiteSeer 10%: Varying Spanning Trees
CiteSeer 10%: Varying Training Sets
CiteSeer 10%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.7
0.6
0.5
0.4
0.3
0.2
0.1
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


Cora 10%: Varying Spanning Trees
Cora 10%: Varying Training Sets
Cora 10%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.1
0.2
0.3
0.4
0.5
0.6
0.7
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


Pubmed 10%: Varying Spanning Trees
Pubmed 10%: Varying Training Sets
Pubmed 10%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.8
0.7
0.6
0.5
0.4
0.3
0.2
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


USPS 10%: Varying Spanning Trees
USPS 10%: Varying Training Sets
USPS 10%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.2
0.4
0.6
0.8
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


RCV1 10%: Varying Spanning Trees
RCV1 10%: Varying Training Sets
RCV1 10%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST

2 Experiments analysis

The previous plots are very dense and a bit difficult to read. Here we break them down in four questions. Furthermore, we focus on the 10% training set size, for the results look similar at 5%.

2.1 Variance of all methods

Table 2: 100 times the MCC standard deviation of each method as the perturbation level increases from 0% to 20%.

dataset	IMDB			CITESEER			CORA			PUBMED			USPS-10			RCV1-10			average
dimension	tree	train	flip	tree	train	flip	tree	train	flip	tree	train	flip	tree	train	flip	tree	train	flip	
SHAZOO+RST	5.1	3.7 5.8	4.4 6.8	2.7	2.8 2.7	4.0 0.0	2.2	3.0 2.8	2.4 2.1	0.0 3.6	1.8 1.3	1.9 3.9	0.0 3.7	4.8 1.8	3.4 2.9	0.0 1.8	2.3 2.1	2.9 3.1	0.0 2.7	2.9
SHAZOO+MST	5.1	4.5 7.0	5.7 6.5	0.0 0.1	0.3 3.2	3.0 0.0	3.1	2.5 1.1	3.2 2.1	0.0 4.3	0.0 0.0	1.3 2.3	0.0 2.4	0.0 0.0	0.7 2.8	0.0 2.4	0.0 0.0	2.1 2.6	0.0 2.2	2.1
RTA+RST	4.5	3.6 5.9	3.8 5.2	2.3 3.6	2.7 5.6	4.8 2.0	2.9	2.6 3.5	3.9 2.9	1.2 2.8	2.3 2.0	2.1 3.5	0.9 3.6	8.4 2.8	4.5 3.3	1.4 1.9	2.6 2.9	2.5 3.7	1.2 2.8	3.5
RTA+MST	5.1	3.9 4.9	3.7 5.6	1.5 1.3	1.5 3.5	3.5 1.2	2.9	2.3 2.4	5.3 4.3	0.9 3.7	0.8 0.5	3.3 2.6	1.0 3.0	0.4 0.6	1.8 4.9	0.6 3.6	0.9 0.9	2.6 2.9	0.9 2.3	2.8
WTA+RST	4.4	1.5 3.8	2.7 2.9	0.0 3.1	1.8 4.2	4.1 0.0	2.2	3.1 2.2	3.5 1.7	0.0 3.0	3.5 1.5	1.6 2.9	0.0 2.0	3.8 1.9	2.1 2.4	0.0 2.2	3.4 1.5	2.6 3.0	0.0 1.8	2.6
WTA+MST	4.9	3.6 4.2	2.7 3.3	0.0 0.6	0.8 4.3	4.6 0.0	1.7	1.8 3.0	1.6 1.6	0.0 0.0	0.0 2.9	2.1 2.4	0.0 0.0	0.0 3.4	2.0 2.9	0.0 0.1	0.1 2.6	1.9 1.9	0.0 1.9	2.0

In Table 2, we see that the variance of the MCC is comparable across methods, datasets and perturbation levels. Actually for a given dataset and dimension, we can see that the shape of the curve is very similar for SHAZOO+RST, SHAZOO+MST, RTA+RST and RTA+MST Furthermore, there is no clear pattern than emerges.

2.2 Random vs Minimum Spanning Trees

As showed in Tables 3 and 4, for all three algorithms, MST usually gives better MCC on average and with slightly less variance (except on the CORA dataset) and therefore will be preferred in the following.

Table 3: 100 times the difference of MCC between MST and RST.

		0% perturbation		2.5% perturbation		5% perturbation		10% perturbation		20% perturbation		average	average
method	dataset	average	better	average	better	average	better	average	better	average	better	per dataset	overall
SHAZOO	IMDB	11.76	100.0%	11.53	100.0%	11.79	100.0%	8.85	100.0%	2.31	66.7%	9.25	9.25
	CITESEER	1.91	66.7%	3.02	100.0%	0.17	33.3%	1.40	100.0%	0.79	66.7%	1.46	
	CORA	-4.31	0.0%	-5.13	0.0%	-5.19	0.0%	-3.06	33.3%	-1.71	33.3%	-3.88	
	PUBMED	1.00	66.7%	0.73	100.0%	1.44	100.0%	3.93	100.0%	2.72	100.0%	1.96	
	USPS-10	22.05	100.0%	18.52	100.0%	15.56	100.0%	12.96	100.0%	6.85	100.0%	15.19	
	RCV1-10	12.05	100.0%	11.37	100.0%	11.61	100.0%	8.68	100.0%	6.46	100.0%	10.03	
RTA	IMDB	7.73	94.7%	8.21	99.6%	8.19	100.0%	7.13	100.0%	4.80	99.6%	7.21	7.21
	CITESEER	4.43	98.2%	4.30	99.1%	0.92	40.0%	4.22	99.1%	2.45	97.3%	3.26	
	CORA	-8.90	0.0%	-10.86	0.0%	-9.04	0.0%	-6.76	0.0%	-4.68	0.0%	-8.05	
	PUBMED	1.55	66.2%	1.53	68.4%	3.03	100.0%	3.77	100.0%	3.57	100.0%	2.69	
	USPS-10	23.57	100.0%	19.18	100.0%	15.83	100.0%	12.08	100.0%	7.73	100.0%	15.68	
	RCV1-10	11.06	100.0%	10.10	100.0%	10.88	100.0%	7.99	100.0%	6.16	100.0%	9.24	
WTA	IMDB	7.82	100.0%	7.80	100.0%	6.17	100.0%	6.18	100.0%	3.44	100.0%	6.28	6.28
	CITESEER	6.20	100.0%	4.27	100.0%	5.37	100.0%	4.20	100.0%	2.58	100.0%	4.53	
	CORA	-3.42	0.0%	-3.91	0.0%	-3.21	0.0%	-4.35	0.0%	-1.02	66.7%	-3.18	
	PUBMED	-4.44	0.0%	-4.60	0.0%	-3.01	0.0%	-1.03	33.3%	-1.35	0.0%	-2.89	
	USPS-10	26.69	100.0%	20.11	100.0%	15.44	100.0%	10.45	100.0%	4.73	100.0%	15.48	
	RCV1-10	10.85	100.0%	10.34	100.0%	9.60	100.0%	7.02	100.0%	4.15	100.0%	8.39	

Table 4: 100 times the difference of MCC variance between RST and MST.

