

Properties of combinations of hierarchical watersheds

Deise Santana Maia, Jean Cousty, Laurent Najman, Benjamin Perret

▶ **To cite this version:**

Deise Santana Maia, Jean Cousty, Laurent Najman, Benjamin Perret. Properties of combinations of hierarchical watersheds. Pattern Recognition Letters, 2019, 128, pp.513-520. 10.1016/j.patrec.2019.10.009 . hal-02175824v2

HAL Id: hal-02175824
https://hal.science/hal-02175824v2

Submitted on 6 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Properties of combinations of hierarchical watersheds

Deise Santana Maia[a], Jean Cousty[a], Laurent Najman[a], Benjamin Perret[a]

[a] *Université Paris-Est, LIGM (UMR 8049), CNRS, ENPC, ESIEE Paris, UPEM, F-93162, Noisy-le-Grand, France*

Abstract

In [19], the authors provide an evaluation of combinations of hierarchical watersheds, showing that some combinations outperform individual hierarchical watersheds. In this article, we aim to achieve a deeper understanding of those combinations. We study which of the four combinations evaluated in [19] always result in flattened (simplified) hierarchical watersheds. Then, we provide a sufficient condition for a combination to always output flattened hierarchical watersheds, and a new combining function that outputs flattened hierarchical watersheds. We conclude that, among the combinations assessed in [19], only a particular case of combinations with infimum always result in flattened hierarchical watersheds.

Keywords:
2000 MSC: 41A05, 41A10, 65D05, 65D17 Keyword1, Keyword2, Keyword3

1. Introduction

Hierarchies of segmentations are sequences of nested partitions of image pixels. Since the pioneer work of [31], several methods to build hierarchies of segmentations have been proposed, such as quasi-flat zones hierarchies [27, 24], waterfall and hierarchical watersheds [5, 29, 25], and binary partition trees [2, 33, 3, 21]. Providing a survey is beyond the scope of this paper and the reader can refer to [7] for a review on hierarchies of segmentations.

In the construction of a hierarchy of segmentations, two aspects should be considered: (1) the regions that compose the finest partition of the hierarchy; and (2) the criterion under which those regions are merged along the hierarchy. In this article, we focus on the second aspect of the construction of hierarchies. More precisely, we explore the method to combine hierarchies introduced in [14, 12] and assessed in [19]. Other approaches on combinations of hierarchies of segmentations have already been tackled in [18, 21, 30].

In this article, we consider hierarchies of partitions that are based on the watershed transform [4, 6, 29, 9], which derives from the topographic definition of watersheds lines and catchment basins. Combined with other morphological operators and machine learning techniques, watershed segmentation is an important step for solving practical problems in many application domains such as medicine and biology, computer vision, remote sensing and material science (see *e.g.* [22, 23, 16, 17, 1, 15, 21]).

More precisely, we focus on the watershed based hierarchies of partitions, called hierarchical watersheds, that are built in the framework of edge-weighted graphs and that are optimal in the sense of minimum spanning forests [9, 11, 13, 28]. The link between hierarchical watersheds and the well-known minimum spanning forest optimization problem allows us to consider the results of the studies on the latter in the context of hierarchical watersheds. A first important consequence of this link is the use of minimum spanning tree algorithms to design efficient algorithms to compute hierarchical watersheds [11, 13, 28]. Moreover, the properties of minimum spanning trees induce corresponding properties on hierarchical

This research is partly funded by the Bézout Labex, funded by ANR, reference ANR-10-LABX-58

watersheds. Furthermore, minimum spanning trees and watersheds are linked to a broader range of optimization problems. For instance, in [8], the authors unify graph-cuts, shortest path forests, watersheds and random walkers in a single framework that solves each of those problems when different parameters are given.

Most hierarchies of partitions used in the context of image analysis (*e.g.* [33]) are defined by means of an algorithm rather than by the optimization of a cost function. In turn, a hierarchical watershed optimizes a well-defined cost function for every partition. It is noteworthy that many cost functions used to define image partitions are not adapted to the computation of hierarchies. For instance, the partitions induced by the min-cuts, average-cuts and shortest path forests of a graph for a sequence of decreasing subsets of markers are generally not nested (see examples in Appendix A).

In practice, when a single hierarchical segmentation algorithm is applied to an image, the resulting hierarchy often does not fit the application and cognitive expectation for all regions and areas of the image. In [19], the authors showed that, based on the assessment method proposed in [32], combinations of hierarchical watersheds are superior when compared to the individual hierarchical watersheds. This is illustrated in Figure 1, where we present a combination of two hierarchical watersheds obtained through the method described in [12]. Those hierarchies are driven by regional attribute values based on area [26] and on dynamics [29]. In Figure 1, we also present segmentations with 75 regions extracted from each hierarchy. Hierarchies are represented thanks to their saliency maps, also know as ultrametric watersheds and ultrametric contour maps [29, 2, 18, 11, 12]. In this representation of saliency maps, the darkest contours are the ones that persist at the highest levels of the hierarchies. From the saliency map of the hierarchical watershed based on area shown in Figure 1, we can see that the sky and sea regions are oversegmented at high levels of this hierarchy. On the other hand, spurious regions are preserved at high levels of the hierarchy based on dynamics and the two people, which are significant regions from a cognitive point of view, are merged with the background. We can see that, in the combination of those two hierarchies, the two people are preserved at high levels of this hierarchy and the sky and sea regions are less oversegmented.

This article is a theoretical extension of the experimental study on combinations of hierarchical watersheds presented in [19]. In Figure 2(a), we show a scheme of the method to combine hierarchical watersheds introduced in [14, 12]. In the table of Figure 2(b), we present three of the combinations assessed in [19] and how much the scores of the combinations are improved when compared to the individual hierarchical watershed of highest score in the combination. Knowing that combinations of hierarchical watersheds can achieve superior quantitative results when compared to individual hierarchical watersheds, we present a theoretical study of the combinations of hierarchical watersheds assessed in [19]. Our contributions are threefold: (1) an investigation of which of the combinations assessed in [19] are hierarchical watersheds, and hence which of the combinations assessed in [19] satisfy the optimization property of hierarchical watersheds discussed earlier in the introduction; (2) the introduction of the notion of flattened hierarchical watersheds, which generalizes the notion of hierarchical watersheds to consider hierarchies computed from a partial ordering on a set of markers; and (3) a sufficient condition for a combination of hierarchical watersheds to always result in a flattened hierarchical watershed.

