

Robust Analysis of MAP Inference in Selective Sum-Product Networks

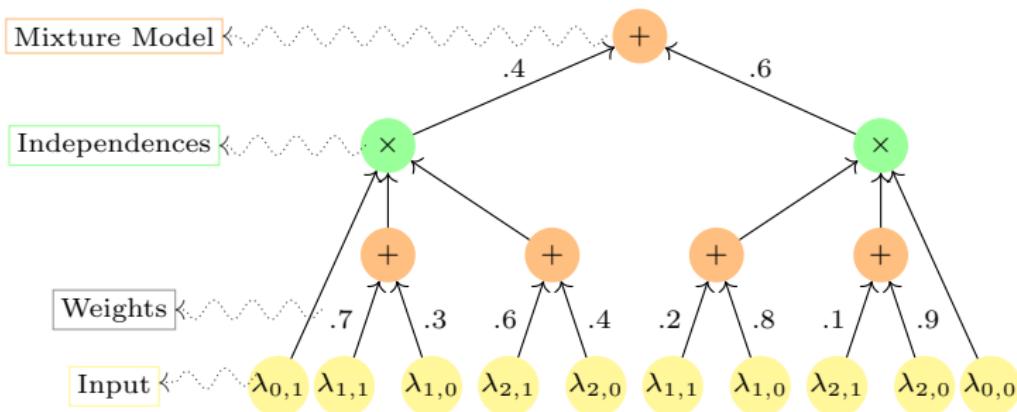
Julissa Villanueva Llerena Denis Deratani Mauá

Institute of Mathematics and Statistics
University of São Paulo
São Paulo, Brazil



Sum-Product Networks

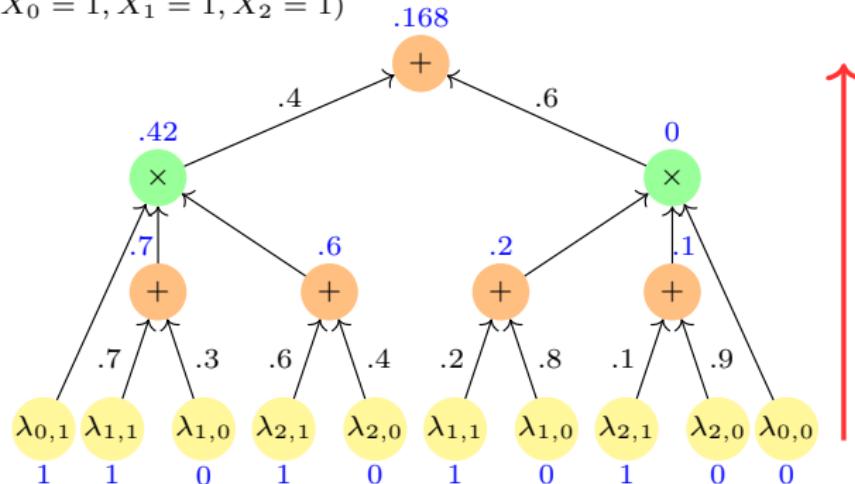
Neural Networks with clear Probabilistic Semantics
(non-linearity due to products)



Sum-Product Networks

- Tractable marginal probability inference
- Intractable Maximum-A-Posteriori (MAP) Inference

