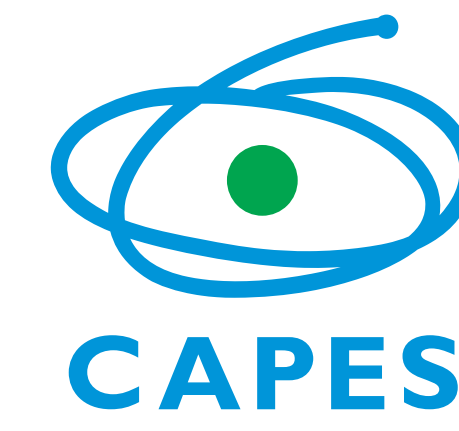


# Fast And Accurate Learning of Probabilistic Circuits by Random Projections

Renato Lui Geh, Denis Deratani Mauá

Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, Brazil

{renatolg, ddm}@ime.usp.br



## 1. Motivation

Correia et al. [2020]: there is a link between DTs and PCs!

- Well known, established results in DTs;
- But mostly in supervised tasks; Khosravi et al. [2020]
- What if we explore this link in an unsupervised context?

Random Projections (RPs):

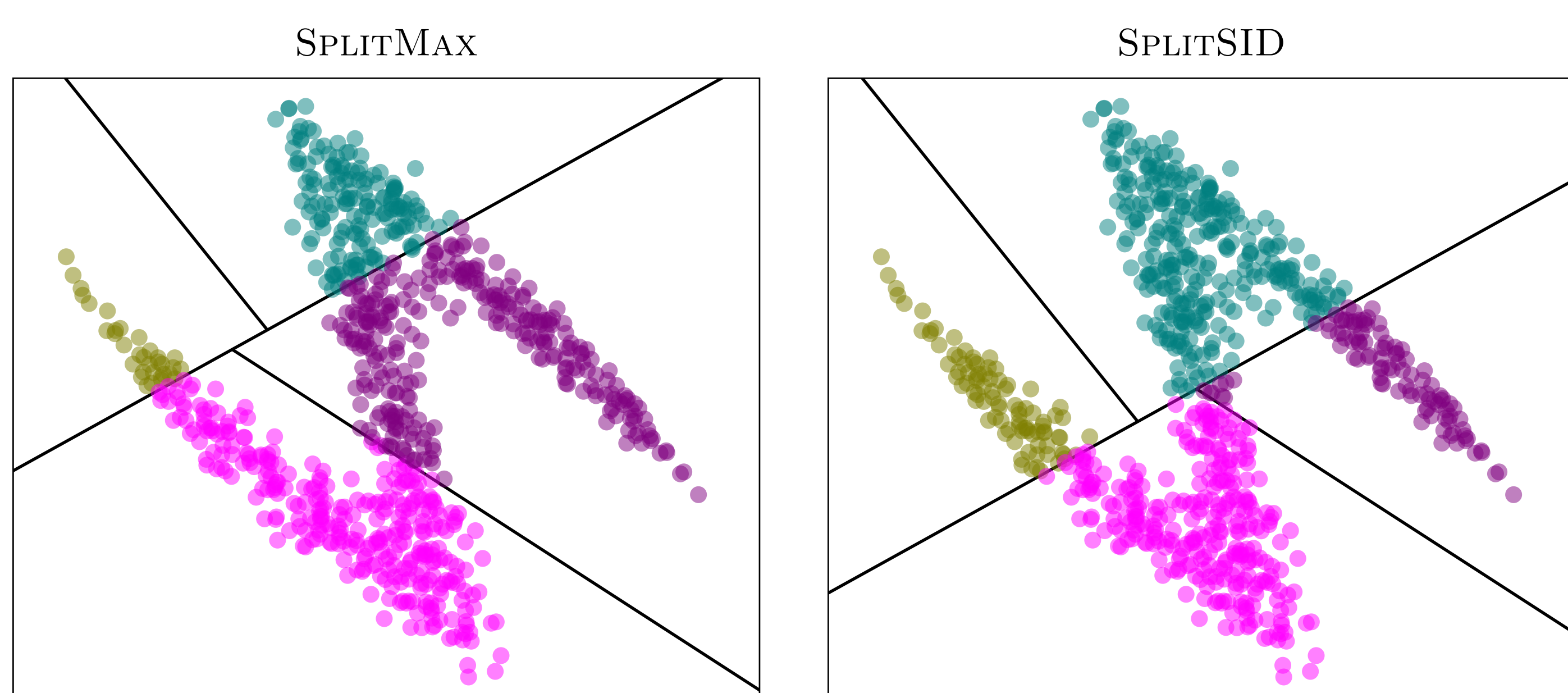
- Interesting theoretical guarantees; Dasgupta and Freund [2008]
- Can we transplant this to PCs?

