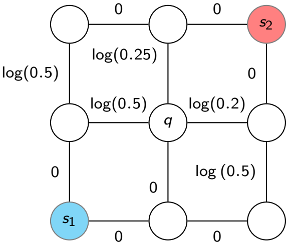
# Probabilistic Watershed: Sampling all spanning forests for seeded segmentation and semi-supervised learning


Enrique Fita Sanmartín, Sebastian Damrich, Fred A. Hamprecht

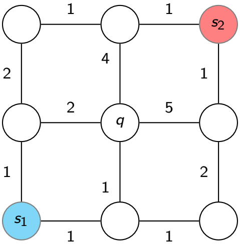
HCI/IWR at Heidelberg University

# Notation: edge cost vs edge weight

$$w(e) = \exp(-\mu c(e)), \ \ \mu \geq 0$$




(a) Graph's edge-costs

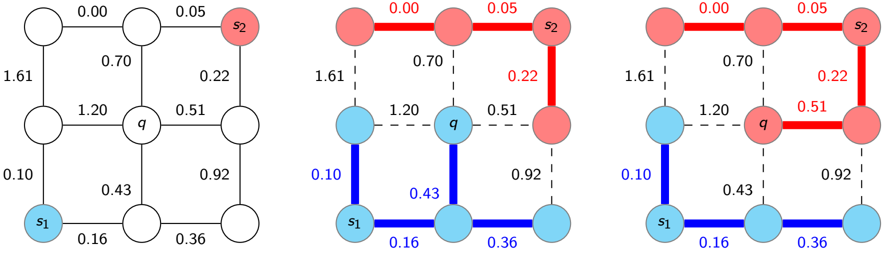



(b) Graph's edge-weights, $\mu = 1$

$$c(G) = \sum_{e \in E_G} c(e) \qquad\qquad w(G) = \prod_{e \in E_G} w(e)$$

# Watershed computes a minimum cost spanning forest (mSF)

## The assignment of $q$ is doubtful




(a) Graph with two seeds　(b) mSF: Cost=1.32　(c) SF: Weight=1.4

Figure: (2a) Graph with edge-costs and two seeds. (2b)The mSF (Watershed segmentation) assigns $q$ to seed $s_1$. (2c) Low cost, but not minimum, spanning forest connecting $q$ to $s_2$. The Watershed has a winner-takes-all behaviour. It only considers the mSF while ignores the rest of seed separating spanning forests

# Question to solve

What is the probability of sampling a forest such that a node of interest is assigned to a certain seed?

# Gibbs Distribution of $\mathcal{F}_{s_1}^{s_2}$

$$\mathcal{F}_{s_1}^{s_2} := \{\text{2-forests separating } s_1 \text{ and } s_2\}$$

**Definition: probability distribution over the forests**

Given $J \in \mathbb{R}_{\geq 0}$

$$\Pr{}^* = \arg\min_{\Pr} \sum_{f \in \mathcal{F}_{s_1}^{s_2}} \Pr(f) c(f), \quad \text{s.t.} \quad \sum_{f \in \mathcal{F}_{s_1}^{s_2}} \Pr(f) = 1 \text{ and } \mathcal{H}(\Pr) = J, \tag{1}$$

where $\mathcal{H}(\Pr)$= entropy Pr

**Solution**: Gibbs Probability distribution

$$\Pr{}^*(f) = \frac{\exp(-\mu c(f))}{\sum_{f' \in \mathcal{F}_{s_1}^{s_2}} \exp(-\mu c(f'))} = \frac{w(f)}{\sum_{f' \in \mathcal{F}_{s_1}^{s_2}} w(f')}$$

# Probabilistic Watershed

$$\mathcal{F}_{s_1}^{s_2} := \{\text{2-forests separating } s_1 \text{ and } s_2\}$$
$$\mathcal{F}_{s_1,q}^{s_2} := \{f \in \mathcal{F}_{s_1}^{s_2} \;:\; s_1 \text{ and } q \text{ are connected}\}$$

## Probability $q$ is connected with $s_1$

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

# Example

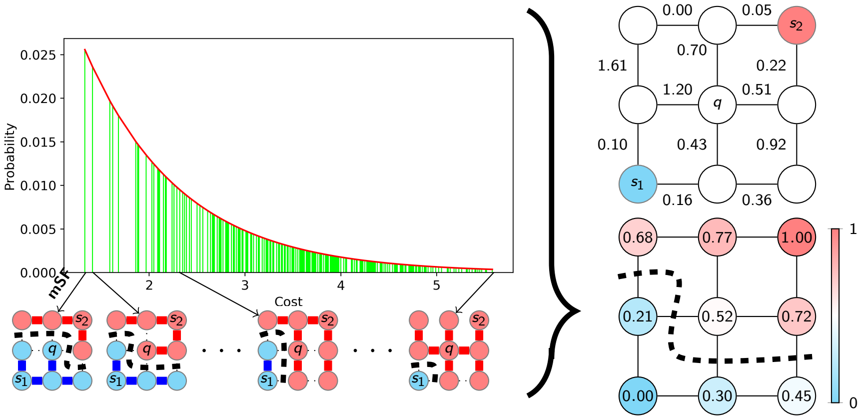



Figure: (**Top left**) Gibbs distribution over all spanning forests. (**Bottom right**) Probabilistic Watershed probabilities for assigning a node to $s_2$. The Probabilistic Watershed computes the expected seed assignment of every node for a Gibbs distribution over all exponentially many spanning forests in closed-form. It thus avoids the winner-takes-all behaviour of the Watershed.