$$\mathbb{P}(X_0 = 1, X_1 = 1, X_2 = 1)$$



Sum-Product Networks Applications

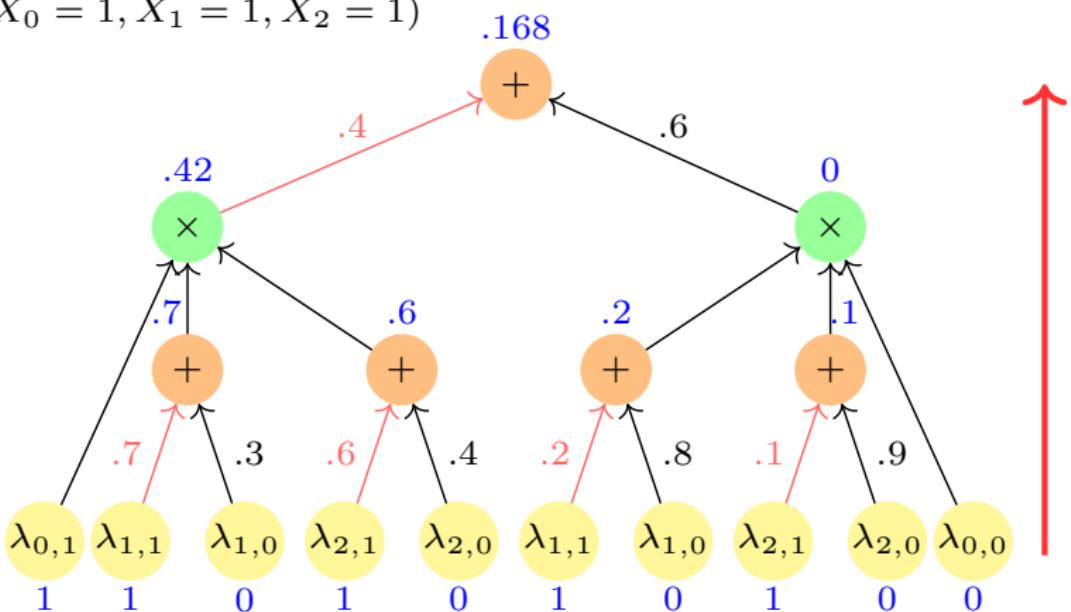


Require (approximate) MAP Inference

ISIP

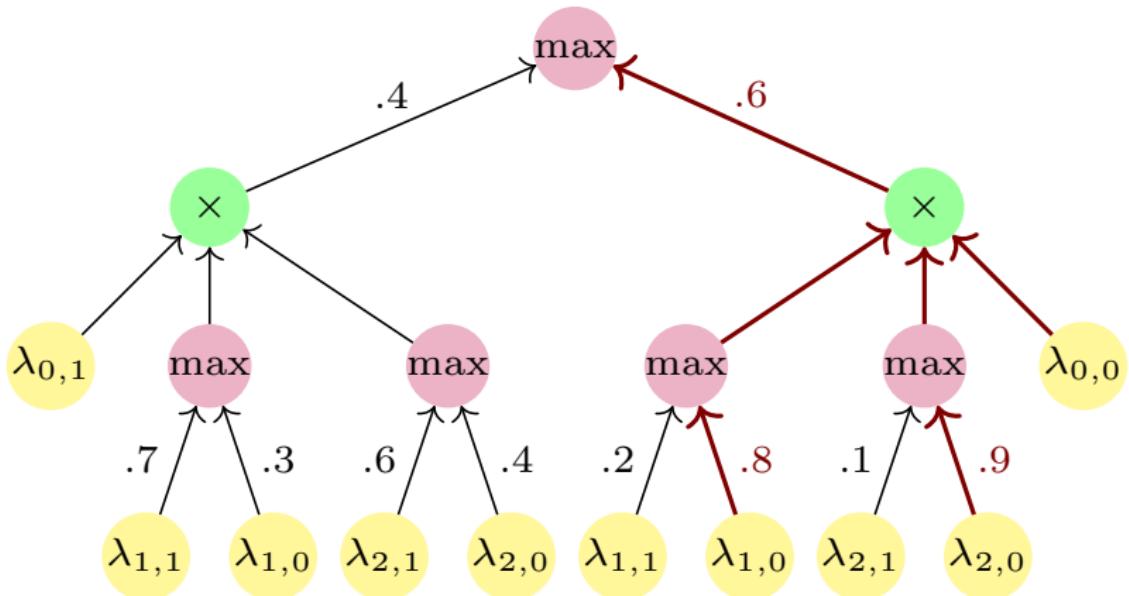
Selective Sum-Product Networks

$$\mathbb{P}(X_0 = 1, X_1 = 1, X_2 = 1)$$



MAP inference Selective Sum-Product Networks

$$\max_{\mathbf{x}} \mathbb{P}(\mathbf{x}) : X_0 = 0, X_1 = 0, X_2 = 0$$



Robust Analysis in Sum-Product Networks

Goal: Distinguish between robust and non-robust MAP inferences

Robustness is assessed via **imprecise mixtures**:

- ϵ -contamination:

$$\mathcal{C}_{\mathbf{w}_i, \epsilon} = \left\{ (1 - \epsilon)\mathbf{w}_i + \epsilon\mathbf{v} : v_j \geq 0, \sum_j v_j = 1 \right\}$$

- Imprecise Dirichlet Model:

$$\mathcal{C}_{N,s} = \left\{ \mathbf{w}_i : w_{ij} = \frac{N_j + s \cdot v_i}{N_i + s}, v_j \geq 0, \sum_j v_j = 1 \right\}$$

Robust Analysis in Sum-Product Networks

Credal Sum-Product Networks

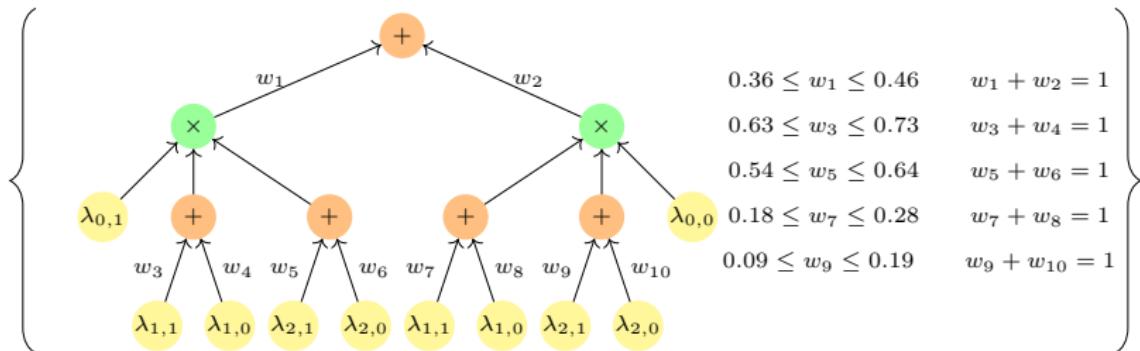


Figura: Credal SPN obtained by 0.1-contamination

Robust Analysis of MAP Inference in Selective Sum-Product Networks

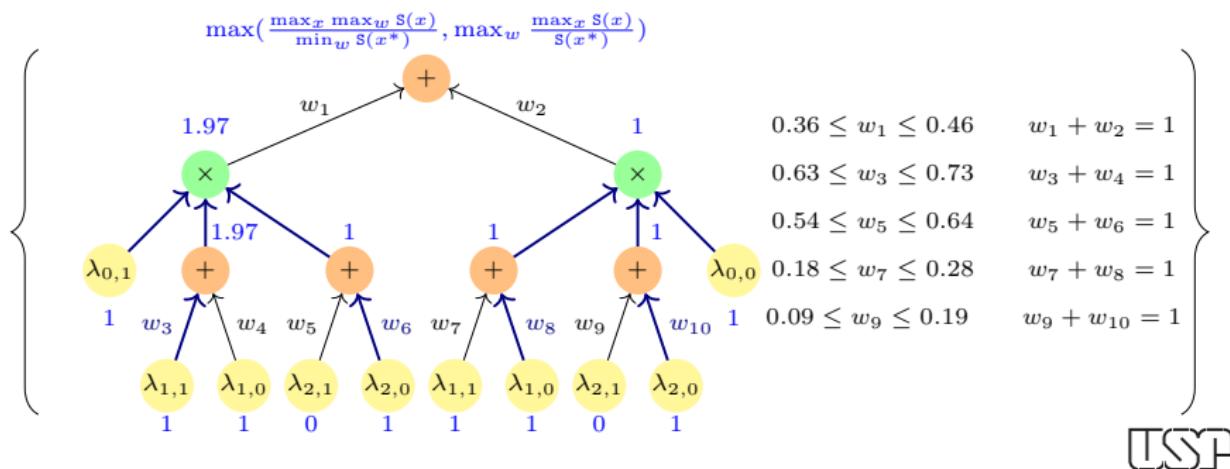
MAP configuration \mathbf{x}^* is **robust** if:

$$\max_{\mathbf{x} \neq \mathbf{x}^*} \max_{\mathbf{w} \in \mathcal{C}} \left(\frac{S_{\mathbf{w}}(\mathbf{x}, \mathbf{e})}{S_{\mathbf{w}}(\mathbf{x}^*, \mathbf{e})} \right) < 1 .$$

