Fast and Accurate Learning of Probabilistic Circuits by Random Projections

Renato L. Geh, Denis D. Mauá | Computer Science Dept., Institute of Mathematics and Statistics, University of São Paulo

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):

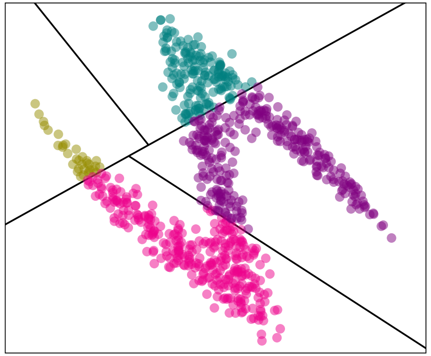
- Interesting theoretical guarantees; Dasgupta and Freund [2008]
- Can we transplant this to PCs?

Idea behind RPs:

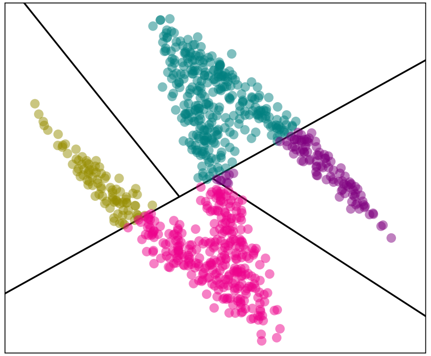
- Construct hyperplane $\mathcal{H}$ with random direction;
- Apply perturbation δ to $\mathcal{H}$ st. divides data somewhat equally;
- Choice of δ:

MAX: by diameter (maximal distance);
SID: by Square Interpoint Distance.

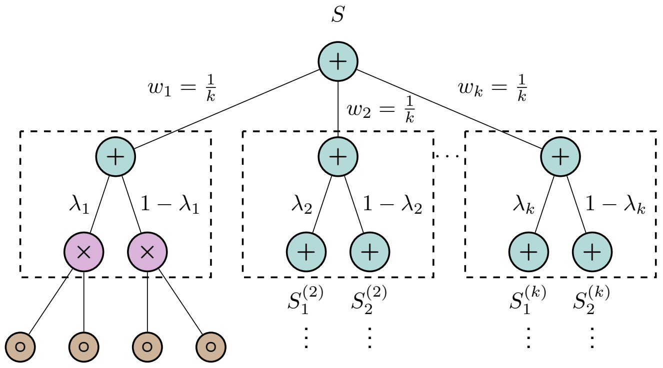
SPLITMAX



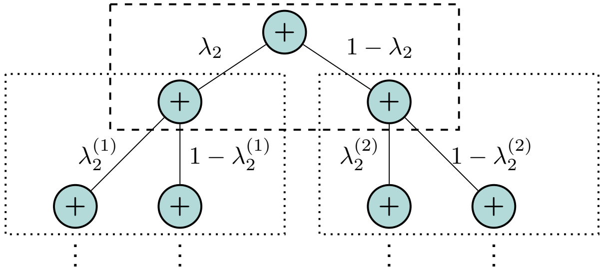
SPLITSID



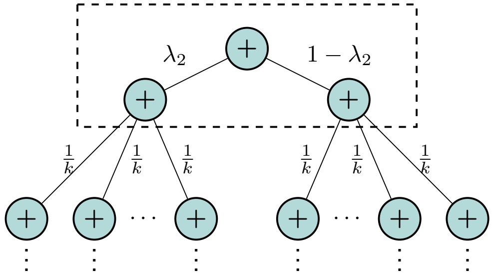
What if we use RPs as sum nodes in PCs?

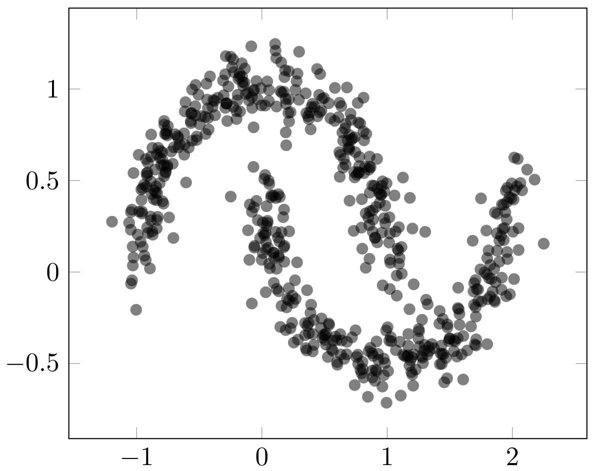


Sum nodes as RPs:

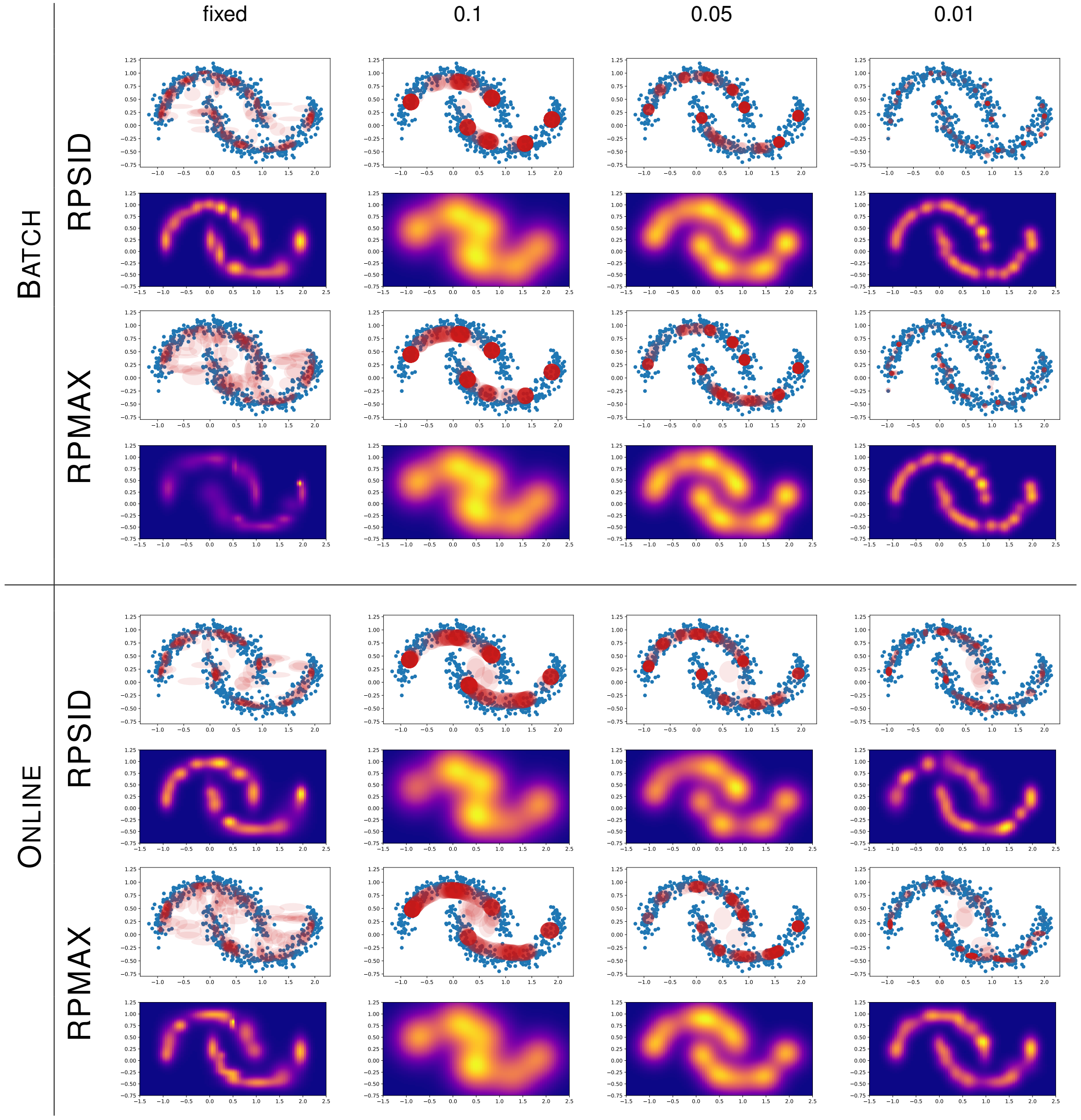


Mixtures of RPs:





EM	$\sigma^2 \geq$	RPMaxS	RPMax	RPSIDS	RPSID
Batch	fixed	-1.083	-1.139	-1.106	-1.074
	0.10	-1.616	-1.616	-1.616	-1.616
	0.05	-1.387	-1.388	-1.387	-1.385
	0.01	-1.034	-1.041	-1.035	-1.032
Online	fixed	-1.089	-1.119	-1.080	-1.108
	0.10	-1.623	-1.627	-1.624	-1.637
	0.05	-1.416	-1.402	-1.408	-1.412
	0.01	-1.139	-1.080	-1.114	-1.105



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], Liang et al. [2017], Mauro et al. [2021]

Dataset	RPMaxSD	RPMaxD	RPSIDSD	RPSIDD	RPMaxS	RPMax	RPSIDS	RPSID	LSPN	Strudel	LPSDD	EXPC
NLTCS	-6.02	-6.02	-6.03	-6.05	$\lvert$-6.01$\rvert$	$\underline{-6.01}$	-6.01	-6.01	-6.11	-6.06	**-5.99**	-6.05
PLANTS	-13.94	-14.06	-14.00	-13.96	-14.07	-13.86	-14.02	-13.94	**-12.97**	$\underline{-12.98}$	$\lvert$-13.02$\rvert$	-14.19
AUDIO	-40.25	-40.25	$\lvert$-40.20$\rvert$	$\underline{-40.20}$	-40.29	-40.25	-40.28	-40.23	-40.50	-41.50	**-39.94**	-40.91
JESTER	$\lvert$-53.05$\rvert$	$\underline{-52.99}$	-53.09	-53.05	-53.21	-53.24	-53.13	-53.06	-75.98	-55.03	**-51.29**	-53.43
NETFLIX	-57.20	-57.21	-57.22	$\underline{-56.89}$	-57.53	-57.40	-57.45	-57.44	$\lvert$-57.02$\rvert$	-58.69	**-55.71**	-57.58
ACCIDENTS	-36.77	-36.67	-36.96	-36.47	-37.41	-37.48	-37.16	-37.39	$\underline{-30.03}$	**-28.73**	$\lvert$-30.16$\rvert$	-31.02
BOOK	-34.84	-34.92	-34.87	-34.82	$\lvert$-34.75$\rvert$	-34.85	$\underline{-34.74}$	**-34.66**	-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	**-82.52**	$\underline{-86.22}$	-88.01	$\lvert$-86.61$\rvert$
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	–