		0% perturbation		2.5% perturbation		5% perturbation		10% perturbation		20% perturbation		average	average
method	dataset	average	better	average	better	average	better	average	better	average	better	per dataset	overall
SHAZOO	IMDB	-0.40	33.3%	0.86	66.7%	1.12	100.0%	1.63	100.0%	-0.59	33.3%	0.53	0.53
	CITESEER	0.67	33.3%	0.64	66.7%	0.70	66.7%	0.85	66.7%	0.88	66.7%	0.75	
	CORA	-0.10	33.3%	0.08	33.3%	-0.37	33.3%	-1.20	33.3%	0.31	33.3%	-0.26	
	PUBMED	0.80	66.7%	0.69	66.7%	0.54	66.7%	0.88	66.7%	1.39	100.0%	0.86	
	USPS-10	2.49	66.7%	2.11	66.7%	1.83	66.7%	1.12	66.7%	0.42	66.7%	1.59	
	RCV1-101	1.03	66.7%	0.66	66.7%	1.32	100.0%	1.21	66.7%	1.02	100.0%	1.05	
RTA	IMDB	0.39	67.1%	1.33	88.4%	0.47	66.2%	0.85	68.0%	-0.20	36.9%	0.57	0.57
	CITESEER	1.76	92.0%	1.48	92.4%	0.18	57.8%	0.85	72.0%	0.86	77.8%	1.03	
	CORA	-0.23	54.7%	-0.65	29.8%	-1.17	29.8%	-0.32	54.2%	-0.37	35.6%	-0.55	
	PUBMED	0.08	44.9%	0.02	42.7%	-0.24	32.9%	0.50	52.9%	0.99	90.7%	0.27	
	USPS-10	3.81	100.0%	2.57	82.7%	2.91	84.4%	1.60	67.6%	-0.39	33.8%	2.10	
	RCV1-101	0.64	73.8%	0.48	63.1%	0.63	64.0%	0.37	48.0%	1.11	93.3%	0.64	
WTA	IMDB	-0.33	0.0%	-1.09	33.3%	1.26	66.7%	0.51	66.7%	-0.78	33.3%	-0.09	-0.09
	CITESEER	0.80	33.3%	1.18	66.7%	0.73	66.7%	1.88	100.0%	0.32	66.7%	0.98	
	CORA	0.61	66.7%	0.22	33.3%	0.44	66.7%	0.13	33.3%	0.62	100.0%	0.40	
	PUBMED	0.70	33.3%	0.55	33.3%	0.87	66.7%	0.06	66.7%	0.69	66.7%	0.57	
	USPS-10	0.82	33.3%	0.07	33.3%	0.95	66.7%	1.22	66.7%	0.52	66.7%	0.72	
	RCV1-101	1.09	66.7%	0.72	33.3%	0.73	33.3%	0.82	66.7%	0.80	66.7%	0.83	

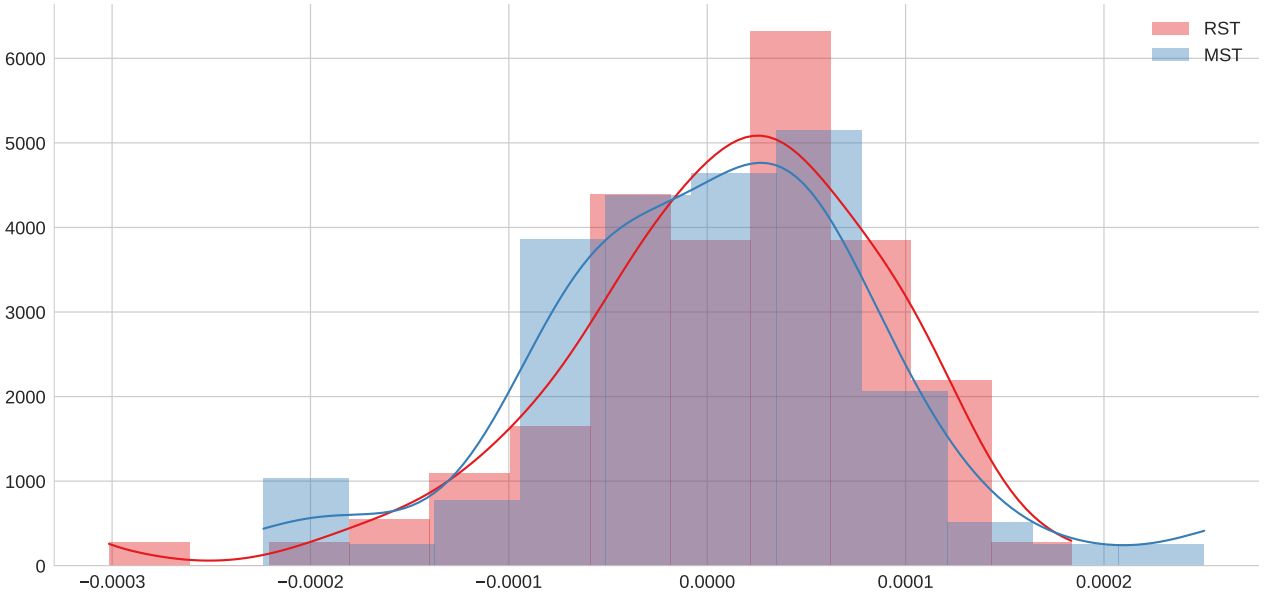
2.3 Influence of the number of presentation

For each dataset, perturbation level and dimension, I have a plot with 15 results for $n_o = 11$ presentation orders, 15 results for $n_o = 33$ presentation orders and so on for $n_o \in [11, 33, 57, 79, 101]$. I computed a least square regression for all this plots, taking as output variable either the 75 averages MCC (Figure 1a), the 75 MCC standard deviations (ie the size of the error bar) (Figure 1b) and the 5 standard deviations among the 15 averages MCC for a given n_o (Figure 1c).

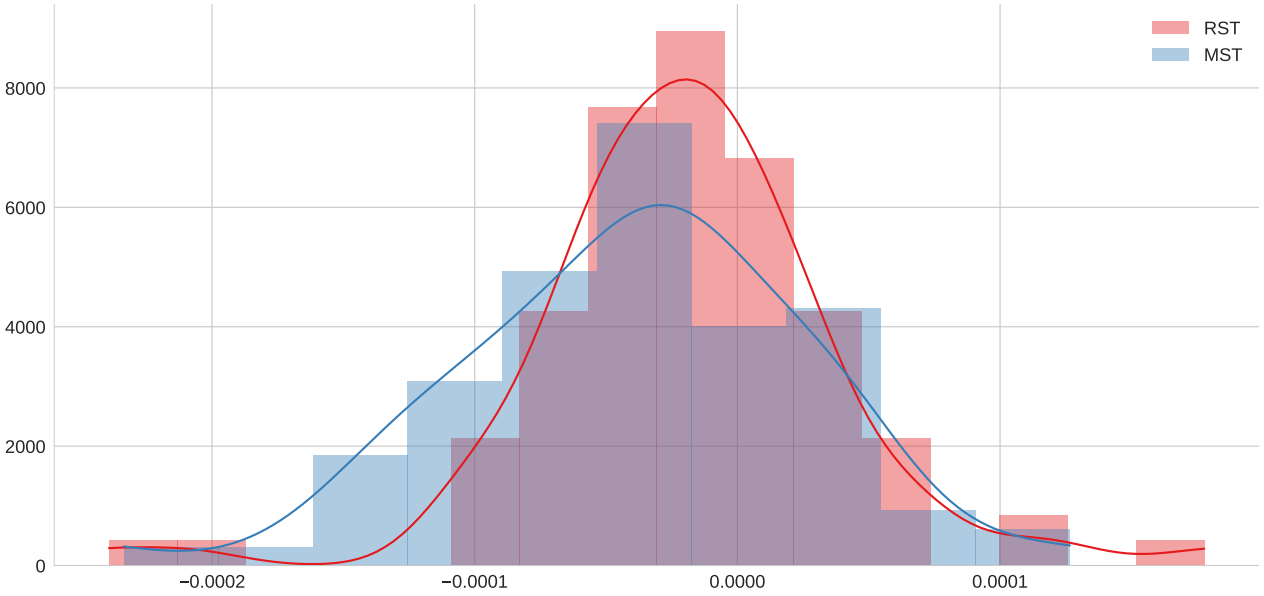
What this shows is that increasing the number of presentation orders has no consistent effect on the MCC performance, whether in terms of higher average or lower variance. On the other hands, because there is only 117 presentation orders to choose from, increasing the number of presentation brings us closer to the "true" performance of these 117 presentations (in which case the variance would be 0 for there would be a single choice).