I.e. \mathbf{x}^* is still MAP inference in any perturbed SPN

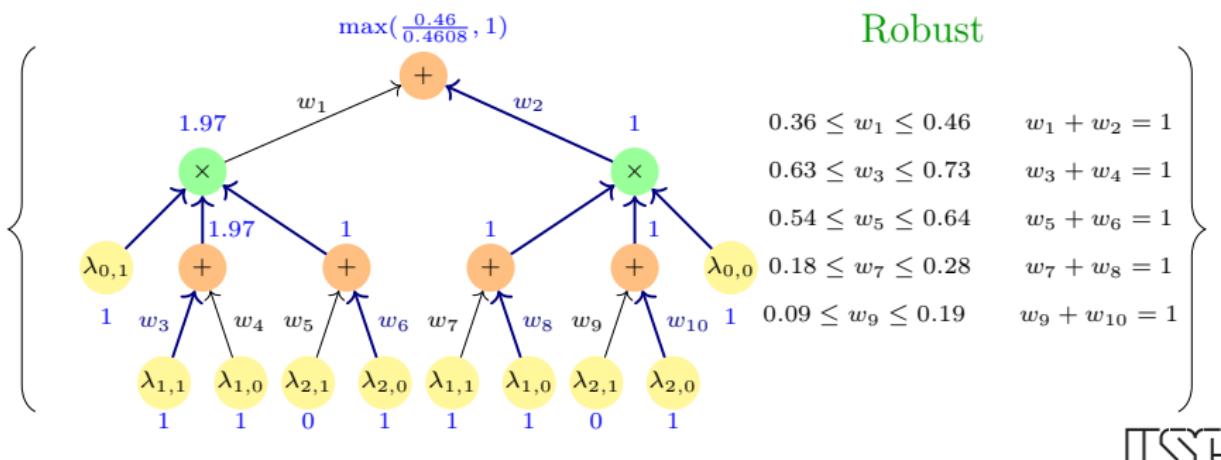
Robust Analysis of MAP Inference in Selective Sum-Product Networks

Let $\mathbf{x}^* : X_0 = 0, X_1 = 0$ be the MAP inference and $e : X_2 = 0$ be the evidence, we decide if \mathbf{x}^* is robust for 0.1 contamination as follow:

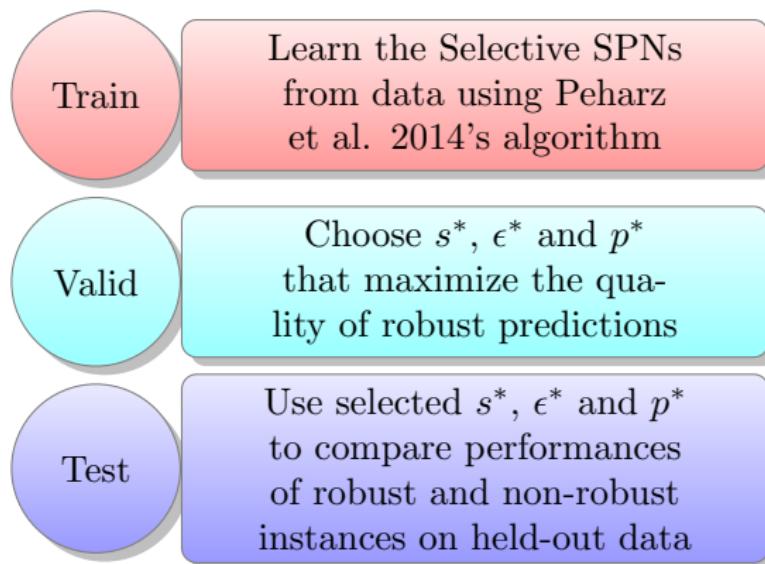


Robust Analysis of MAP Inference in Selective Sum-Product Networks

Let $\mathbf{x}^* : X_0 = 0, X_1 = 0$ be the MAP inference and $e : X_2 = 0$ be the evidence, we decide if \mathbf{x}^* is robust for 0.1 contamination as follow:



Experiments



Results: Multilabel Classification

Dataset	Accuracy			Exact Match			
	<i>Robust</i>	\neg <i>Robust</i>	Δ <i>Acc</i>	<i>Robust</i>	\neg <i>Robust</i>	Δ <i>EM</i>	
Arts	ϵ^* <i>s*</i> <i>p*</i>	0.88 0.107 0.81	0.196 0.351 0.159	0.634 -0.244 0.651	0.833 0.089 0.75	0.143 0.247 0.107	0.69 -0.158 0.643
	ϵ^* <i>s*</i> <i>p*</i>	0.751 0.781 0.762	0.582 0.641 0.581	0.169 0.14 0.181	0.617 0.642 0.62	0.598 0.469 0.392	0.019 0.173 0.228
	ϵ^* <i>s*</i> <i>p*</i>	0.595 0.686 0.574	0.41 0.413 0.391	0.185 0.273 0.183	0.238 0.308 0.176	0.163 0.16 0.176	0.075 0.148 0
Flags	ϵ^* <i>s*</i> <i>p*</i>	0.917 0.917 0.917	0.468 0.468 0.468	0.449 0.449 0.449	0.5 0.5 0.5	0.118 0.1 0.118	0.382 0.4 0.382
	ϵ^* <i>s*</i> <i>p*</i>	0.667 0.637 0.655	0.557 0.482 0.552	0.11 0.155 0.103	0.5 0.537 0.552	0.416 0.304 0.409	0.084 0.233 0.143
	ϵ^* <i>s*</i> <i>p*</i>	0.203 0.203 0.211	- - 0.198	- - 0.013	0.146 0.146 0.155	- - 0.14	- - 0.015
Human	ϵ^* <i>s*</i> <i>p*</i>	0.331 0.367 0.345	0.217 0.205 0.212	0.114 0.205 0.132	0.324 0.362 0.338	0.213 0.205 0.208	0.111 0.205 0.13
	ϵ^* <i>s*</i> <i>p*</i>	0.857 0.929 0.923	0.277 0.293 0.276	0.58 0.636 0.647	0.857 0.929 0.923	0.212 0.293 0.211	0.645 0.636 0.712
	ϵ^* <i>s*</i> <i>p*</i>	0.436 0.436 0.425	0.419 0.419 0.203	0.017 0.017 0.222	0.071 0.071 0.099	0.098 0.098 0	-0.027 -0.027 0.099

Results: Density Estimation

Dataset	Exact Match			Hamming Score			
	<i>Robust</i>	\neg <i>Robust</i>	ΔEM	<i>Robust</i>	\neg <i>Robust</i>	ΔHS	
Jester	ϵ^*	0.259	0.001	0.258	0.913	0.692	0.221
	s^*	0.076	.001	0.075	0.755	0.692	0.063
	p^*	0.016	0.001	0.015	0.858	0.687	0.171
NLTCs	ϵ^*	0.736	0.248	0.488	0.934	0.77	0.164
	s^*	0.736	0.248	0.488	0.934	0.77	0.164
	p^*	0.736	0.248	0.488	0.933	0.77	0.163
MSNBC	ϵ^*	1	0.247	0.753	1	0.783	0.217
	s^*	0.464	0.215	0.249	0.875	0.769	0.106
	p^*	0.727	0.247	0.48	0.92	0.783	0.137

See you at the Poster Session!

Global Sensitivity Analysis of MAP Inference in Selective Sum-Product Networks

Julissa Villanueva Llerena and Denis Derafchi Mazi
Departments of Computer Science, University of São Paulo, Brazil
gdlle@icmc.usp.br

1. Introduction

- **Sum-Product Networks**: deep generative probabilistic models with tractable inference by second machine-learning module.
- **Qualitative Global Sensitivity Analysis**: a method to analyze how changes in input variables affect the output probability.
- **Qualitative Global Sensitivity Analysis of MAP**: a method to analyze how changes in input variables are propagated independently.
- **Goal**: Trustable global sensitivity analysis of neural prediction.