$$\mathcal{F}_{s_1}^{s_2} := \{\text{2-forests separating } s_1 \text{ and } s_2\}$$
$$\mathcal{F}_{s_1,q}^{s_2} := \{f \in \mathcal{F}_{s_1}^{s_2} \;:\; s_1 \text{ and } q \text{ are connected}\}$$

**Probability $q$ is connected to $s_1$**

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

**Question**: How do we compute $w(\mathcal{F}_{s_1}^{s_2})$ and $w(\mathcal{F}_{s_1,q}^{s_2})$?
**Answer:** Matrix Tree Theorem (MTT).

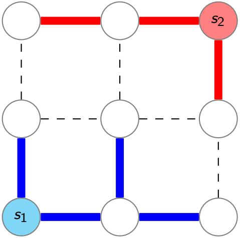
## Matrix Tree Theorem (MTT)

For any weighted multigraph $G$, the sum of the weights of the spanning trees of $G$, $w(\mathcal{T})$, is equal to

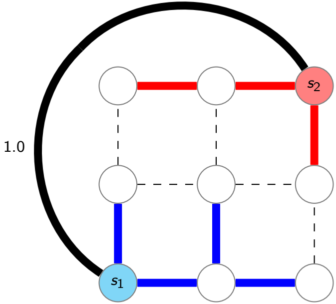
$$w(\mathcal{T}) := \sum_{t \in \mathcal{T}} \prod_{e \in E_t} w(e) = \frac{1}{|V|} \det(L + \frac{1}{|V|} \mathbb{1}\mathbb{1}^\top) = \det(L^{[v]}),$$

where $\mathbb{1}$ is the column of 1. $L^{[v]}$ is the matrix obtained from the Laplacian, $L$, after removing the row and column corresponding to an arbitrary node $v$.

The relation between forests and trees allows the use of the MTT.




(a) Forest.




(b) Tree

# Computation of $w(\mathcal{F}_{s_1}^{s_2})$

$r_{uv}^{\text{eff}}$ = effective resistance distance between the nodes $u$ and $v$.

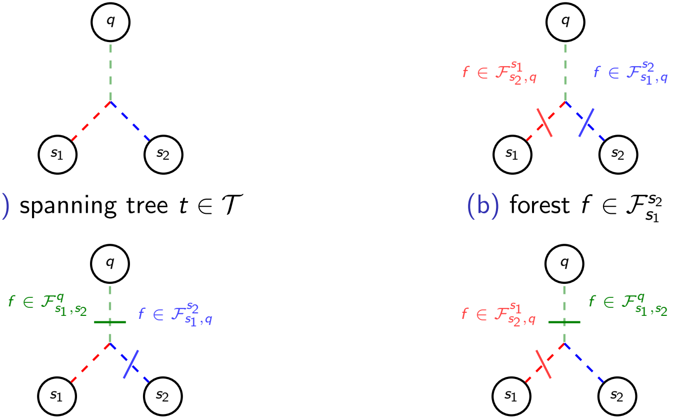
**Theorem (Consequence of the MTT)**

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

# Computation of $w(\mathcal{F}^q_{s_1,s_2})$ and $w(\mathcal{F}^q_{s_2,s_1})$


(a) spanning tree $t \in \mathcal{T}$ (b) forest $f \in \mathcal{F}^{s_2}_{s_1}$

Figure: $\mathcal{F}^{s_2}_{s_1,q}$ and $\mathcal{F}^{s_1}_{s_2,q}$ form a partition of $\mathcal{F}^{s_2}_{s_1}$ since $q$ must be connected either to $s_1$ or $s_2$, but not to both. Given an spanning tree (dashed lines represent all possible spanning trees), to form a forest in $\mathcal{F}^{s_2}_{s_1}$ we will have to remove an edge from the red part, forming a forest in $\mathcal{F}^{s_1}_{s_2,q}$, or from the blue part, forming a forest in $\mathcal{F}^{s_2}_{s_1,q}$.




