Probabilistic Watershed:
Sampling all spanning forests for seeded segmentation and semi-supervised learning

Enrique Fita Sanmartín, Sebastian Damrich, Fred A. Hamprecht

HCI/IWR at Heidelberg University

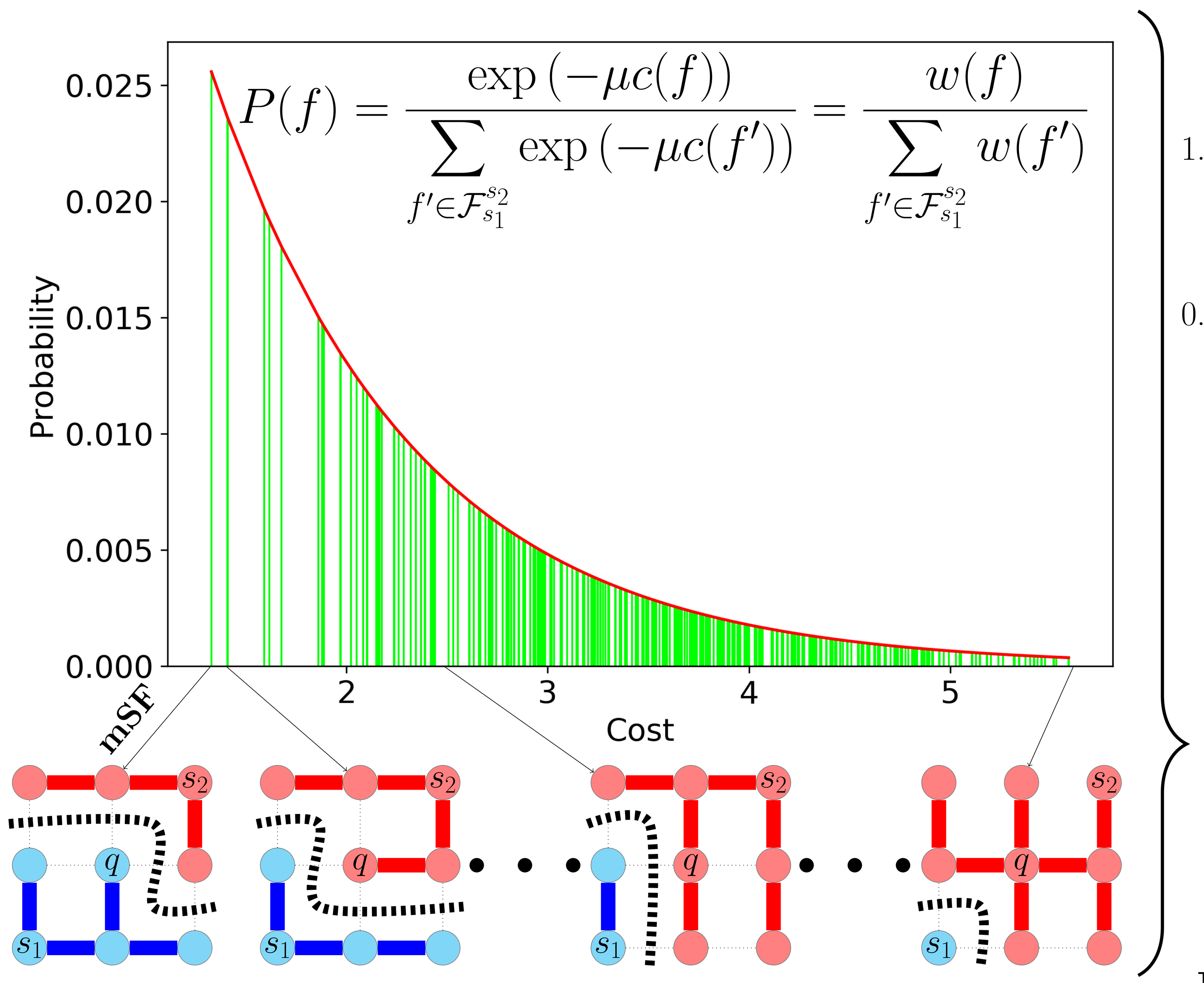
Approach

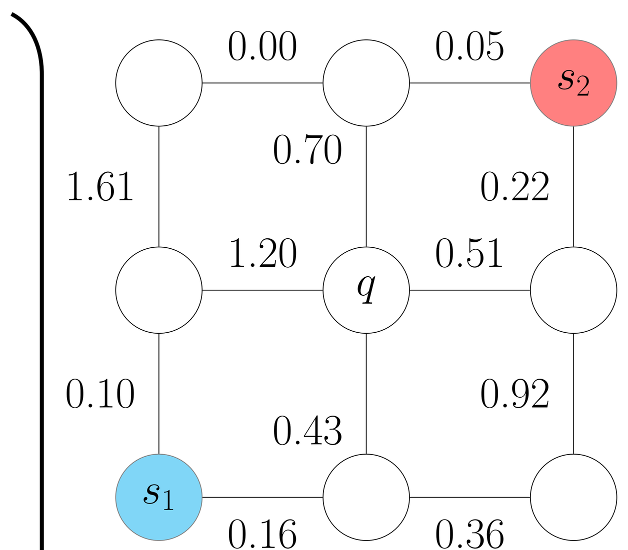
Question

What is the probability of sampling a Gibbs-distributed forest such that a node of interest is connected to a certain seed? What is the relation between the probability and the cost of a forest?

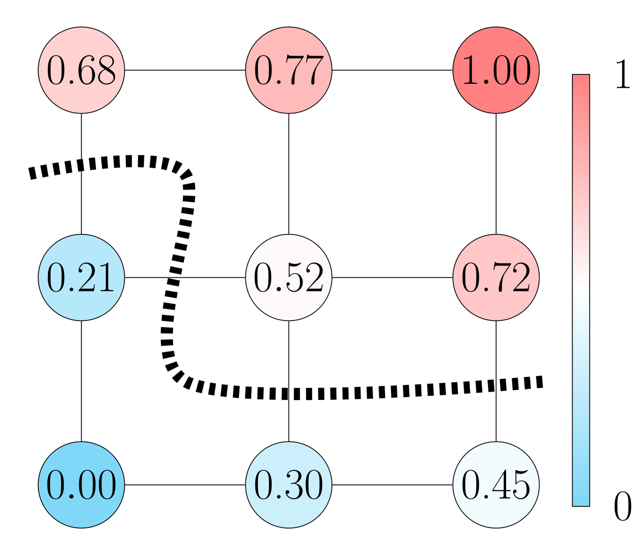
- Seeded **Watershed** computes a minimum spanning cost forest (**mSF**)→'winner-takes-all".