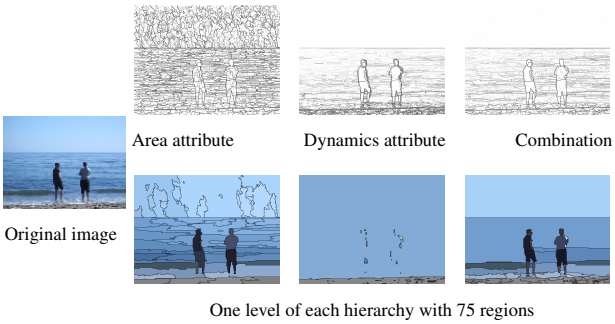
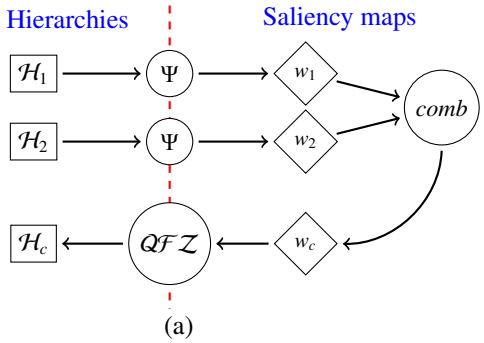


Figure 1: Hierarchical watersheds based on area and dynamics and their combination by average.

In section 2, we present the basic notions for handling hierarchies with graphs, and we review the method to combine hierarchies of segmentations assessed in [19]. In section 3, we present our contributions on the properties of combinations of hierarchical watersheds. We first review hierarchical watersheds and present a generalization of the characterization of hierarchical watersheds introduced in [20]. Then, based on this generalized characteri-



(a)

comb \ $\mathcal{H}_1$ and $\mathcal{H}_2$	**area & dynamics**	**nb. of descendants & dynamics**	**nb. of minima & dynamics**
Supremum	+3.7%	+3.8%	+2.8%
Infimum	+9.3%	+8.8%	+4.9%
Linear combination	+10.7%	+10.7%	+6.2%
Concatenation	+2.3%	+2.3%	+0.9%

(b)

Figure 2: (a) Scheme of the method to combine hierarchical watersheds proposed in [14, 12]. First, given two hierarchical watersheds $\mathcal{H}_1$ and $\mathcal{H}_2$, the saliency maps w_1 and w_2 of respectively $\mathcal{H}_1$ and $\mathcal{H}_2$ are computed. Then, the saliency maps w_1 and w_2 are combined, resulting in the map w_c. Finally, the resulting hierarchy is the quasi-flat zones hierarchy of w_c. (b) Twelve of the combinations assessed in [19] and how much (in percentage) the score of the combination is higher than the individual hierarchical watershed of highest score. The performance of each hierarchy is measured by the average of the area under the fragmenentation curves for optimal and horizontal cuts as proposed in [32].

zation, we are able to state properties related to combinations of hierarchical watersheds.

2. Combinations of hierarchies

In this section, we first remind the definitions of connected hierarchies of partitions, weighted graphs, hierarchies of quasi-flat zones and saliency maps. Then, we review the technique to combine hierarchical watersheds assessed in [19].

2.1. Connected hierarchies of partitions

Let V be a set. A *partition* (of V) is a set $\mathbf{P}$ of non empty disjoint subsets of V whose union is V. Any element of a partition $\mathbf{P}$ is called a *region of* $\mathbf{P}$. Let $\mathbf{P}_1$ and $\mathbf{P}_2$ be two partitions. We say that $\mathbf{P}_1$ is a *refinement* of $\mathbf{P}_2$ if every element of $\mathbf{P}_1$ is included in an element of $\mathbf{P}_2$. A *hierarchy (of partitions)* is a sequence $\mathcal{H} = (\mathbf{P}_0, \ldots, \mathbf{P}_\ell)$ of partitions such that $\mathbf{P}_{i-1}$ is a refinement of $\mathbf{P}_i$, for any i in $\{1, \ldots, \ell\}$ and such that $\mathbf{P}_n = \{V\}$. Let $\mathcal{H} = (\mathbf{P}_0, \ldots, \mathbf{P}_\ell)$ be a hierarchy of partitions. Any region of a partition $\mathbf{P}$ of $\mathcal{H}$ is called a region of $\mathcal{H}$.

A hierarchy of partitions can be represented as a tree whose nodes correspond to regions, as shown on Figure 3. Given a hierarchy $\mathcal{H}$ and two regions X and Y of $\mathcal{H}$, we say that X *is a parent of* Y *(or that* Y *is a child of* X*)* if $Y \subset X$ and X is minimal for this property, *i.e.*, if there is a region Z such that $Y \subseteq Z \subset X$, then we have $Y = Z$. It

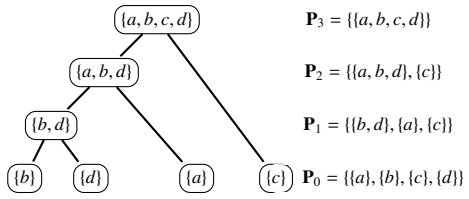


Figure 3: Representation of the hierarchy $\mathcal{H} = (\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3)$ on the set $V = \{a, b, c, d\}$. The partition $\mathbf{P}_0$ of $\mathcal{H}$ contain all the leaf regions of $\mathcal{H}$.