2. Selective Sum-Product Networks & MAP Inference

A **Selective SPN** is one of the following:

- An indicator variable $\delta_{\{x_1, x_2, \dots, x_n\}}$ mapping every $\{x_1, x_2, \dots, x_n\}$ to either 0 or 1;
- A product $\prod_{i=1}^n \delta_{\{x_i\}}$ of indicator SPNs with degree one;
- A weighted sum $\sum_{i=1}^n w_i \delta_{\{x_i\}}$ of indicator SPNs, where $\sum_i w_i = 1$, $w_i \geq 0$, and $w_i \neq 0$.

Maximum a Posteriori (MAP) Inference

Given $E(X)$ representing joint distribution $P(X|Z)$, and $\pi^*(X)$ representing $P(Z|X)$.
 This is the mode of the posterior probability. $\text{MAP}(X) = \arg\max_{X \in \mathcal{X}} P(X|Z)$

3. Global Sensitivity Analysis

Let Δ_w be an SPP whose weight vector is w . **Qualitative SPP** is defined as follows. The SPP is represented by $\delta_{\{w_1, w_2, \dots, w_n\}}$ with node-level speed: $\langle \Delta_w, w \cdot C \rangle$ obtained by:

- $\langle \Delta_w, w \cdot C \rangle = \sum_{i=1}^n w_i \langle \delta_{\{x_i\}}, w \cdot C \rangle$ if $w_i \neq 0$;
- $\langle \Delta_w, w \cdot C \rangle = \sum_{i=1}^n w_i \langle \delta_{\{x_i\}}, w \cdot C \rangle / \sum_{i=1}^n w_i$ if $w_i = 0$.

4. Robust MAP Inference

A **MAP** definition $\pi^*(X)$ is robust with respect to a small SPP Δ_w , $w \in \mathbb{C}$, if:

$$\left| \frac{\pi^*(X)}{\pi^*(X') - \pi^*(X)} \right| \leq \frac{\left| \frac{\pi^*(X)}{\pi^*(X)} \right|}{1 + \left| \frac{\pi^*(X)}{\pi^*(X)} \right|} \quad \text{for all } X, X' \in \mathcal{X} \text{ such that } \Delta_w \subset \Delta_{w'} \text{ and } \Delta_{w'} \subset \Delta_{w''}$$

5. Results

Dataset	Model	MAP	Qualitative GSA	Qualitative GSA + Robust MAP
Adult	MAP	0.62	0.62	0.62
Adult	Qualitative GSA	0.62	0.62	0.62
Adult	Qualitative GSA + Robust MAP	0.62	0.62	0.62
Iris	MAP	0.97	0.97	0.97
Iris	Qualitative GSA	0.97	0.97	0.97
Iris	Qualitative GSA + Robust MAP	0.97	0.97	0.97
Wine	MAP	0.87	0.87	0.87
Wine	Qualitative GSA	0.87	0.87	0.87
Wine	Qualitative GSA + Robust MAP	0.87	0.87	0.87
Spambase	MAP	0.99	0.99	0.99
Spambase	Qualitative GSA	0.99	0.99	0.99
Spambase	Qualitative GSA + Robust MAP	0.99	0.99	0.99
Image	MAP	0.99	0.99	0.99
Image	Qualitative GSA	0.99	0.99	0.99
Image	Qualitative GSA + Robust MAP	0.99	0.99	0.99

6. References and Acknowledgments

[1] Carlini, Michael, Levine, and David Balduzzi. "Model uncertainty analysis." In *NeurIPS*, pp. 1–10. 2018.

[2] Derafchi, M., Villanueva Llerena, and Denis Derafchi Mazi. "Qualitative Global Sensitivity Analysis of Selective Sum-Product Networks." In *NeurIPS*, 2020.

[3] Derafchi, M., Villanueva Llerena, and Denis Derafchi Mazi. "Qualitative Global Sensitivity Analysis of Selective Sum-Product Networks." *International Journal of Approximate Reasoning*, 120, 1–16, 2020.

[4] Derafchi, M., Villanueva Llerena, and Denis Derafchi Mazi. "Qualitative Global Sensitivity Analysis for a family of neural networks." In *NeurIPS*, 2020.