- The **Probabilistic Watershed** (**ProbWS**) takes all spanning forests into account according to their cost by means of a Gibbs distribution over the spanning forests.

$$P(q\sim s_1) := \sum_{f\in\mathcal{F}^{s_2}_{s_1,q}} P(f) = \frac{\sum_{f\in\mathcal{F}^{s_2}_{s_1,q}} w(f)}{\sum_{f'\in\mathcal{F}^{s_2}_{s_1}} w(f')} = \frac{w\left(\mathcal{F}^{s_2}_{s_1,q}\right)}{w\left(\mathcal{F}^{s_2}_{s_1}\right)}.$$



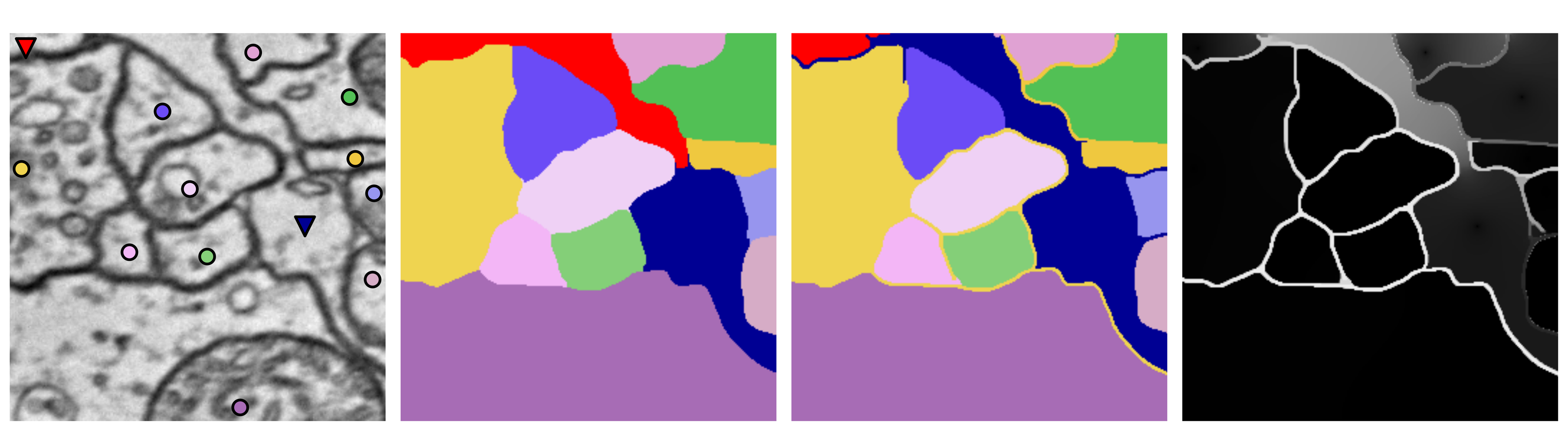


Graph with edge-costs



ProbWS Probabilities

- Applications: **Seeded segmentation**.