Idea behind RPs:

- Construct hyperplane  $\mathcal{H}$  with random direction;
- Apply perturbation  $\delta$  to  $\mathcal{H}$  st. divides data somewhat equally;
- Choice of  $\delta$ :

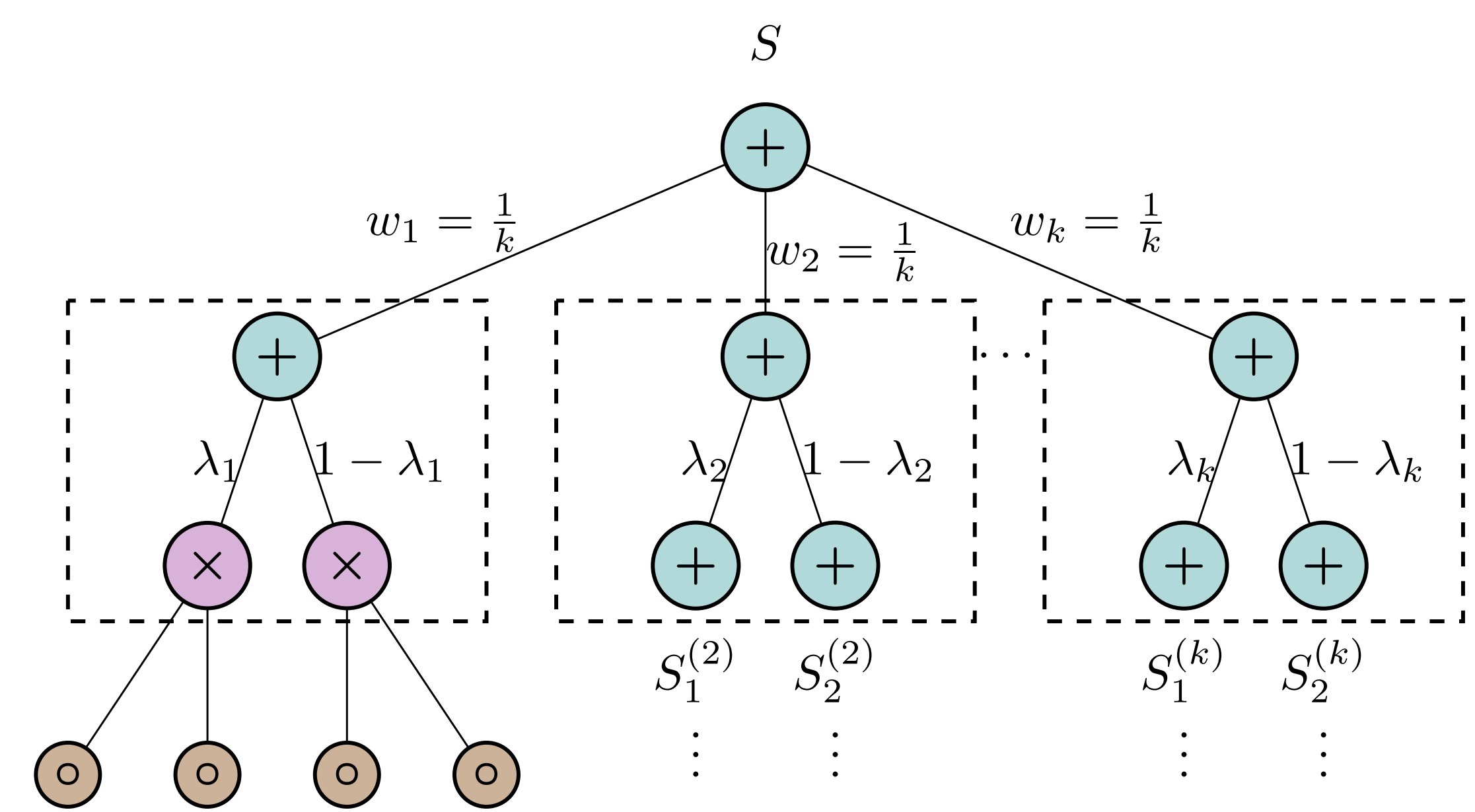
**Max:** by diameter (maximal distance);