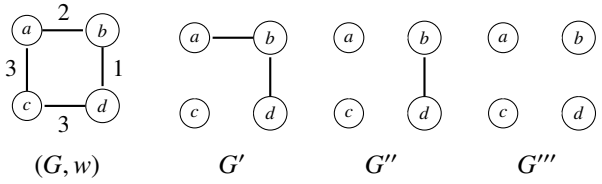


Figure 4: From left to right: a weighted graph (G, w) and three subgraphs of G. The partitions $\boldsymbol{P_0}$, $\boldsymbol{P_1}$, $\boldsymbol{P_2}$ and $\boldsymbol{P_3}$ of Figure 3 are the set of connected components of the graphs G''', G'', G' and G, respectively.

can be seen that any region X of $\mathcal{H}$ such that $X \neq V$ has exactly one parent. For any region R of $\mathcal{H}$, if R is not the parent of any region of $\mathcal{H}$, we say that R *is a leaf region (of* $\mathcal{H}$*)*. Otherwise, we say that R *is a non-leaf region*.

In Figure 3, the regions of a hierarchy $\mathcal{H}$ are linked to their parents (and to their children) by straight lines.

A (undirected) *graph* is a pair $G = (V, E)$, where V is a finite set and E is a set of pairs of distinct elements of V,

i.e., $E \subseteq \{\{x, y\} \subseteq V \mid x \neq y\}$. Each element of V is called a *vertex (of G)*, and each element of E is called an *edge (of G)*. To simplify the notations, the set of vertices and edges of any graph G will be also denoted by $V(G)$ and $E(G)$, respectively.

Let $G = (V, E)$ be a graph and let X be a subset of V. A sequence $\pi = (x_0, \ldots, x_\ell)$ of elements of X is a *path (in X) from x_0 to x_ℓ* if $\{x_{i-1}, x_i\}$ is an edge of G for any i in $\{1, \ldots, \ell\}$. The subset X of V is said to be *connected for G* if, for any x and y in X, there exists a path from x to y. The subset X is a *connected component* of G if X is connected and if X is maximal for this property, *i.e.*, for any connected subset Y of V such that $X \subseteq Y$, we have $X = Y$. In the following, we denote by $CC(G)$ the set of all connected components of G. It is well known that this set $CC(G)$ of all connected components of G is a partition of V.

Let $G = (V, E)$ be a graph. *A partition of V is connected (for G)* if each of its regions is connected for G and *a hierarchy on V is connected (for G)* if each of its partitions is connected.

For example, the hierarchy $\mathcal{H}$ of Figure 3 is connected for the graph G of Figure 4. The partitions $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$ and $\mathbf{P}_3$ of $\mathcal{H}$ are the sets of connected components of G''', G'', G' and G of Figure 4, respectively.

2.2. Saliency maps and quasi-flat zones

Let G be a graph. If w is a map from the edge set of G to the set $\mathbb{R}$ of real numbers, then the pair (G, w) is called an *(edge) weighted graph*. If (G, w) is a weighted graph, for any edge u of G, the value $w(u)$ is called the *weight of u (for w)*.

Important notation: in the sequel of this paper, the symbol (G, w) denotes a weighted graph whose vertex set is connected. To shorten the notation, the vertex set of G is denoted by V and its edge set is denoted by E. Without loss of generality, we also assume that the range of w is the set $\mathbb{E}$ of all integers from 0 to $|E|-1$ (otherwise, one could always consider an increasing one-to-one correspondence from the set $\{w(u) \mid u \in E\}$ into the subset $\{0, ..., |\{w(u) \mid u \in E\}| - 1\}$ of $\mathbb{E}$).

As established in [12], a connected hierarchy can be equivalently treated by means of a weighted graph through the notion of a saliency map. Given a hierarchy $\mathcal{H} = (\mathbf{P}_0, \ldots, \mathbf{P}_\ell)$ connected for G, the *saliency map of $\mathcal{H}$* is the map from E into $\{0, \ldots, \ell\}$, denoted by $\Phi(\mathcal{H})$, such that, for any edge $u = \{x, y\}$ in E, the value $\Phi(\mathcal{H})(u)$ is the lowest value i in $\{0, \ldots, \ell\}$ such that x and y belong to a same region of $\mathbf{P}_i$. It follows that any connected hierarchy has a unique saliency map.

For instance, the weight map w depicted in Figure 4 is the saliency map of the hierarchy $\mathcal{H}$ of Figure 3.

Given a hierarchy $\mathcal{H}$ and its saliency map $\Phi(\mathcal{H})$, we can recover $\mathcal{H}$ from $\Phi(\mathcal{H})$ through the notion of quasi-flat zones. Let λ be any element in $\mathbb{R}$. The *λ-level set of (G, w)*, denoted by G_λ, is the graph $(V, E_\lambda(G))$ such that $E_\lambda(G) = \{u \in E(G) \mid w(u) \leq \lambda\}$. The sequence

$$\mathcal{QFZ}(w) = (CC(G_\lambda) \mid \lambda \in \mathbb{E}) \tag{1}$$

is a hierarchy called the *quasi-flat zones (QFZ) hierarchy (of w)* [27, 24, 34, 13].

In Figure 4 we present a weighted graph (G, w) whose QFZ hierarchy $\mathcal{QFZ}(w)$ is precisely the sequence $(CC(G'''), CC(G''), CC(G'), CC(G))$ of connected components of the subgraphs of G shown in Figure 4. Therefore, the hierarchy of Figure 3 is also the quasi-flat zones hierarchy of w.

Any hierarchy $\mathcal{H}$ connected for G is the QFZ hierarchy of its own saliency map, *i.e.*, $\mathcal{H} = \mathcal{QFZ}(\Phi(\mathcal{H}))$.

For instance, the weight map w depicted in Figure 4 is the saliency map of the hierarchy $\mathcal{H}$ of Figure 3 and, conversely, $\mathcal{H}$ is the QFZ hierarchy of w.

