Smoothing Structured Decomposable Circuits

Andy Shih[1], Guy Van den Broeck[2], Paul Beame[3], Antoine Amarilli[4]

[1]Stanford University, [2]University of California, Los Angeles, [3]University of Washington, [4]LTCI Télécom Paris, IP Paris

Probabilistic circuits

Are state-of-the-art for:

- Inference algorithms
  - ✳ PGMs
  - ✳ Probabilistic Programs
- Discrete density estimation

Can encode complex distributions

Can answer advanced queries

but only if they are smooth!

	SPN	AC	PSDD
Decomposability	✓	✓	✓(S)
Determinsim	✗	✓	✓
Smoothness	✓	✓	✓
Pr(evid)	✓	✓	✓
Marginal	✓	✓	✓
MPE	✗	✓	✓
Marginal MAP	✗	✗	✓*
Expectation	✗	✗	✓*

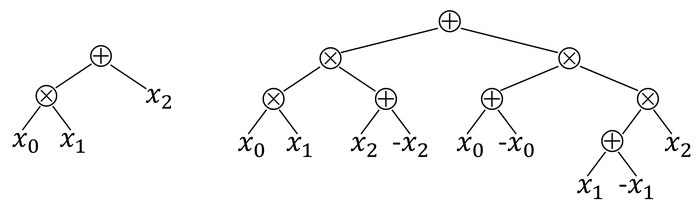
Results

Task	Operations	Complexity
Smoothing	$\oplus, \otimes$	$O(m \cdot \alpha(m, n))$
Smoothing*	$\oplus, \otimes$	$\Omega(m \cdot \alpha(m, n))^*$
All-Marginal	$\oplus, \ominus, \otimes, \oslash$	$\Theta(m)$

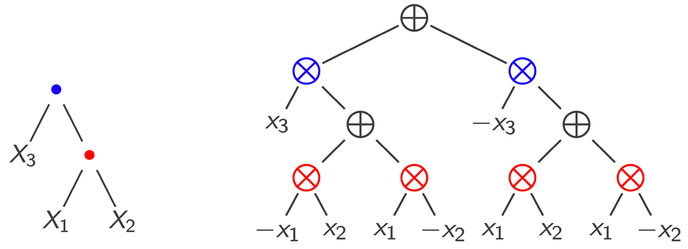
Definition

A circuit is **smooth** if for every pair of children c_1 and c_2 of a $\oplus$-gate, $vars_{c_1} = vars_{c_2}$.

Not Smooth ***Smooth***



Structured Decomposability

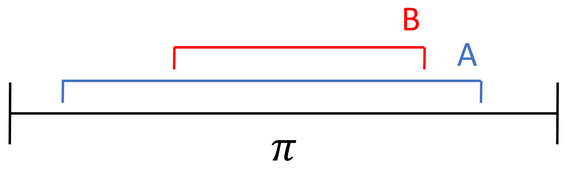


(a) vtree v over $X_1X_2X_3$ (b) circuit decomposes according to v

Missing variables: 2 intervals

Traverse vtree in-order to get ordering π.

Variables of a subtree is an interval in π.



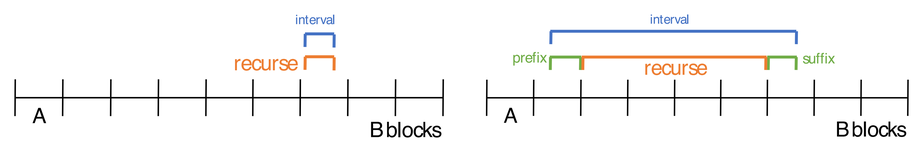
A\B forms 2 intervals

Semigroup Range-Sum

Theorem

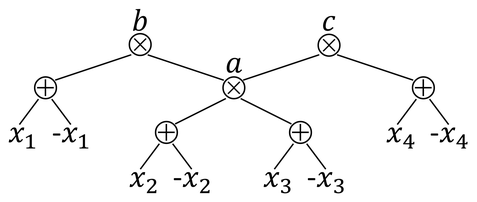
Given n variables defined over a semigroup and m intervals, the sum of all intervals can be computed using $O(m \cdot \alpha(m, n))$ additions [Chazelle and Rosenberg 1989].

Split N variables into B blocks of size A



Trace additions using circuits

```
a = z(x_2) + z(x_3)
b = a + z(x_1)
c = a + z(x_4)
output b, c
```



Efficient smoothing / all-marginals

Table 2: Experiments on smoothing hand-crafted circuits and experiments on computing All-Marginals as part of the collapsed sampling algorithm. Sizes are reported in thousands (k).

(a) Time (in seconds) taken to smooth circuits.

Size	Naive	Ours	Speedup ×
40k	0.82 ± 0.01	0.04 ± 0.01	21 ± 1
416k	50 ± 0.3	0.31 ± 0.01	161 ± 6
1,620k	293 ± 2	0.74 ± 0.04	390 ± 30
8,500k	6050 ± 20	4.13 ± 0.09	1470 ± 40

(b) Number of $\oplus, \ominus, \otimes, \oslash$ operations to compute All-Marginals when sampling the Segmentation-11 network.

Size	Naive	Ours	Impr %
100k	$28{,}494 \pm 598$	$20{,}207 \pm 411$	29 ± 3
200k	$55{,}875 \pm 1{,}198$	$36{,}101 \pm 1{,}522$	35 ± 5
400k	$86{,}886 \pm 6{,}330$	$56{,}094 \pm 817$	35 ± 6