**SID:** by Square Interpoint Distance.

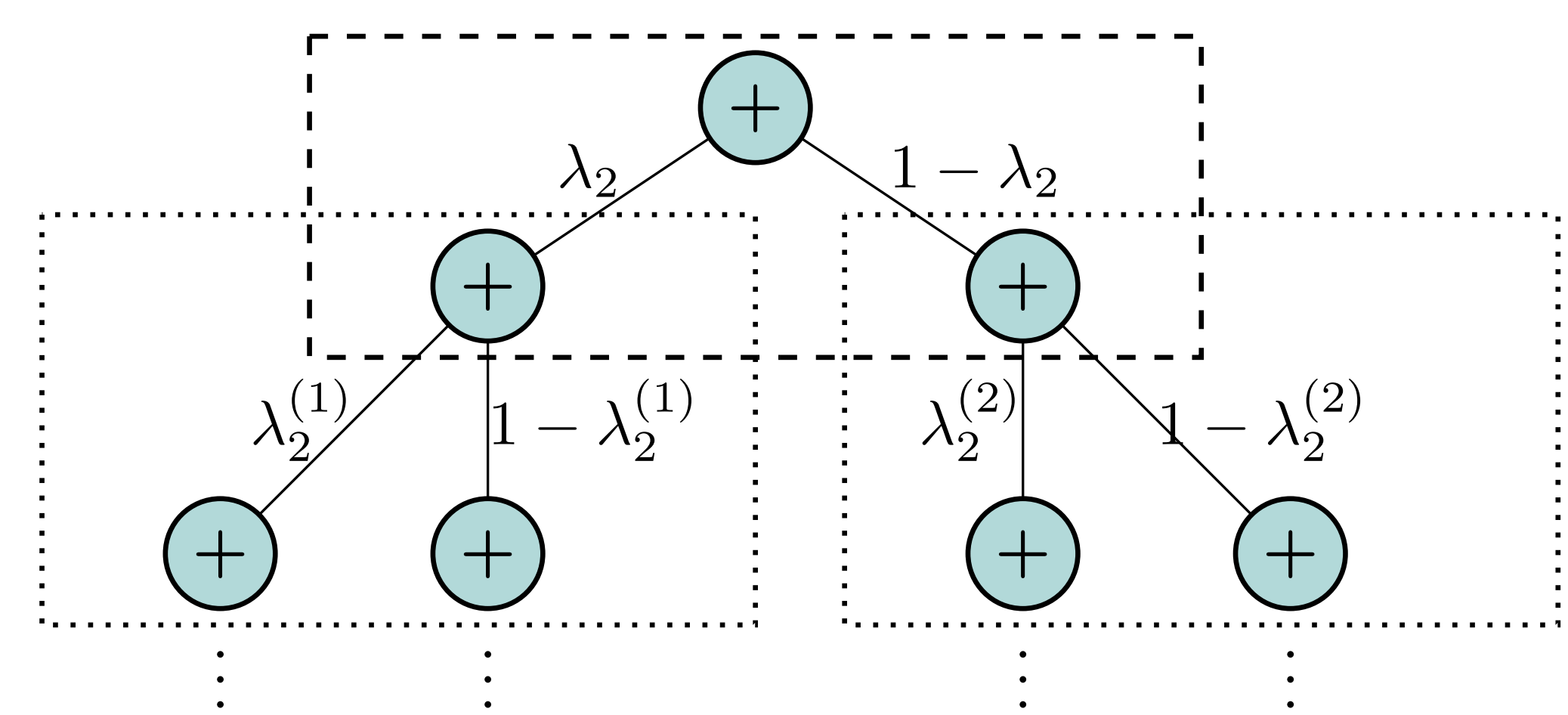


What if we use RPs as sum nodes in PCs?