2.3. Hierarchical watersheds

In [9], the authors formalize watersheds in the framework of weighted graphs and show the optimality of watersheds in the sense of minimum spanning forests. In this section, we present hierarchical watersheds following the definition of hierarchies of minimum spanning forests presented in [11, 13].

We say that the graph $G = (V, E)$ *is a forest* if, for any edge u in E, the number of connected components of the graph $(V, E \setminus \{u\})$ is greater than the number of connected components of G. Given another graph G', we say that *G' is a subgraph of G*, denoted by $G' \sqsubseteq G$, if $V(G') \subseteq V$ and $E(G') \subseteq E$. Let G'' be a subgraph of G and let G' be a subgraph of G''. The graph G'' is a *Minimum Spanning Forest (MSF) of G rooted in G'* if:

1. the graphs G and G'' have the same set of vertices, *i.e.*, $V(G'') = V$; and

2. each connected component of G'' includes exactly one connected component of G'; and

3. the sum of the weight of the edges of G'' is minimal among all subgraphs of G for which the above conditions 1 and 2 hold true.

A MSF of (G, w) rooted in a single vertex of G is a tree (connected forest) called a *Minimum Spanning Tree (MST) of* (G, w).

Let k be a value in $\mathbb{R}$. A connected subgraph G' of G is a *(regional) minimum* (of w) at level k if:

1. $E(G') \neq \emptyset$; and

2. for any edge u in $E(G')$, the weight of u is equal to k; and

3. for any edge $\{x, y\}$ in $E \setminus E(G')$ such that $|\{x, y\} \cap V(G')| \geq 1$, the weight of $\{x, y\}$ is strictly greater than k.

Important notation: in the sequel of this article, we denote by n the number of minima of w. Any sequence of minima of w considered in this article is a sequence of n pairwise distinct minima of w.

Let $\{G_1, \ldots, G_\ell\}$ be a set of graphs. We denote by $\sqcup\{G_1, \ldots, G_\ell\}$ the graph $(\cup\{V(G_j) \mid j \in \{1, \ldots, \ell\}\}, \cup\{E(G_j) \mid j \in \{1, \ldots, \ell\}\})$. In the following, we define hierarchical watersheds based on minimum spanning forests following the definition of [11, 13].

Definition 1 (hierarchical watershed [11, 13]). *Let $\mathcal{S} = (M_1, \ldots, M_n)$ be a sequence of minima of w. Let $(G_0, \ldots, G_{n-1})$ be a sequence of subgraphs of G such that:*

1. *for any $i \in \{0, \ldots, n-1\}$, the graph G_i is a MSF of G rooted in $\sqcup\{M_j \mid j \in \{i+1, \ldots, n\}\}$; and*

2. *for any $i \in \{1, \ldots, n-1\}$, $G_{i-1} \sqsubseteq G_i$.*

The sequence $\mathcal{T} = (CC(G_0), \ldots, CC(G_{n-1}))$ is called a hierarchical watershed of (G, w) for $\mathcal{S}$. *Given a hierarchy $\mathcal{H}$, we say that $\mathcal{H}$* is a hierarchical watershed of (G, w) *if there exists a sequence $\mathcal{S} = (M_1, \ldots, M_n)$ of minima of w such that $\mathcal{H}$ is a hierarchical watershed for $\mathcal{S}$.*

A weighted graph (G, w) and two hierarchical watersheds $\mathcal{H}_1$ and $\mathcal{H}_2$ of (G, w) are illustrated in Figure 5. It can be verified that $\mathcal{H}_1$ and $\mathcal{H}_2$ are the hierarchical watersheds of (G, w) for the sequences $\mathcal{S}_1 = (A, B, D, C, E)$ and $\mathcal{S}_2 = (D, C, E, A, B)$, respectively.

2.4. Combination of hierarchies

Combining partitions and, *a fortiori*, hierarchies is not straightforward. This problem has been tackled in [18, 21, 14, 12] thanks to the use of saliency maps and we follow the same approach. More precisely, in order to combine two hierarchies $\mathcal{H}_1$ and $\mathcal{H}_2$, we proceed in three steps: first the saliency maps of $\mathcal{H}_1$ and $\mathcal{H}_2$ are considered, then the two saliency maps are combined to obtain new weights on the edges of G, and, finally, the combination of hierarchies is the QFZ hierarchy of the new weight map.

Let $\mathcal{F}$ be the set of all maps from E into $\mathbb{R}$. Any map C from $\mathcal{F}^2$ into $\mathcal{F}$ is called a *combining function*.

Given two hierarchies $\mathcal{H}_1$ and $\mathcal{H}_2$ and a combining function C, the *combination of $\mathcal{H}_1$ and $\mathcal{H}_2$ by C* is the hierarchy $\mathcal{H}_C(\mathcal{H}_1, \mathcal{H}_2)$ defined by:

$$\mathcal{H}_C(\mathcal{H}_1, \mathcal{H}_2) = \mathcal{QFZ}(C(\Phi(\mathcal{H}_1), \Phi(\mathcal{H}_2))). \tag{2}$$

In [19], the authors consider three classical functions in the instantiation of the combining function (supremum, infimum and linear combination), and a new function called *concatenation*. Given two maps f and g in $\mathcal{F}$, the supremum, infimum and linear combination of f and g, respectively denoted by $\curlyvee(f, g)$, $\curlywedge(f, g)$ and $\boxplus_{\alpha,\beta}(f, g)$, are defined for each edge u in E as:

$$\begin{aligned} \curlyvee(f, g)(u) &= max(f(u), g(u)) \\ \curlywedge(f, g)(u) &= min(f(u), g(u)) \\ \boxplus_{\alpha,\beta}(f, g)(u) &= \alpha f(u) + \beta g(u) \end{aligned} \tag{3}$$

where $\alpha, \beta \in \mathbb{N}$ are parameters of the linear combination.

One example of a combination of hierarchies by infimum is shown in Figure 6.