2.4 Influence of the number of presentation

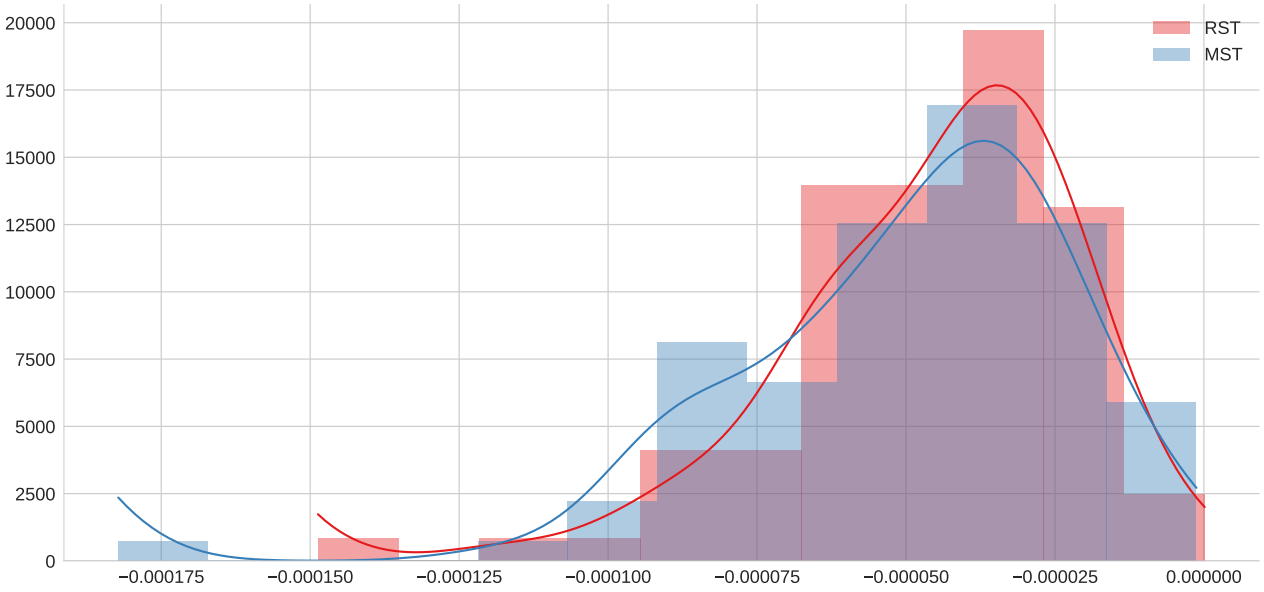
Finally in Tables 5 and 6, we compare SHAZOO+MST and RTA+MST. Unfortunately, it seems that with a single tree and default learning rate, RTA is almost always worse than SHAZOO, even though the gap decreases when the perturbation increases.



(a) slope of mean MCC w.r.t to number of presentation orders.



(b) slope of MCC standard deviation w.r.t to number of presentation orders.



(c) slope of standard deviation across MCC w.r.t to number of presentation orders.

Figure 1

Table 5: 100 times the MCC difference between RTA and SHAZOO (i.e positive values mean RTA is performing better).

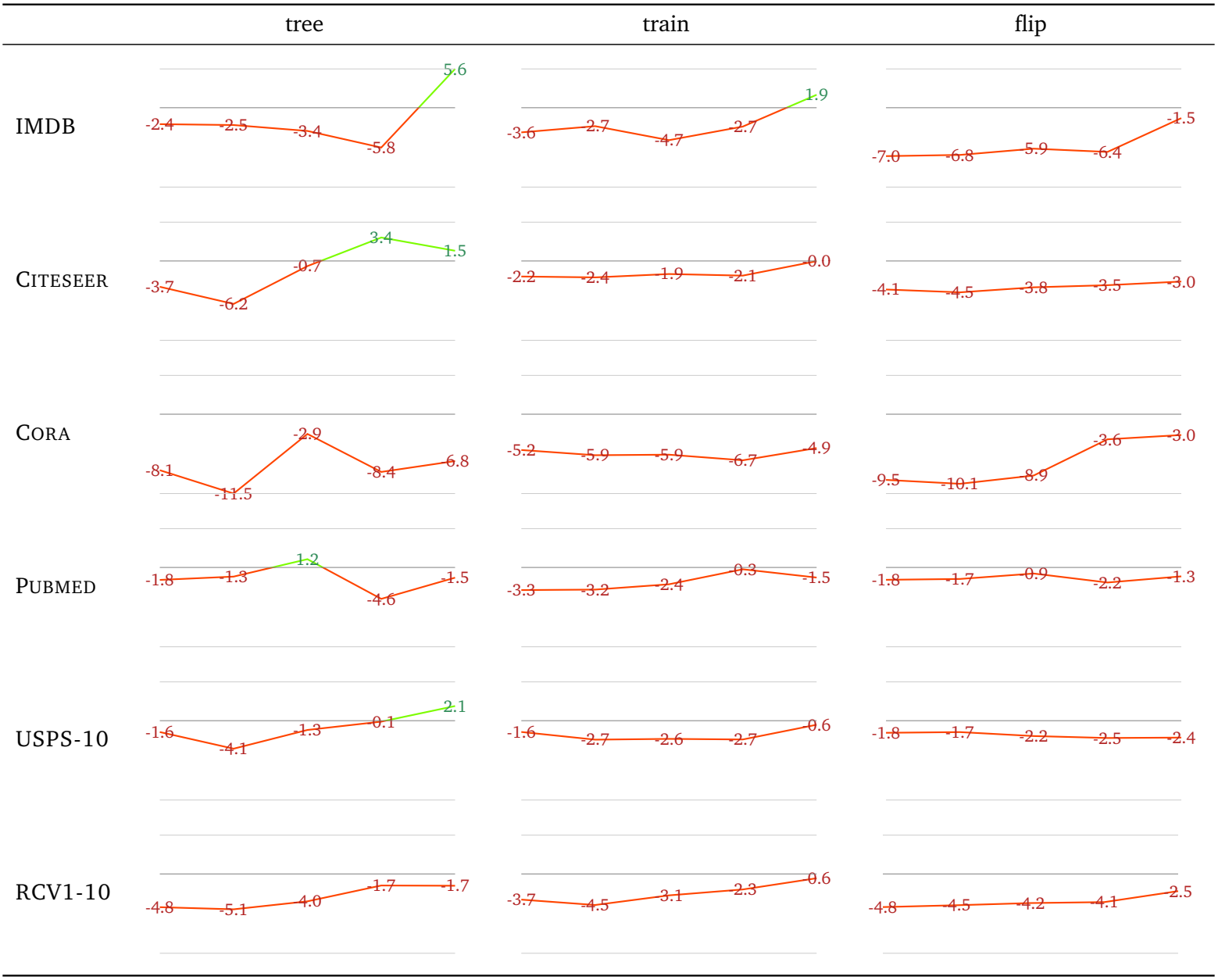


Table 6: 100 times the MCC difference between RTA and SHAZOO, averaged over the 3 dimensions source of randomness, and then over datasets or perturbation levels.