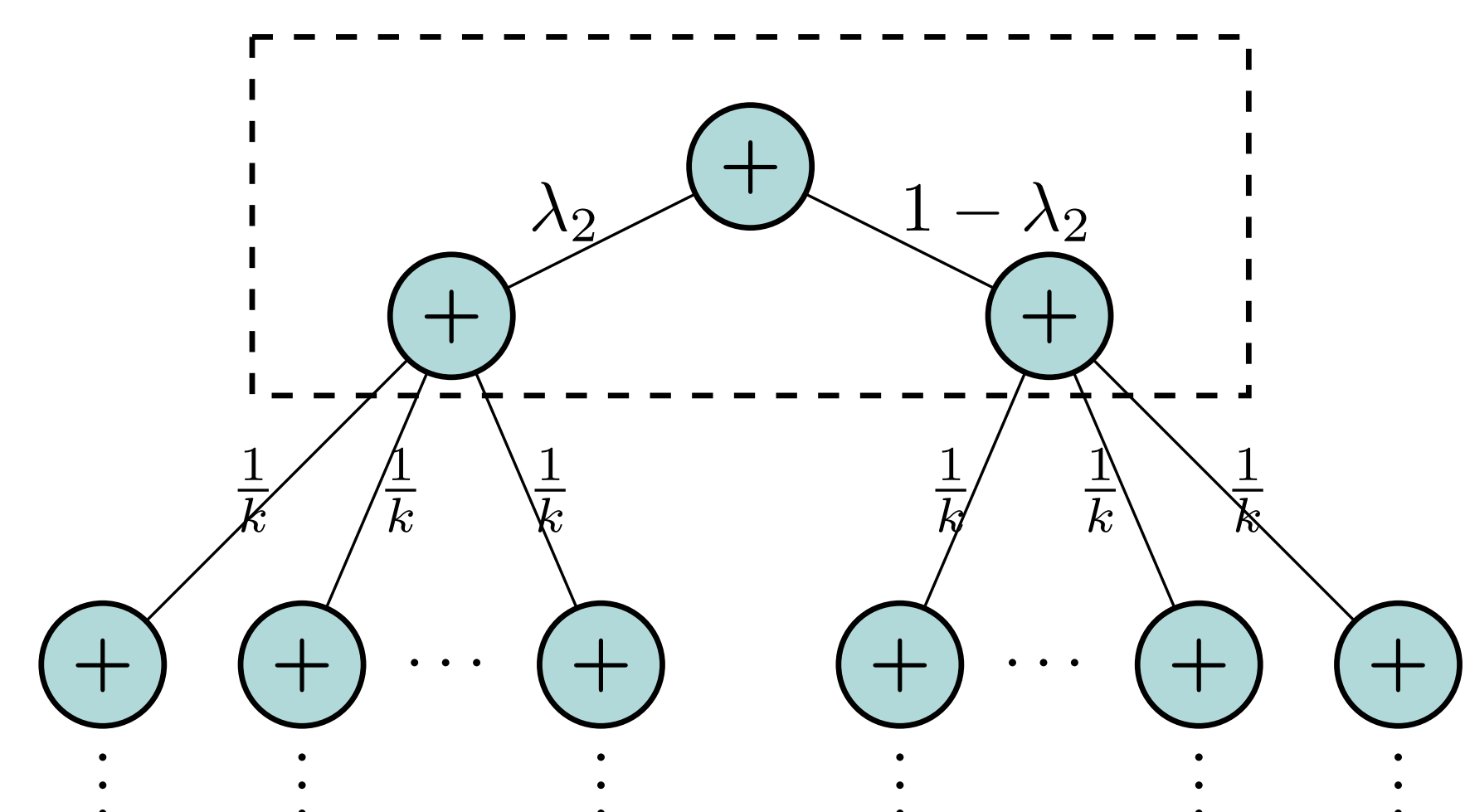
## 2. Learning PCs by RPs



Sum nodes as RPs:



Mixtures of RPs:



## 3. Experiments and Further Work

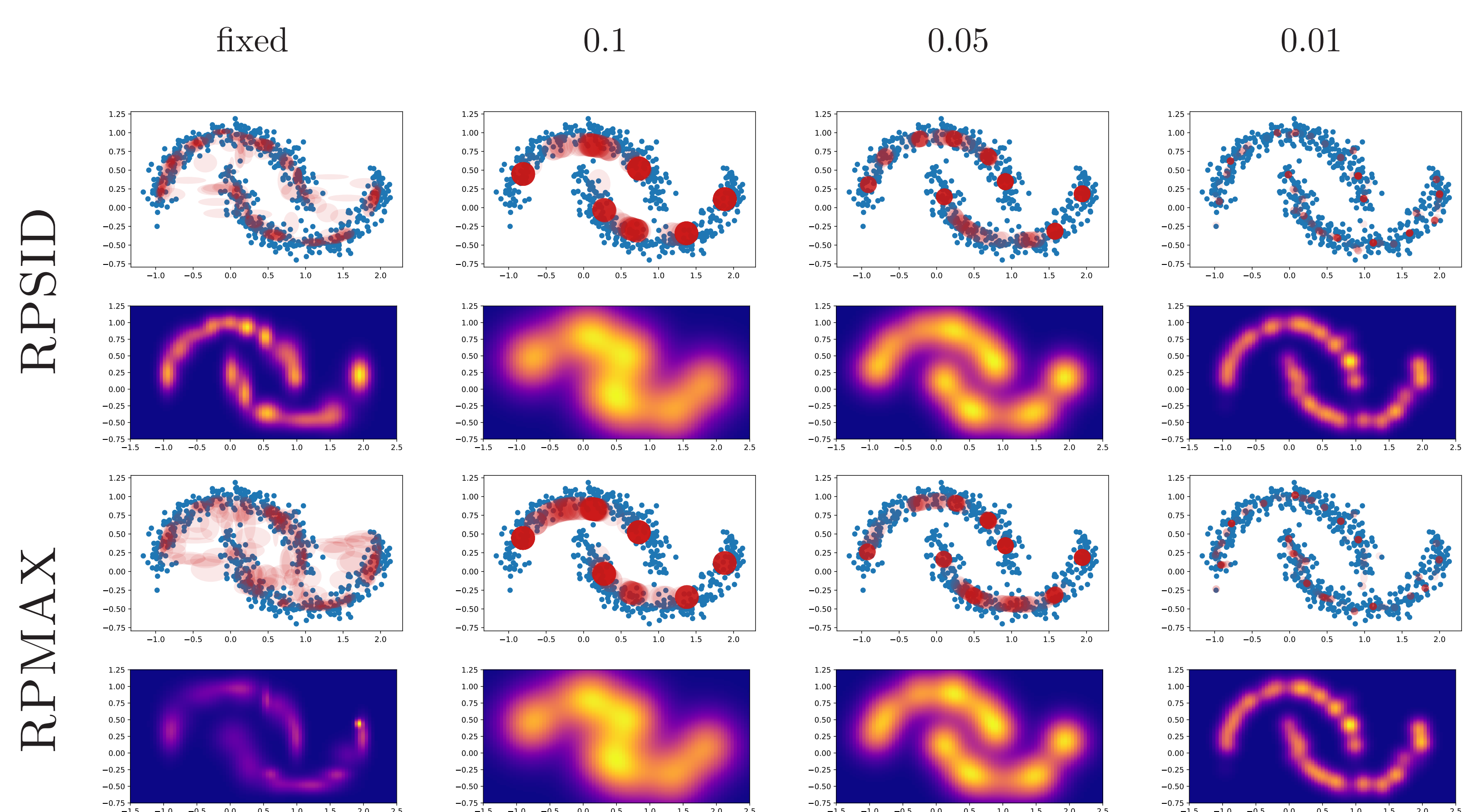
Experiments:

- Synthetic dataset: 2-moons;
- Benchmark datasets: 20 binary density datasets;
- Almost never best, but promising results given its simplicity.

Further work:

- RP theoretical guarantees extend to PCs?
- Are there other desirable statistical properties?
- RPs in more elaborate structural learners?
- More complex leaf distributions (CLTs, other PCs)?

Gens and Domingos [2013], Dang et al. [2020], Mauro et al. [2021]



Dataset	RPMaSD	RPMaD	RPSIDSD	RPSIDD	RPMaS	RPMa	RPSIDS	RPSID	LSPN	Strudel	LPSDD	EXPC
NLTCS	-6.02	-6.02	-6.03	-6.05	[-6.01]	<u>-6.01</u>	-6.01	-6.01	-6.11	-6.06	<b>-5.99</b>	-6.05
PLANTS	-13.94	-14.06	-14.00	-13.96	-14.07	-13.86	-14.02	-13.94	<b>-12.97</b>	<u>-12.98</u>	[-13.02]	-14.19
AUDIO	-40.25	-40.25	[-40.20]	<u>-40.20</u>	-40.29	-40.25	-40.28	-40.23	-40.50	-41.50	<b>-39.94</b>	-40.91
JESTER	[-53.05]	<u>-52.99</u>	-53.09	-53.05	-53.21	-53.24	-53.13	-53.06	-75.98	-55.03	<b>-51.29</b>	-53.43
NETFLIX	-57.20	-57.21	-57.22	<u>-56.89</u>	-57.53	-57.40	-57.45	-57.44	[-57.02]	-58.69	<b>-55.71</b>	-57.58
ACCIDENTS	-36.77	-36.67	-36.96	-36.47	-37.41	-37.48	-37.16	-37.39	<u>-30.03</u>	<b>-28.73</b>	[-30.16]	-31.02
BOOK	-34.84	-34.92	-34.87	-34.82	[-34.75]	-34.85	<u>-34.74</u>	<b>-34.66</b>	-35.88	-34.99	-34.97	-34.75
DNA	-97.89	-96.86	-97.28	-97.64	-97.45	-97.17	-97.47	-96.68	<b>-82.52</b>	<u>-86.22</u>	-88.01	[-86.61]
NLTCS	3s	2s	5s	3s	3s	2s	6s	3s	7m	3m	6m	-
PLANTS	8s	4s	12s	7s	7s	4s	12s	7s	50m	41m	26m	-
AUDIO	9s	4s	14s	8s	8s	4s	14s	8s	2h	33m	51m	-
JESTER	5s	3s	9s	5s	4s	3s	9s	5s	52m	24m	37m	-
NETFLIX	8s	4s	15s	8s	7s	4s	14s	8s	1h	14m	33m	-
ACCIDENTS	7s	4s	13s	7s	7s	4s	13s	7s	47m	20m	41m	-
BOOK	12s	6s	18s	10s	9s	6s	17s	10s	>3h	8m	1.3h	-
DNA	1s	1s	3s	2s	2s	1s	2s	1s	>3h	>3h	>3h	-