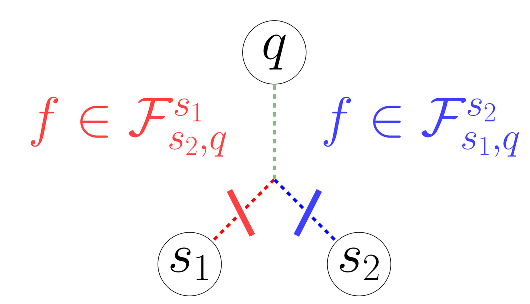
(a) Image with seeds (b) Watershed (c) ProbWS (d) Uncertainty

Computation $w(\mathcal{F}^{s_2}_{s_1})$, $w(\mathcal{F}^{s_1}_{s_2,q})$ and $w(\mathcal{F}^{s_2}_{s_1,q})$

- r^{eff}_{uv} := Effective resistance distance between the nodes u and v.
- As a consequence of the **Matrix Tree Theorem** [3] we are able to find a **closed-form computation of the weight of the set of forests**:

$$w(\mathcal{F}^{s_2}_{s_1}) = w(\mathcal{T})r^{\text{eff}}_{s_1s_2} \propto r^{\text{eff}}_{s_1s_2}.$$

- We define the following linear system:



$$\begin{aligned} w(\mathcal{F}^{s_2}_{s_1,q}) + w(\mathcal{F}^{q}_{s_1,s_2}) &= w(\mathcal{F}^{q}_{s_2}) \\ w(\mathcal{F}^{q}_{s_1,s_2}) + w(\mathcal{F}^{s_1}_{s_2,q}) &= w(\mathcal{F}^{q}_{s_1}) \\ w(\mathcal{F}^{s_2}_{s_1,q}) + w(\mathcal{F}^{s_1}_{s_2,q}) &= w(\mathcal{F}^{s_2}_{s_1}). \end{aligned}$$

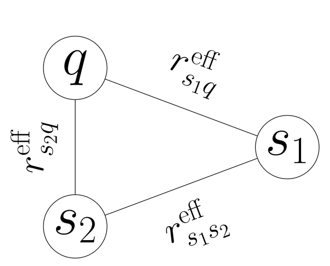
$$f\in\mathcal{F}^{s_2}_{s_1} = \mathcal{F}^{s_1}_{s_2,q} \cup \mathcal{F}^{s_2}_{s_1,q}$$

Closed-form ProbWS

$$P(q\sim s_1) = \frac{w(\mathcal{F}^{q}_{s_2}) + w(\mathcal{F}^{s_2}_{s_1}) - w(\mathcal{F}^{q}_{s_1})}{2w(\mathcal{F}^{s_2}_{s_1})} = \frac{r^{\text{eff}}_{s_2q} + r^{\text{eff}}_{s_2s_1} - r^{\text{eff}}_{s_1q}}{2r^{\text{eff}}_{s_2s_1}}.$$

.

- ProbWS is proportional to the effective resistance's triangle inequality gap.



$$P(q\sim s_1) \geq P(q\sim s_2) \iff r^{\text{eff}}_{s_1q} \leq r^{\text{eff}}_{s_2q}$$

Random Walker = ProbWS

The probability, $x^{s_2}_q$, that a Random Walker starting at node q reaches first s_2 before reaching s_1 ([2]) is equal to the probability defined by the ProbWS.