	IMDB	CITESEER	CORA	PUBMED	USPS-10	RCV1-10	average
0% perturbation	−4.32	−3.37	−7.61	−2.29	−1.68	−4.45	−3.95
2.5% perturbation	−4.00	−4.38	−9.17	−2.07	−2.82	−4.72	−4.53
5% perturbation	−4.67	−2.15	−5.87	−0.71	−2.07	−3.76	−3.20
10% perturbation	−4.97	−0.74	−6.22	−2.33	−1.79	−2.67	−3.12
20% perturbation	1.99	−0.51	−4.90	−1.41	−0.31	−1.61	−1.12
average	−3.19	−2.23	−6.75	−1.76	−1.73	−3.44	−3.19

3 Varying the rate at which we update γ in case of mistake

As we saw previously, SHAZOO+MST is more competitive than RTA+MST. Here we explore how using update of the node guiltiness of the form $a\gamma$, $a \in \mathbb{R}^+$ changes RTA behavior. For the record, all results presented so far were computed with a arbitrarily set to 1.5. As a starting point, I explore 33 values of a between 0 and 6 for a single MST tree, a single perturbation of level of 20% and 10 training sets of size 10%[3]. Setting $a = 0$ makes RTA equivalent to SHAZOO, as confirmed in Figure 2, where we also see that values above 2 make no sense. In Table 7, we therefore focus on the range $[1/2, 1]$ where we can indeed obtain modest gains (except again on the CORA dataset). Those gains are even smaller with a 5% perturbation (see Table 8) but at least RTA remains comparable with SHAZOO. A similar conclusions hold at 35% perturbation, see Table 9.

Table 7: Difference of 100 times average MCC (and standard deviations) between RTA and SHAZOO for 3 small γ multipliers. In both cases, positive values means than RTA is better than SHAZOO, since we want high average and low variance.

	multiplier	IMDB	CITESEER	CORA	PUBMED	USPS-10	RCV1-10	average	average w/o CORA
average MCC	0.562	0.31	0.06	−2.02	0.86	0.11	0.27	−0.07	0.32
of	0.75	1.31	−0.08	−2.83	1.32	0.19	0.38	0.05	0.62
RTA − SHAZOO	0.938	2.34	−0.31	−2.68	0.42	0.20	0.79	0.13	0.69
MCC variance	0.562	0.28	−0.12	−1.90	0.53	−0.02	0.11	−0.19	0.16
of	0.75	1.37	−0.55	−1.34	0.23	−0.26	−0.50	−0.17	0.06
SHAZOO − RTA	0.938	1.81	−0.47	−1.72	−0.01	−0.45	−0.23	−0.18	0.13

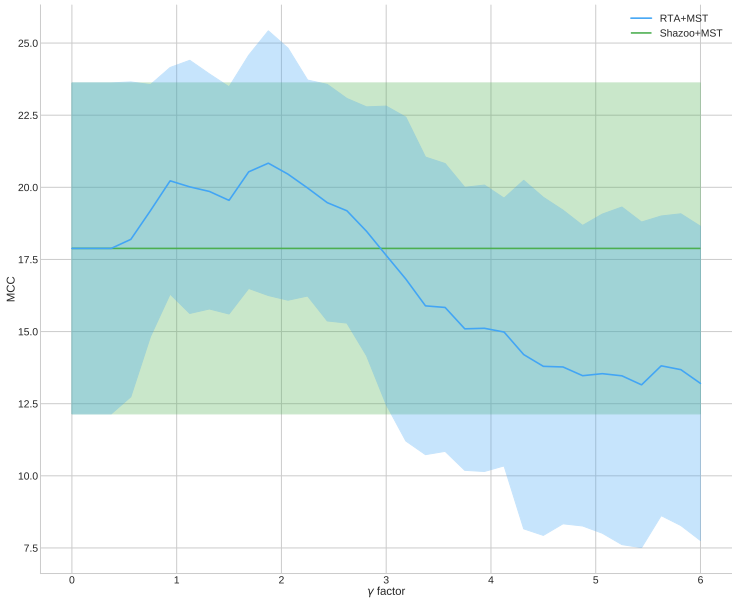
Table 8: Same as Table 7 but at 5% perturbation level.

	multiplier	IMDB	CITESEER	CORA	PUBMED	USPS-10	RCV1-10	average	average w/o CORA
average MCC	0.562	0.02	0.26	−0.72	0.70	0.02	−0.03	0.04	0.19
of	0.75	−0.99	0.06	−1.00	0.64	0.26	−0.29	−0.22	−0.07
RTA − SHAZOO	0.938	−2.35	−0.36	−1.02	−0.36	0.36	−0.07	−0.64	−0.56
MCC variance	0.562	−0.02	0.29	−1.63	−0.60	−0.01	0.02	−0.32	−0.06
of	0.75	−0.76	0.16	−2.94	−0.94	0.21	0.12	−0.69	−0.24
SHAZOO − RTA	0.938	−0.80	0.09	−2.38	−1.21	0.15	−0.49	−0.77	−0.45

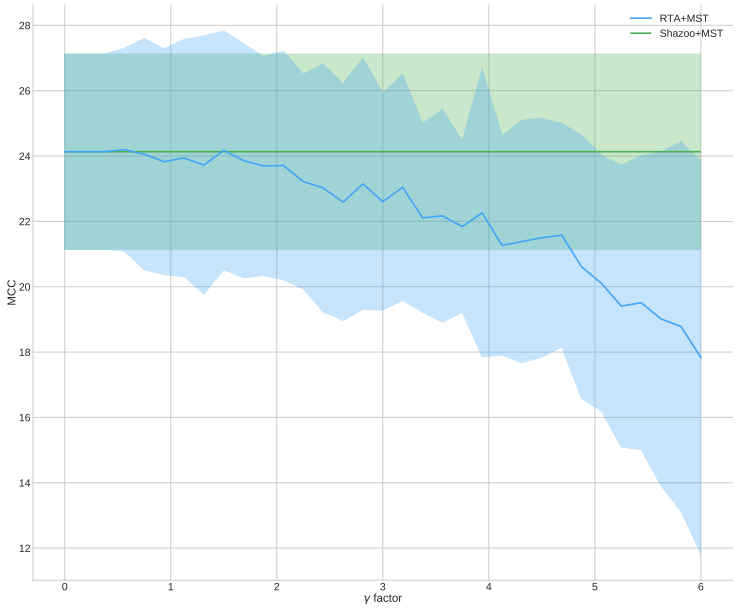
Table 9: Same as Table 7 but at 35% perturbation level.

	multiplier	IMDB	CITESEER	CORA	PUBMED	USPS-10	RCV1-10	average	average w/o CORA
average MCC	0.562	0.86	−0.94	−1.82	−0.58	0.13	−0.04	−0.40	−0.12
of	0.75	1.47	−1.37	−1.59	−1.33	0.07	−0.15	−0.48	−0.26
RTA − SHAZOO	0.938	0.55	−1.97	−1.42	−2.19	−0.03	0.05	−0.83	−0.72
MCC variance	0.562	0.28	−0.36	−1.37	0.44	0.03	−0.01	−0.17	0.07
of	0.75	0.48	−0.51	−0.84	0.22	0.02	−0.09	−0.12	0.02
SHAZOO − RTA	0.938	−0.51	−0.28	−0.49	0.69	0.14	−0.08	−0.09	−0.01