The purpose of the concatenation is to combine higher levels of a hierarchy with lower levels of another hierarchy. This type of combination is useful when a hierarchy $\mathcal{H}_1$ succeeds at describing the small details of an image at lower levels, but fails at filtering the small regions to capture the main large objects at higher levels of the hierarchy. Therefore, it can be interesting to concatenate $\mathcal{H}_1$

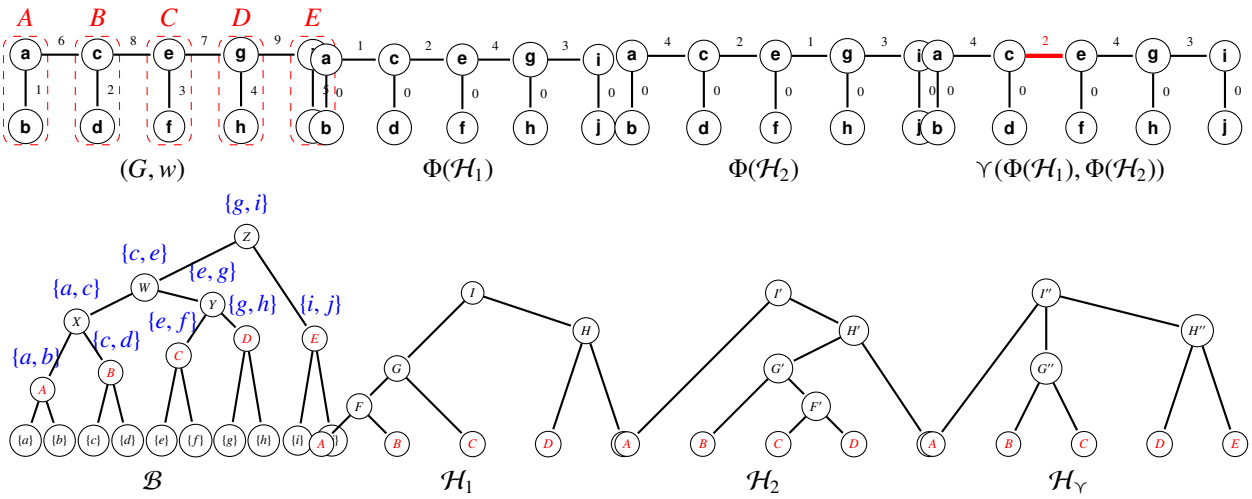


Figure 5: First line: a weighted graph (G, w) with five minima delimited by the dashed rectangles, the saliency maps $\Phi(\mathcal{H}_1)$ and $\Phi(\mathcal{H}_2)$ of two hierarchical watersheds of (G, w), and the combination $\Upsilon(\Phi(\mathcal{H}_1), \Phi(\mathcal{H}_2))$ of $\Phi(\mathcal{H}_1)$ and $\Phi(\mathcal{H}_2)$ with supremum. Second line: the unique binary partition hierarchy $\mathcal{B}$ of (G, w), the hierarchical watershed $\mathcal{H}_1$ of (G, w) for the sequence $\mathcal{S}_1 = (A, B, D, C, E)$, the hierarchical watershed $\mathcal{H}_2$ of (G, w) for the sequence $\mathcal{S}_2 = (D, C, E, A, B)$, and the resulting combinations of $\mathcal{H}_1$ and $\mathcal{H}_2$ with supremum: $\mathcal{H}_\Upsilon = \mathcal{QFZ}(G, \Upsilon(\Phi(\mathcal{H}_1), \Phi(\mathcal{H}_2)))$.

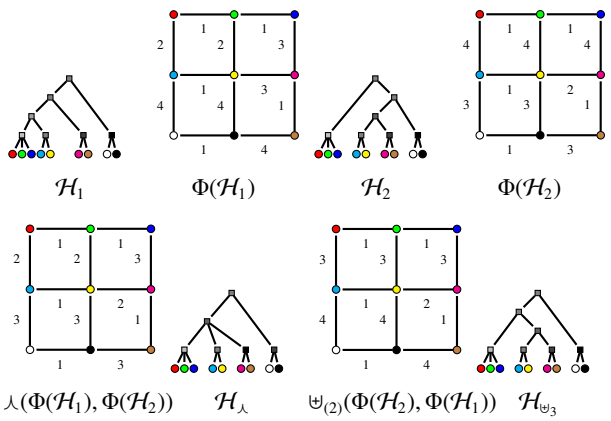


Figure 6: Two hierarchies $\mathcal{H}_1$ and $\mathcal{H}_2$, their saliency maps $\Phi(\mathcal{H}_1)$ and $\Phi(\mathcal{H}_2)$, and their combinations with infimum ($\curlywedge$) and concatenation ($\uplus_3$).

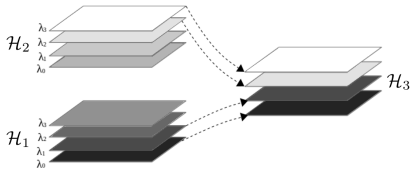


Figure 7: Concatenation of the hierarchies $\mathcal{H}_1$ and $\mathcal{H}_2$ at level $\lambda = 2$.

with another hierarchy $\mathcal{H}_2$ whose high level sets describe well the important regions in the image. This general idea is represented in Figure 7. Given two hierarchies $\mathcal{H}_1$ and $\mathcal{H}_2$, we aim to obtain a new hierarchy $\mathcal{H}_3$ whose high (resp. low) levels correspond approximately to the high (resp. low) levels of $\mathcal{H}_1$ (resp. $\mathcal{H}_2$). In order to define the concatenation of hierarchies, we first define the double threshold function. Given any map f in $\mathcal{F}$ and given two parameters α and β in $\mathbb{R}$ such that $\alpha < \beta$, we denote by $T(f, \alpha, \beta)$ the *double threshold of f by (α, β)* such that, for any edge u of G:

$$T(f, \alpha, \beta)(u) = \begin{cases} 0 & if\ f(u) < \alpha \\ \beta & if\ f(u) > \beta \\ f(u) & otherwise \end{cases} \tag{4}$$

Let f and g be two maps in $\mathcal{F}$. Given a threshold value λ, the *concatenation of f and g at level λ*, for any edge u, is given by:

$$\uplus_\lambda (f, g)(u) = max(T(f, 0, \lambda)(u), T(g, \lambda, \infty)(u)) \quad (5)$$

In Figure 6, we show the concatenation of the hierarchies $\mathcal{H}_2$ and $\mathcal{H}_1$ at level $\lambda = 3$.