$$x^{s_2}_q = P(q\sim s_2).$$

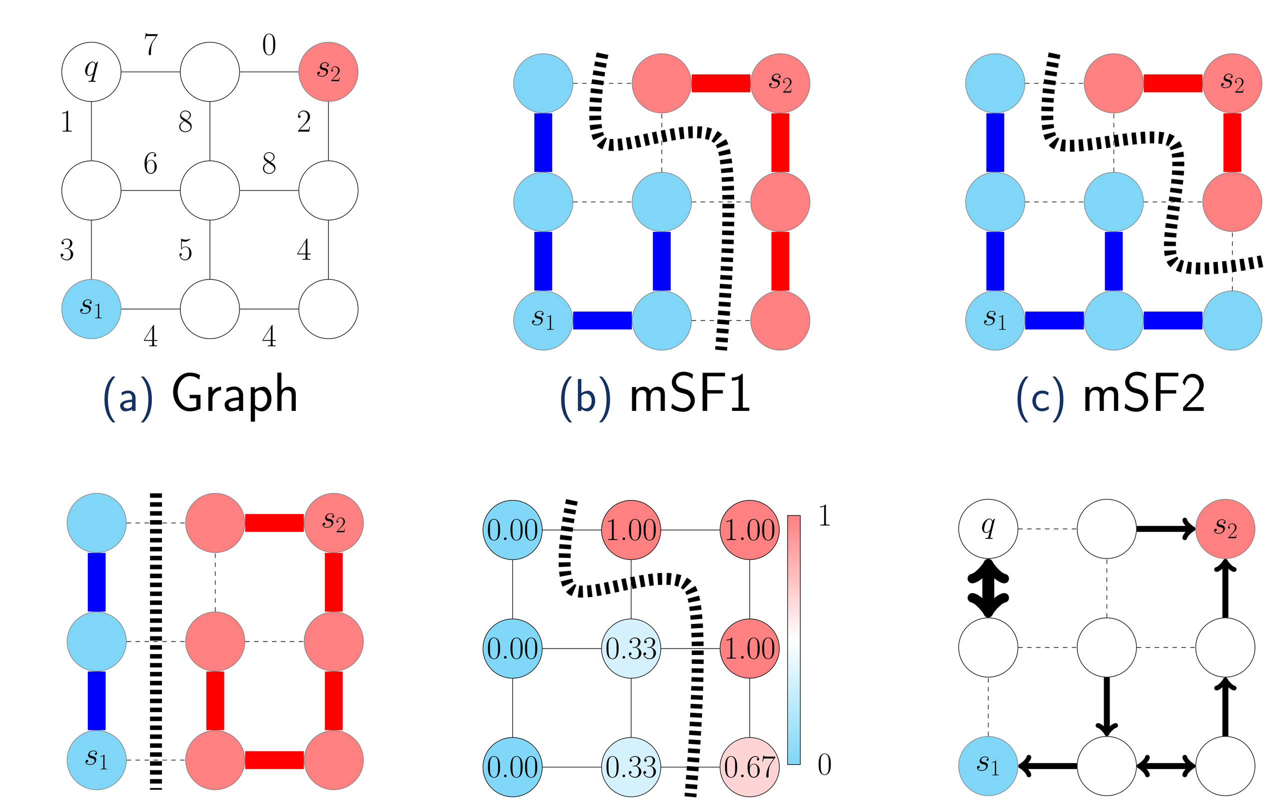
Power Watershed

Power Watershed [1] was introduced as a limit of the Random Walker.

Power Watershed counts mSF

Given $S=\{s_1,s_2\}$ a set of seeds, let us denote the potential of node q being assigned to seed s_1 by the Power Watershed[1] with $\beta=2$ as $x^{\text{PW}_1}_q$. Let further $c_{\min}$ be $\min_{f\in\mathcal{F}^{s_2}_{s_1}} c(f)$. Then

$$x^{\text{PW}_1}_q = P_\infty(q\sim s_1) := \frac{\left|\{f\in\mathcal{F}^{s_2}_{s_1,q} \;:\; c(f)=c_{\min}\}\right|}{\left|\{f\in\mathcal{F}^{s_2}_{s_1} \;:\; c(f)=c_{\min}\}\right|}.$$



(a) Graph (b) mSF1 (c) mSF2 (d) mSF3 (e) $P_\infty(\text{node}\sim s_2)$ (f) RW reachability

Contributions

- Analytical calculus of the Probabilistic Watershed probability, i.e. the probability that a node is assigned to a particular seed in an ensemble of Gibbs distributed spanning forests.
- Probabilistic Watershed (ProbWS) = Random Walker (RW).
- Explicit relation between effective resistance and ProbWS/RW.
- Power Watershed counts minimum cost spanning forests.

References

[1] C. Couprie, L. Grady, L. Najman, and H. Talbot. Power watershed: A unifying graph-based optimization framework. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.

[2] L. Grady. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006.

[3] G. Kirchhoff. Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Annalen der Physik*, 148:497–508, 1847.

Web

Paper