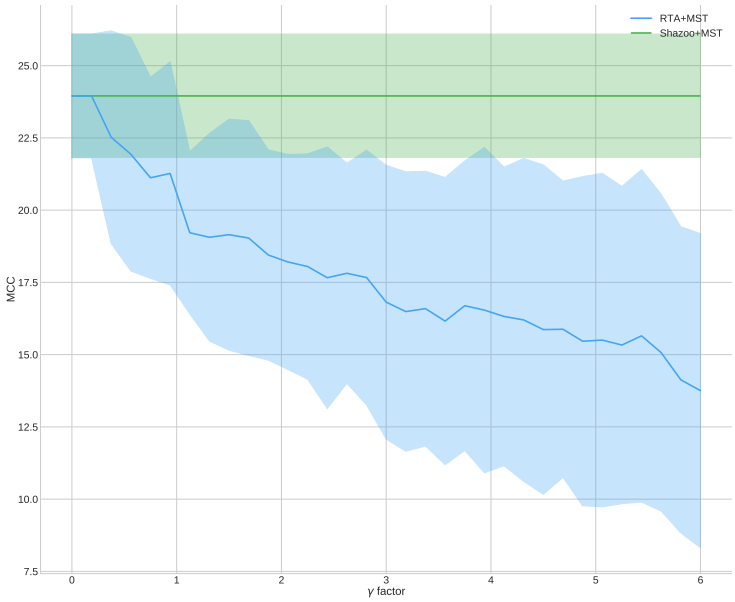
[3]Only the training set varies since all sources of randomness cause similar level of variance.