3. Properties of combinations of hierarchical watersheds

Combining hierarchical watersheds through their saliency maps raises the question whether the combination of hierarchies is closed for the set of hierarchical watersheds. More precisely, given any two hierarchical watersheds $\mathcal{H}_1$ and $\mathcal{H}_2$ of (G, w), is the combination of $\mathcal{H}_1$ and $\mathcal{H}_2$ with a given combining function also a hierarchical watershed of (G, w)?

In this section, we answer to this question for combinations of hierarchical watersheds with supremum, infimum, concatenation and linear combination with parameters α and β different from zero. For the linear combination, we consider the case where $\alpha = 1$ and $\beta = 1$, which will be denoted by *average*. We show that combinations with any of those functions do not result in hierarchical watersheds in general. However, in the particular case where we consider a unique ordering on the edges of (G, w), combinations of hierarchical watersheds with infimum have the noteworthy property of being flattened (simplified) hierarchical watersheds, *i.e.*, hierarchies resulting from removing (and/or repeating) partitions of a departing hierarchical watershed. For this particular case, we present a sufficient condition for a combining function to always output flattened hierarchical watersheds.

Definition 2 (flattening of hierarchies). *Let $\mathcal{H}$ and $\mathcal{H}'$ be two hierarchies on V such that any partition of $\mathcal{H}$ is a partition of $\mathcal{H}'$. We say that $\mathcal{H}$* is a flattening of $\mathcal{H}'$.

Let $\mathcal{H}$ and $\mathcal{H}'$ be two hierarchies on V such that $\mathcal{H}$ is a flattening of $\mathcal{H}'$. If $\mathcal{H}'$ is a hierarchical watershed of (G, w), then we say that $\mathcal{H}$ *is a flattened hierarchical watershed of* (G, w). The reader may note that there can be repeated partitions in both $\mathcal{H}$ and $\mathcal{H}'$. Hence, the hierarchy $\mathcal{H}$ can have more partitions than $\mathcal{H}'$, but $\mathcal{H}$ has less distinct partitions than $\mathcal{H}'$.

To compute a hierarchical watershed of (G, w), a sequence of minima of w is often defined by extinction values [36]. When distinct minima of w have the same extinction value, the order between those minima is defined arbitrarily. Let G' be the MSF of (G, w) rooted in the minima of w. By Definition 1, we may say that a hierarchical watershed of (G, w) can be obtained by filtering, one by one, the connected components of G'. Now, let us consider a framework in which the minima with equal extinction values are treated simultaneously. In this new framework, the connected components of G' rooted in minima of w with equal extinction values are filtered out simultaneously. We can affirm that the resulting partitions of this framework are a subset of the partitions of a hierarchical watershed of (G, w), and hence a flattened hierarchical watershed. Thus, we can see that the notion of flattened hierarchical watersheds, even though not formally defined previously, arise naturally in the context of marker-based watershed segmentation.

It is noteworthy that, alike a hierarchical watershed, all partitions of a flattened hierarchical watershed are optimal in the sense of minima spanning forests.

To state the properties of combinations of hierarchical watersheds aforementioned, we review the characterization of hierarchical watersheds proposed in [20]. To this end, we first remind the definition of binary partition hierarchies by altitude ordering, which is closely linked to hierarchical watersheds [13].

3.1. Binary partition hierarchies by altitude ordering

Binary partition trees [33] are widely used for hierarchical image representation. In this section, we describe the particular case where regions linked by the lowest edge weights are the first regions to be merged in the hierarchy [13], which is deeply linked to single-linkage clustering [12].

Let $\prec$ be an altitude ordering (on E) for w, *i.e.*, $\prec$ is a binary relation that is transitive and trichotomous (for any u and v in E only one of the relations $u \prec v$, $v \prec u$ and $v = u$ holds true) such that, for any u and v in E, if $w(u) < w(v)$, then $u \prec v$. Let k be any element in $\{1, \ldots, |E|\}$. We denote by $u_k^\prec$ the k-th element of E with respect to $\prec$. We set $\mathbf{B}_0 = \{\{x\} | x \in V\}$. The *$k$-partition of V (by the ordering $\prec$)* is defined by $\mathbf{B}_k = \{\mathbf{B}_{k-1}^y \cup \mathbf{B}_{k-1}^x\} \cup (\mathbf{B}_{k-1} \setminus \{\mathbf{B}_{k-1}^x, \mathbf{B}_{k-1}^y\})$ where $u_k^\prec = \{x, y\}$ and $\mathbf{B}_{k-1}^x$ and $\mathbf{B}_{k-1}^y$ are the regions of $\mathbf{B}_{k-1}$ that contain x and y, respectively. The sequence

$(\mathbf{B}_i \mid i = 0 \text{ or } \mathbf{B}_i \neq \mathbf{B}_{i-1})$ is a hierarchy on V. This hierarchy $(\mathbf{B}_i \mid i = 0 \text{ or } \mathbf{B}_i \neq \mathbf{B}_{i-1})$, denoted by $\mathcal{B}_\prec$, is called the *binary partition hierarchy (by altitude ordering) of* (G, w) *by* $\prec$.