(a) spanning tree $t \in \mathcal{T}$

(b) forest $f \in \mathcal{F}^{s_2}_{s_1}$

(c) forest $f \in \mathcal{F}^{q}_{s_2}$

(d) forest $f \in \mathcal{F}^{q}_{s_1}$

Figure: Analogously (c) $\mathcal{F}^{q}_{s_2,s_1}$ and $\mathcal{F}^{s_2}_{q,s_1}$ form a partition of $\mathcal{F}^{q}_{s_2}$, and (d) $\mathcal{F}^{q}_{s_1,s_2}$ and $\mathcal{F}^{s_1}_{q,s_2}$ form a partition of $\mathcal{F}^{q}_{s_1}$.

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

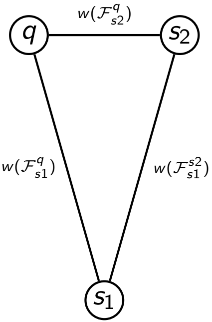
Linear system with 3 unknowns and 3 equations.

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

# Probabilistic Watershed corresponds to the effective resistance distance

## Probability measures the gap of the triangle inequality

$$\Pr(q \sim s_1) = \frac{w(\mathcal{F}_{s_2}^{q}) + w(\mathcal{F}_{s_1}^{s_2}) - w(\mathcal{F}_{s_1}^{q})}{2w(\mathcal{F}_{s_1}^{s_2})} = \frac{r_{s_2 q}^{\text{eff}} + r_{s_2 s_1}^{\text{eff}} - r_{s_1 q}^{\text{eff}}}{2r_{s_2 s_1}^{\text{eff}}}. \qquad (2)$$




$$\Pr(q \sim s_1) \geq \Pr(q \sim s_2) \iff r_{s_1 q}^{\text{eff}} \leq r_{s_2 q}^{\text{eff}}.$$

# Random Walker[Grady, 2006] = Probabilistic Watershed

**Theorem**

The probability, $x_q^{s_2}$, that a Random Walker starting at node $q$ reaches first $s_2$ before reaching $s_1$ ([Grady, 2006]) is equal to the probability defined by the Probabilistic Watershed.

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

# Power Watershed [Couprie et al., 2011]

## Power Watershed minimization problem

$$x^* = \arg\min_x \sum_{e=(v,u)\in E} (w(e))^\alpha \left(|x_v - x_u|\right)^\beta, \text{ s.t. } x_{s_1} = 1,\ x_{s_2} = 0, \quad (3)$$

Random Walker: $\alpha = 1,\ \beta = 2$
Power Watershed: $\alpha \to \infty,\ \beta = 2$

## $\alpha$ and $\mu$ are equivalent

$$w(e) = \exp(-\mu c(e)) \to (w(e))^\alpha = \exp(-\mu\alpha c(e)) = \exp(-\mu_\alpha c(e))$$

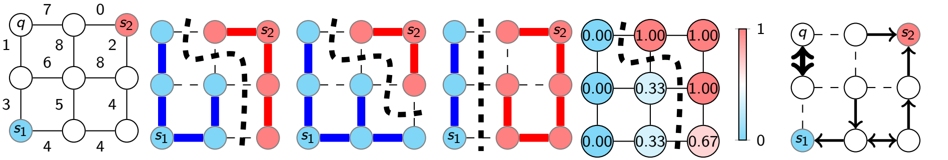
# Power Watershed [Couprie et al., 2011]

## Power Watershed counts minimum cost/maximum weight spanning forests

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 with $\beta = 2$ as $x_q^{\mathrm{PW}_1}$. Let further $w_{\max}$ be $\max_{f \in \mathcal{F}_{s_1}^{s_2}} w(f)$. Then

$$x_q^{\mathrm{PW}_1} = \Pr_\infty(q \sim s_1) := \frac{\left|\{f \in \mathcal{F}_{s_1,q}^{s_2} \ : \ w(f) = w_{\max}\}\right|}{\left|\{f \in \mathcal{F}_{s_1}^{s_2} \ : \ w(f) = w_{\max}\}\right|}.$$

# Power Watershed counts mSF




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

Figure: Forest-interpretation of Power Watershed. The Power Watershed computes the ratio between the mSFs connecting a node to $s_2$ and all possible mSFs. (7f) indicates the allowed Random Walker transitions when $\mu \to \infty$ with headed arrows. The Random Walker interpretation of the Power Watershed breaks down in the limit case since a Random Walker starting at node $q$ does not reach any seed, but oscillates along the bold arrow.

# Conclusion

- Probabilistic Watershed = Random Walker.
- Random Walker/Probabilistic Watershed probabilities are proportional to the triangle inequality gap.

$$\Pr(q \sim s_1) \propto r^{\text{eff}}_{s_2 q} + r^{\text{eff}}_{s_2 s_1} - r^{\text{eff}}_{s_1 q}.$$

- Power Watershed counts minimum cost spanning forests.

# References


Couprie, C., Grady, L., Najman, L., and Talbot, H. (2011).
Power watershed: A unifying graph-based optimization framework.
*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33:1384–1399.

Grady, L. (2006).
Random walks for image segmentation.
*IEEE Transactions on Pattern Analysis and Machine Intelligence*.

# The End