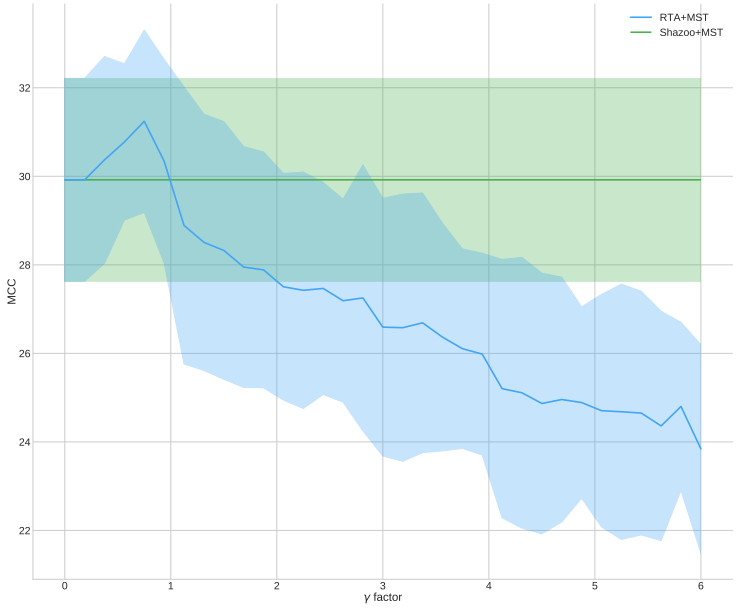
(a) IMDB



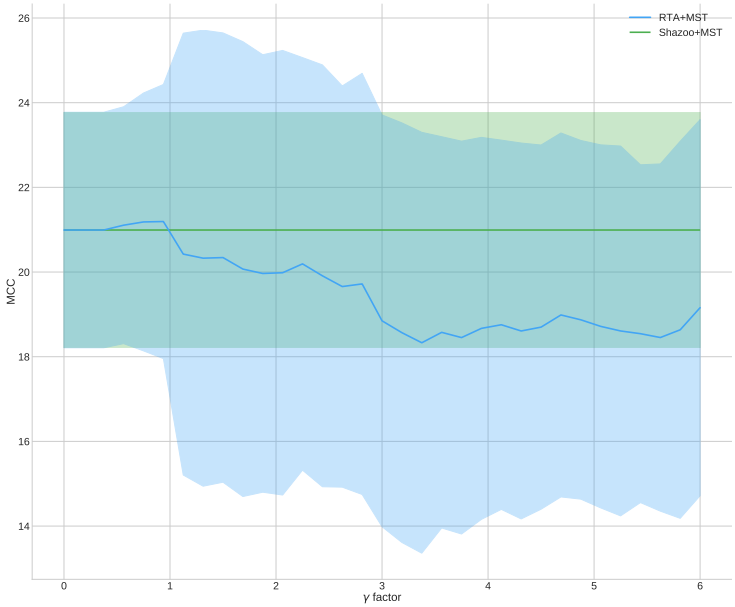
(b) CITESEER



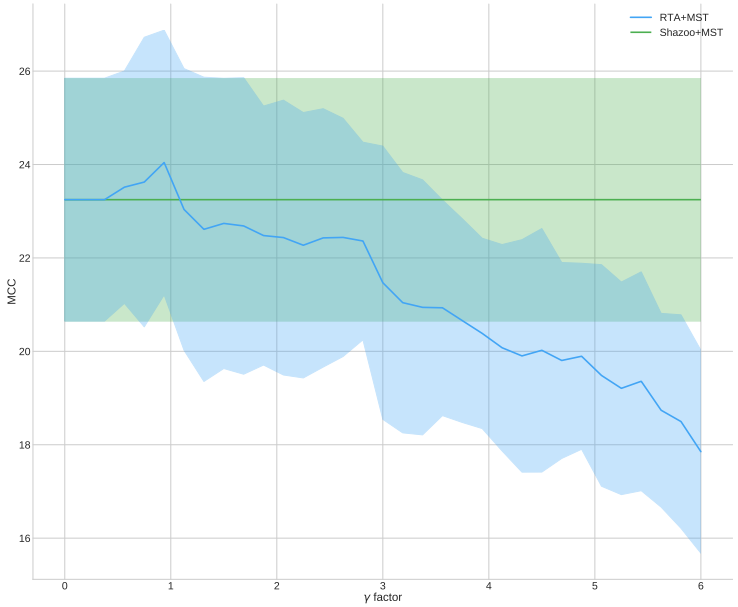
(c) CORA



(d) PUBMED



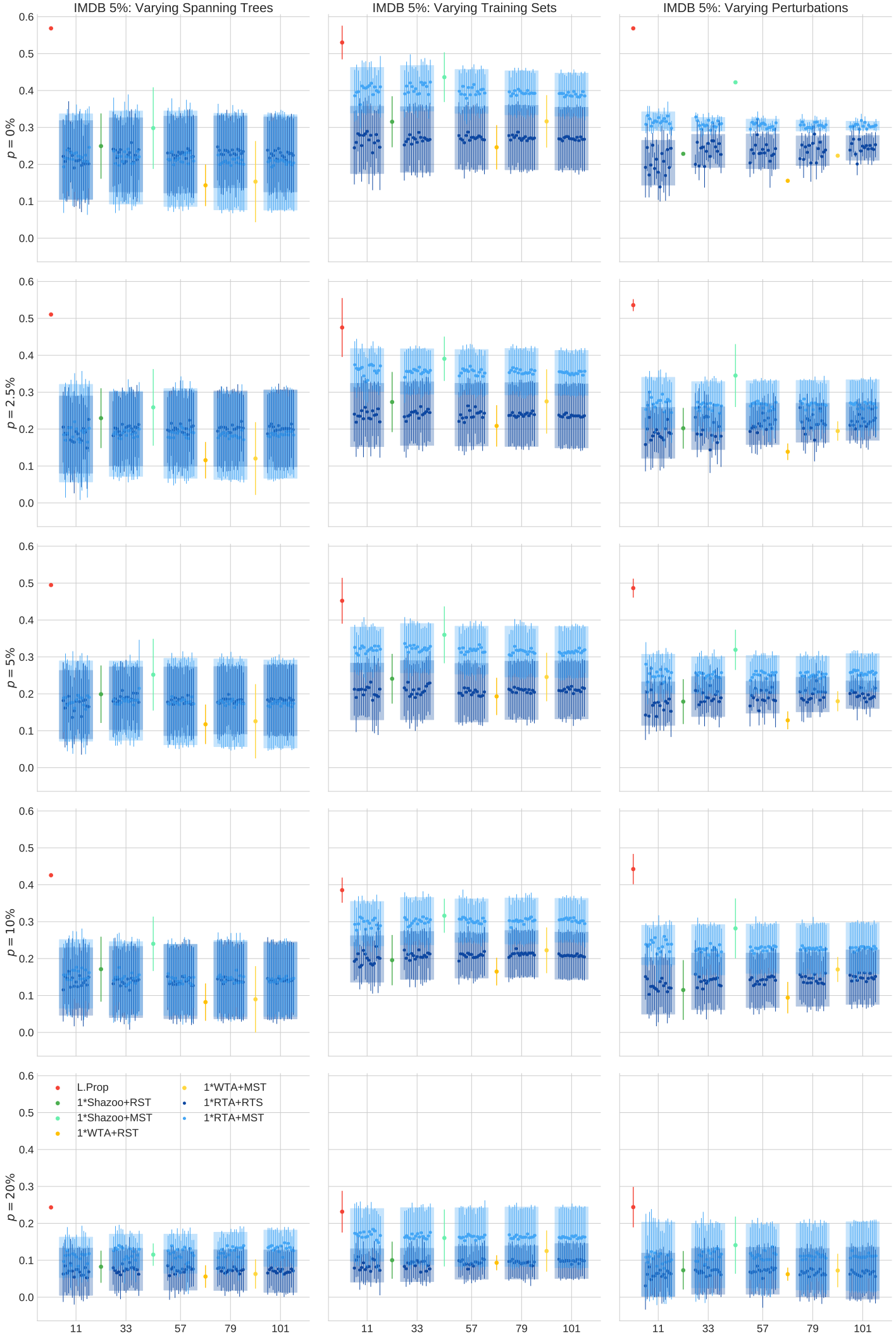
(e) USPS-10

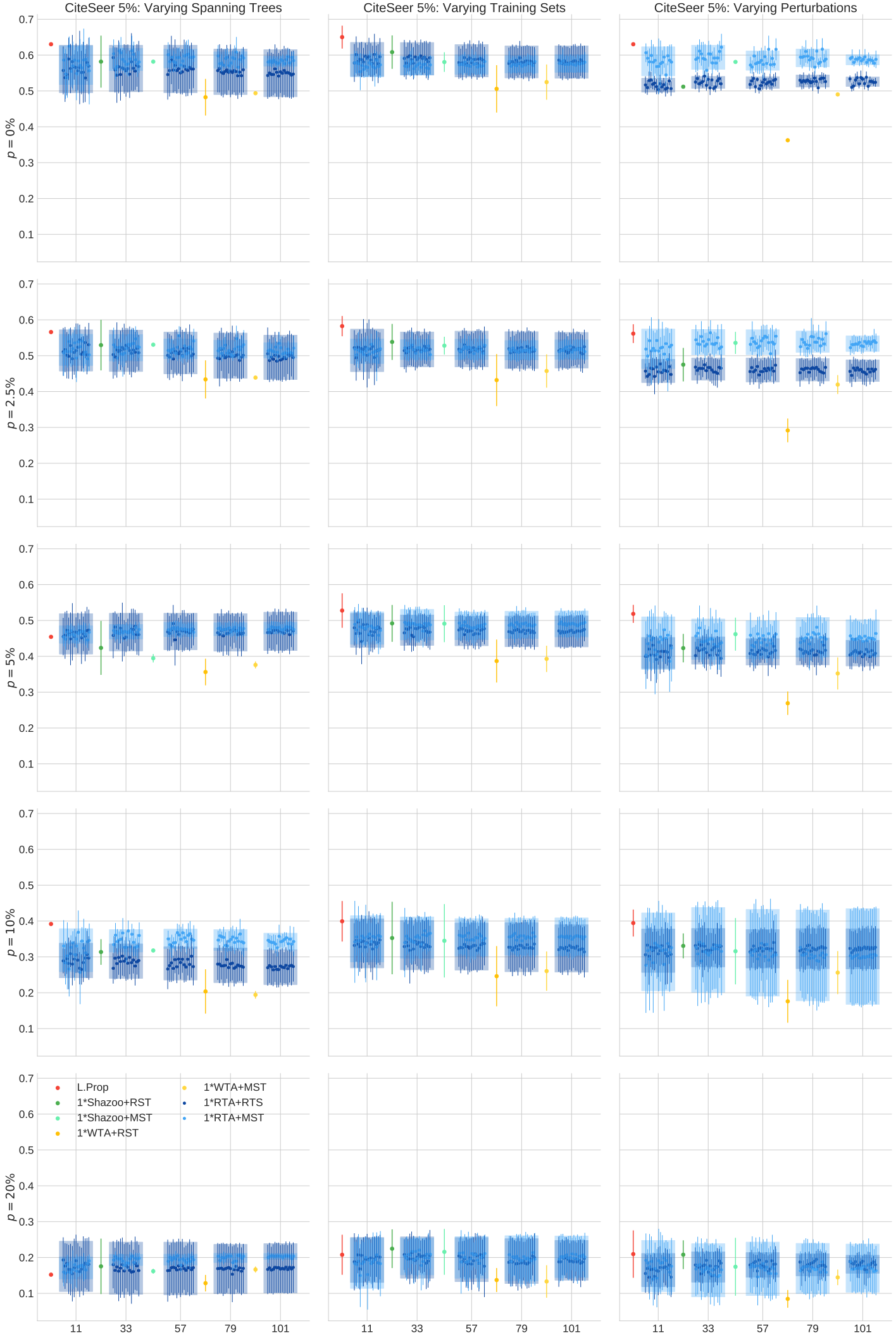