Let us consider the graph (G, w) of Figure 4. Let $\prec$ be an altitude ordering for w such that $\{b, d\} \prec \{a, b\} \prec \{a, c\} \prec \{c, d\}$. We have $\mathbf{B}_0 = \{\{a\}, \{b\}, \{c\}, \{d\}\}$. For k from 1 to 4, the k-partition of $V(G)$ by $\prec$ are the following: $\mathbf{B}_1 = \{\{b, d\}, \{a\}, \{c\}\}$, $\mathbf{B}_2 = \{\{a, b, d\}, \{c\}\}$, $\mathbf{B}_3 = \{\{a, b, d, c\}\}$ and $\mathbf{B}_4 = \{\{a, b, d, c\}\}$. Since the partitions $\mathbf{B}_3$ and $\mathbf{B}_4$ are equal, the binary partition hierarchy of (G, w) by $\prec$ is the sequence $(\mathbf{B}_0, \mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3)$.

Let $\mathcal{B}$ be a hierarchy on V. We say that $\mathcal{B}$ *is a binary partition hierarchy of* (G, w) if there is an altitude ordering $\prec$ for w such that $\mathcal{B}$ is the binary partition hierarchy of (G, w) by $\prec$.

Let $\prec$ be an altitude ordering for w. We can associate any non-leaf region X of the binary partition hierarchy $\mathcal{B}_\prec$ of (G, w) by $\prec$ to the lowest rank r such that $\mathbf{B}_r$ contains X. This rank is called the *rank of* X and it is denoted by $r(X)$. Let X be a non-leaf region of $\mathcal{B}_\prec$, the *building edge of* X is the $r(X)^{th}$ edge for $\prec$. Given an edge u in E, we denote the region of $\mathcal{B}_\prec$ whose building edge is u by R_u. If u is the building edge of a region of $\mathcal{B}_\prec$, we say that u *is a building edge of* $\mathcal{B}_\prec$. The set of all building edges of $\mathcal{B}_\prec$ is denoted by $E(\mathcal{B}_\prec)$.

In Figure 5, we present a weighted graph (G, w) and a binary partition hierarchy $\mathcal{B}$ of (G, w). The building edge of each non-leaf region R of $\mathcal{B}$ is shown above the node that represents R.

Let $\mathcal{B}$ be a binary partition hierarchy of (G, w). Let X and Y be two distinct regions of $\mathcal{B}$. If the parent of X is equal to the parent of Y, we say that X is a sibling of Y, that Y is a sibling of X and that X and Y are siblings. It can be seen that X has exactly one sibling and we denote this unique sibling of X by $sibling(X)$.

Important remark: by abuse of notation, when no confusion is possible, we will denote the set of vertices of any minima of w as a minima of w.

As established in [28], given a binary partition hierarchy $\mathcal{B}$ of (G, w), the minima of w can be extracted from $\mathcal{B}$ as well as the watershed-cut edges for $\mathcal{B}$, whose definition is given bellow.

Definition 3 (watershed-cut edge). *Let $\mathcal{B}$ be a binary partition hierarchy of (G, w), let u be a building edge of $\mathcal{B}$ such that each child of R_u contains at least one minimum of w. Then we say that u* is a watershed-cut edge (of (G, w)) for $\mathcal{B}$.

3.2. Characterization of hierarchical watersheds

In [20], the authors present a characterization of hierarchical watersheds for weighted graphs which are trees with pairwise distinct edge weights. They define one-side increasing maps and establish their link with hierarchical watersheds. In this section, we generalize their definition of one-side increasing maps (in Definition 4) as well as their characterization of hierarchical watersheds (in Theorem 5) to any weighted graph [1].

Definition 4 (one-side increasing map). *Let $\mathcal{B}$ be a binary partition hierarchy of (G, w). Let f be a map from E into $\mathbb{R}$. We say that f is* one-side increasing for $\mathcal{B}$ *if:*

1. $\{f(e) \mid e \in E(\mathcal{B})\} = \{0, \ldots, n-1\}$;
2. *for any edge u in $E(\mathcal{B})$, $f(u) > 0$ if and only if u is a watershed-cut edge for $\mathcal{B}$; and*
3. *for any edge u in $E(\mathcal{B})$, there exists a child R of R_u such that $f(u) \geq \vee\{f(v)$ such that R_v is included in $R\}$, where $\vee\{\} = 0$.*

Important remark: for simplicity's sake, only the building edges of $\mathcal{B}$ are considered in the second statement of Definition 4. But, since we are handling saliency maps of hierarchies, this statement applies to every edge in E.

Theorem 5. *Let $\mathcal{H}$ be a hierarchy on V. The hierarchy $\mathcal{H}$ is a hierarchical watershed of (G, w) if and only if there is an altitude ordering $\prec$ for w such that the saliency map $\Phi(\mathcal{H})$ of $\mathcal{H}$ is one-side increasing for $\mathcal{B}_\prec$.*

For instance, we can verify in Figure 5 that the saliency maps of the hierarchical watersheds $\mathcal{H}_1$ and $\mathcal{H}_2$ of (G, w) are one-side increasing for the binary partition hierarchy $\mathcal{B}$.

[1] The proofs of the theorem and properties can be found in Appendix B.

3.3. Combinations with infimum, supremum, average and concatenation

As established in the following property, combinations of hierarchical watersheds with infimum, supremum, average and concatenation are not hierarchical watersheds nor flattened hierarchical watersheds in general.

Property 6. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hierarchical watersheds of (G, w). The combination of $\mathcal{H}_1$ and $\mathcal{H}_2$ with supremum (resp. average, concatenation and infimum) is not a flattened hierarchical watershed of (G, w) in general.*