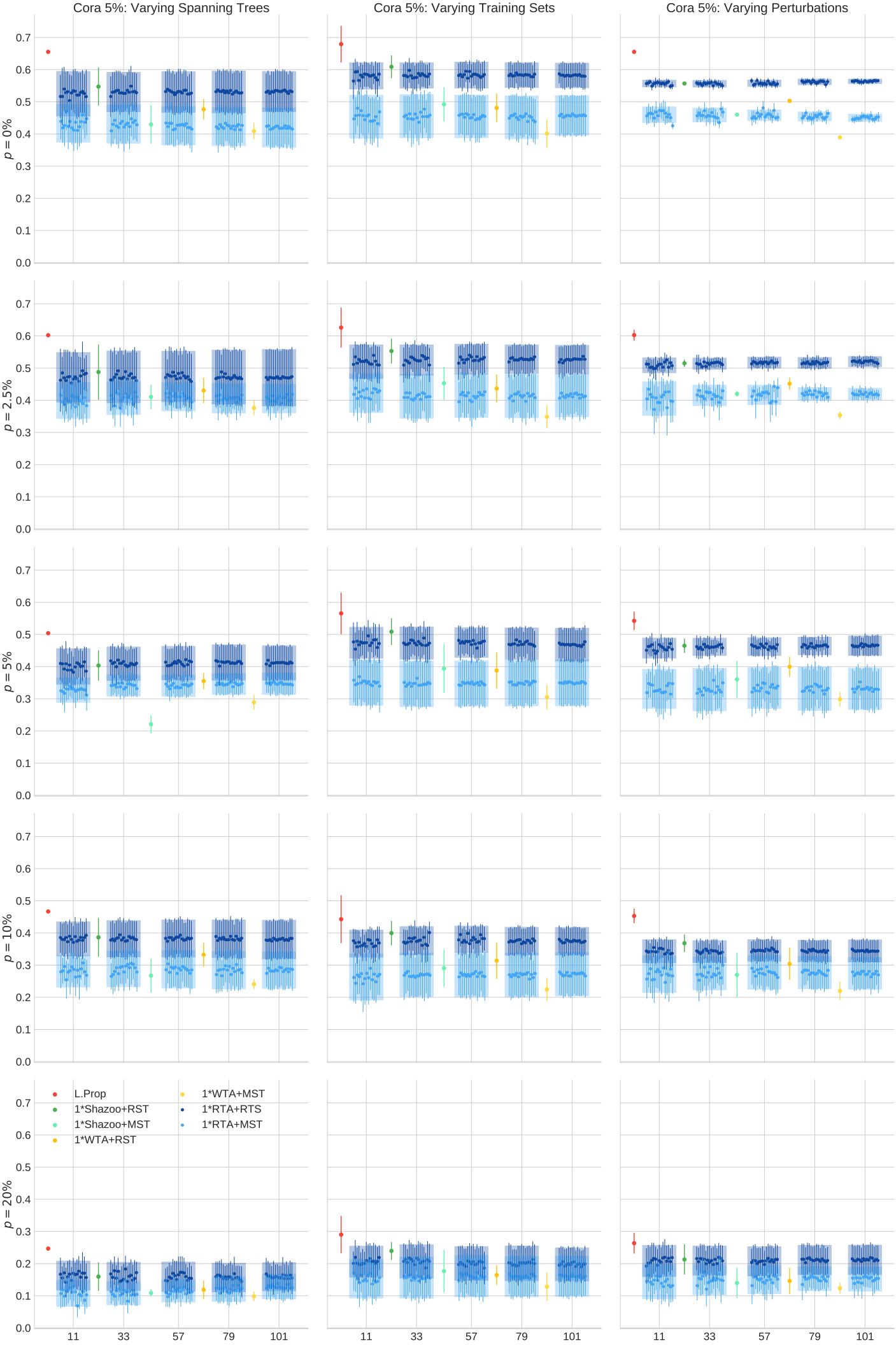
(f) RCV1-10

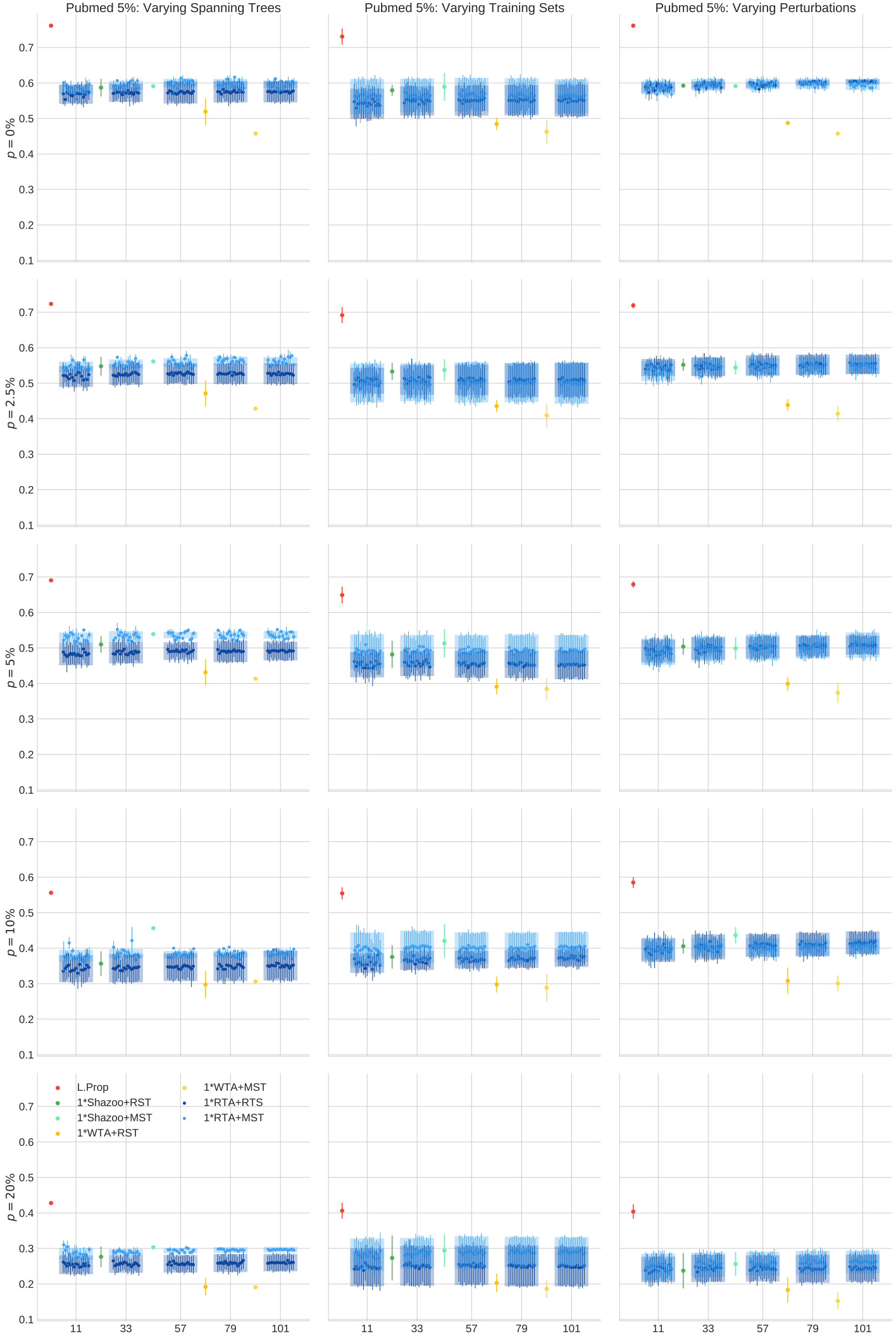
Figure 2: Average MCC over 10 training set of size 10% with a single MST tree and a single 20% perturbation as a function of γ multiplier.

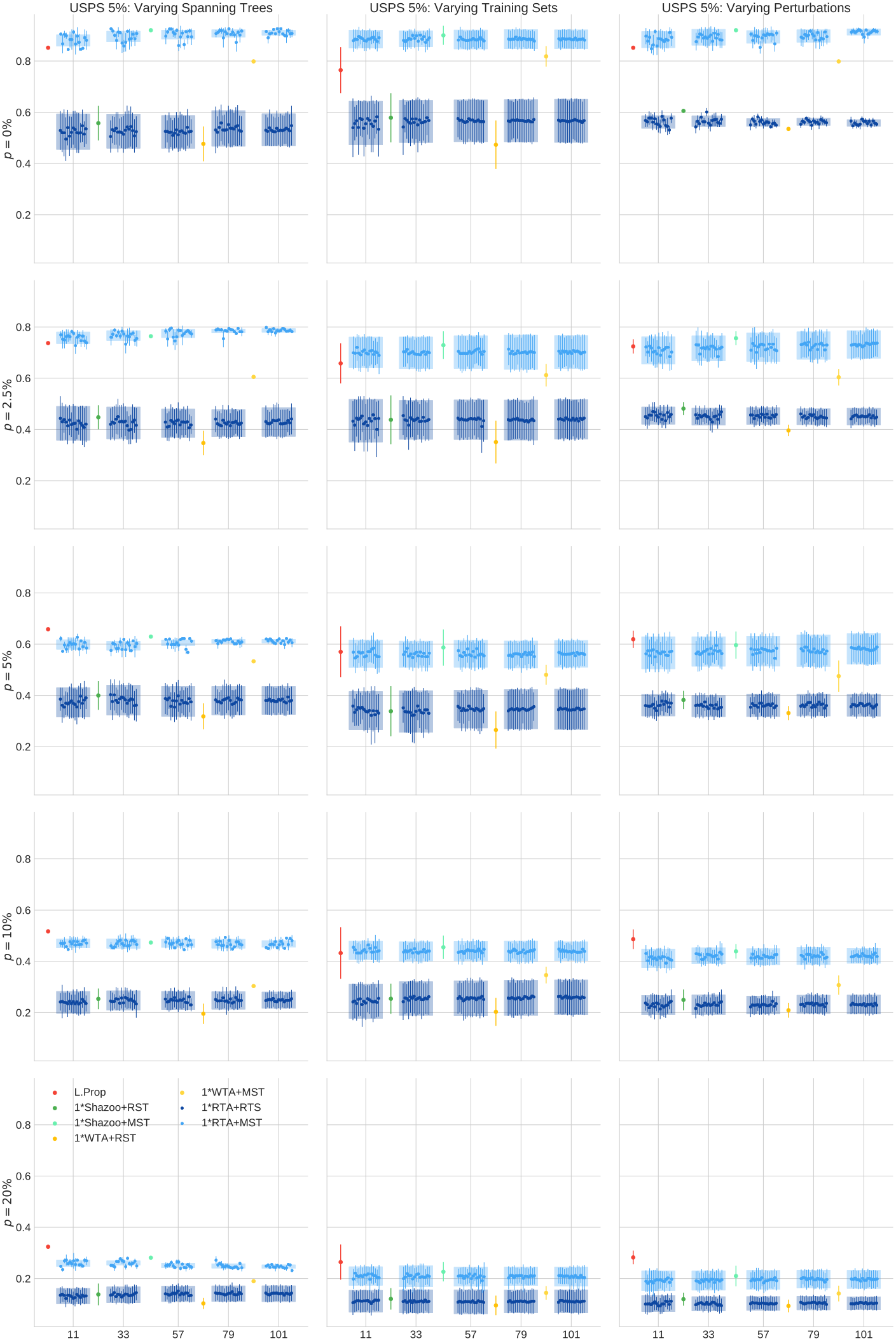
4 Additional plots for 5% training set size (with $a = 1.5$ update rule)

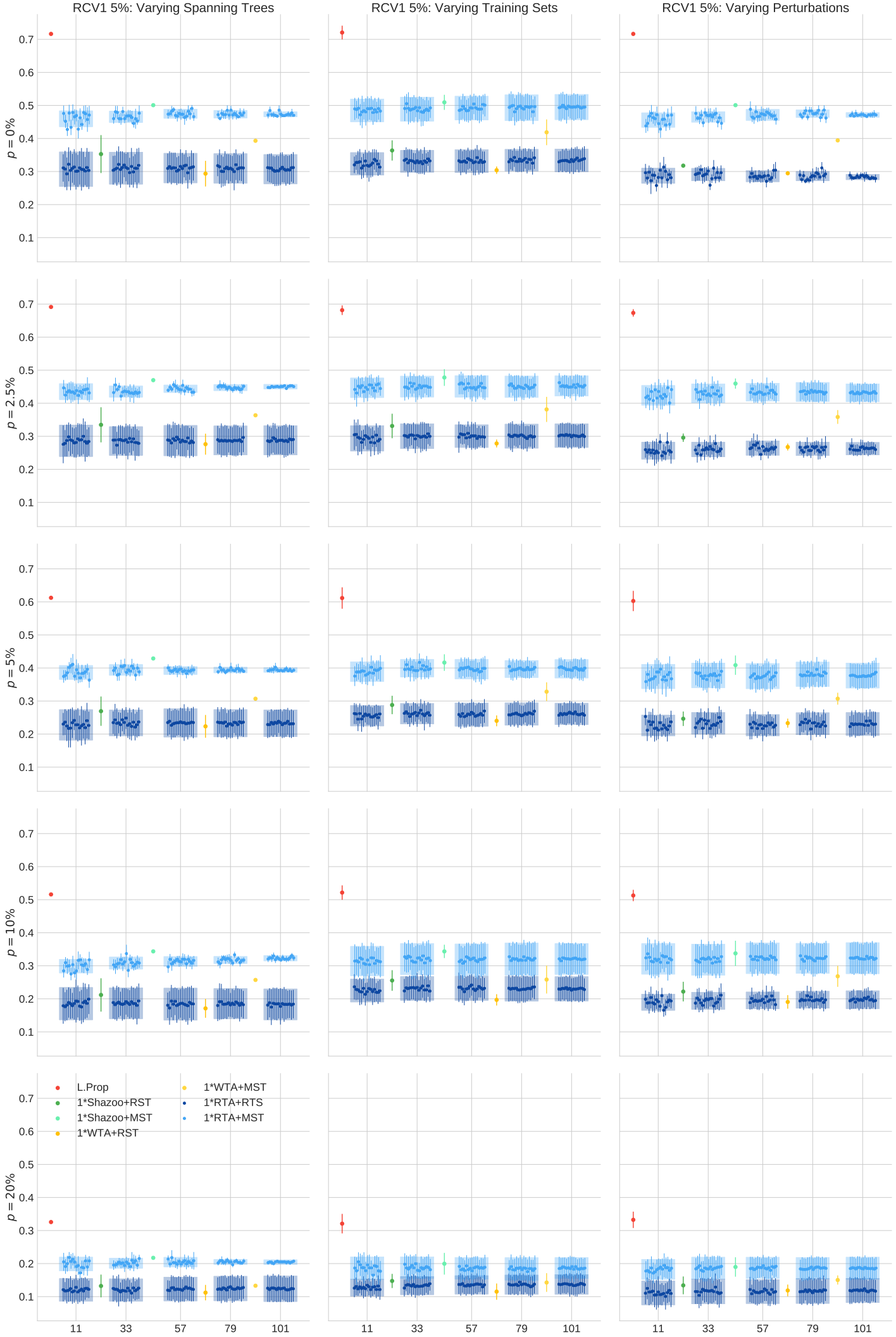

IMDB 5%: Varying Spanning Trees
IMDB 5%: Varying Training Sets
IMDB 5%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.0
0.1
0.2
0.3
0.4
0.5
0.6
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


CiteSeer 5%: Varying Spanning Trees
CiteSeer 5%: Varying Training Sets
CiteSeer 5%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.7
0.6
0.5
0.4
0.3
0.2
0.1
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


Cora 5%: Varying Spanning Trees
Cora 5%: Varying Training Sets
Cora 5%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


Pubmed 5%: Varying Spanning Trees
Pubmed 5%: Varying Training Sets
Pubmed 5%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.1
0.2
0.3
0.4
0.5
0.6
0.7
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


USPS 5%: Varying Spanning Trees
USPS 5%: Varying Training Sets
USPS 5%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.2
0.4
0.6
0.8
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST


RCV1 5%: Varying Spanning Trees
RCV1 5%: Varying Training Sets
RCV1 5%: Varying Perturbations
p = 0%
p = 2.5%
p = 5%
p = 10%
p = 20%
0.7
0.6
0.5
0.4
0.3
0.2
0.1
11
33
57
79
101
L.Prop
1*Shazoo+RST
1*Shazoo+MST
1*WTA+RST
1*WTA+MST
1*RTA+RTS
1*RTA+MST