Proof 1 (of Property 6). *Let us consider the hierarchical watersheds $\mathcal{H}_1$ and $\mathcal{H}_2$ of the graph (G, w) and their combination $\mathcal{H}_\curlyvee$ with supremum depicted in Figure 5. The first pair of regions to be merged in $\mathcal{H}_\curlyvee$ are the minima B and C. However, there is no hierarchical watershed of (G, w) such that B and C are the first minima to be merged. By contradiction, let us assume that there is a sequence of minima $\mathcal{S}'$ of (G, w) such that, in the hierarchical watershed $\mathcal{H}'$ of (G, w) for $\mathcal{S}'$, B and C are the first minima to be merged. Then, we can infer that either B or C is the first minimum in the sequence $\mathcal{S}'$. However, if B were the first minimim, then it would be merged to the minimum A because the weight of the edge linking A and B is 6, which is lower than the weight of the edge linking B and C. On the other hand, if C were the first minimum in the sequence $\mathcal{S}'$, then C would be first merged to D due to the same reason. Hence, there is no hierarchical watershed $\mathcal{H}'$ of (G, w) such that partition $\{A, B \cup C, D, E\}$ of $\mathcal{H}_\curlyvee$ is a partition of $\mathcal{H}'$. Therefore, $\mathcal{H}_\curlyvee$ is not a flattened hierarchical watershed of (G, w). The same situation is found in the combinations with average and concatenation of the maps $\Phi(\mathcal{H}_1)$ and $\Phi(\mathcal{H}_2)$ shown in Figure 8. In both cases, the minima B and C are the first to be merged in the hierarchy. The counter example for infimum is presented in Figure 9, where we show two maps f_1 and f_2 which are one-side increasing for distinct binary partition hierarchies of the graph (G', w'). The reader can verify that the QFZ hierarchy of the combination of f_1 and f_2 with infimum is not a flattened hierarchical watershed of (G', w').* □

The counter-example of Figure 9 considers two maps which are one-side increasing for distinct binary partition hierarchies of (G', w'). However, as established in the following property, if the input maps are one-side increasing for the same binary partition hierarchy of (G, w), then their combination with infimum is a flattened hierarchical watershed of (G, w).

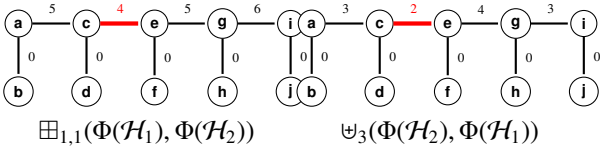


Figure 8: Combination of the saliency maps $\Phi(\mathcal{H}_1)$ and $\Phi(\mathcal{H}_2)$ of Figure 5 with average and concatenation ($\lambda = 3$).

Property 7. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hierarchical watersheds of (G, w) and let $\mathcal{B}$ be a binary partition hierarchy of (G, w) such that both $\Phi(\mathcal{H}_1)$ and $\Phi(\mathcal{H}_2)$ are one-side increasing for $\mathcal{B}$. Then the hierarchy $\mathcal{H}_\curlywedge = \mathcal{QFZ}(G, \curlywedge(\Phi(\mathcal{H}_1), \Phi(\mathcal{H}_2)))$ is a flattened hierarchical watershed of (G, w).*

Important remark: in the processing of graph-based image segmentation, it is common to consider a raster scanning of the edges of any input graph. Hence, a unique altitude ordering $\prec$ for w and, consequently, a unique binary partition hierarchy of (G, w) are used by the algorithm.

3.4. Generalization to other combining functions

In this section, we are interested in generalizing the analysis of combination functions. In the following property, we introduce a sufficient condition for a combining function to always output flattened hierarchical watersheds.

Property 8. *Let C be a combining function, let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hierarchical watersheds of (G, w) and let $\mathcal{B}$ be a binary partition hierarchy of (G, w) such that both $\Phi(\mathcal{H}_1)$ and $\Phi(\mathcal{H}_2)$ are one-side increasing for $\mathcal{B}$. The combination of $\mathcal{H}_1$ and $\mathcal{H}_2$ with C is a flattened hierarchical watershed of (G, w) if $C(0, 0) = 0$ and if, for any a, b, c, d in $\{0, \ldots, n-1\}$, we have:*

1. *$C(a, b) = C(b, a)$; and*
2. *if $min(a, b) < min(c, d)$, then $C(a, b) < C(c, d)$; and*

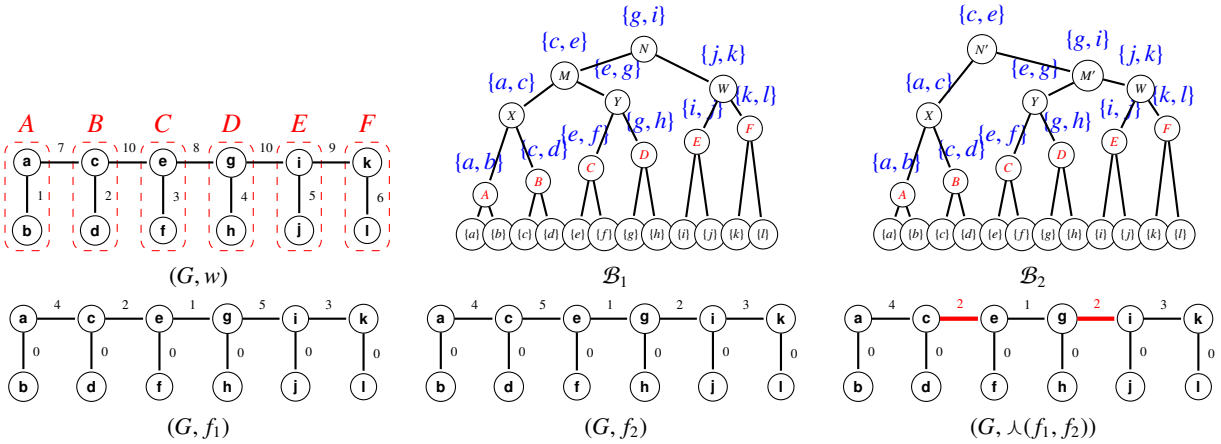


Figure 9: Firs line: a graph (G, w) and its two binary partition hierarchies $\mathcal{B}_1$ and $\mathcal{B}_2$. Second line: a map f_1 which is the saliency map of a hierarchical watershed of (G, w) and which is one-side increasing for $\mathcal{B}_1$, a map f_2 which is the saliency map of a hierarchical watershed of (G, w) and which is one-side increasing for $\mathcal{B}_2$, and the combination of f_1 and f_2 with infimum. The resulting combination is not a flattened hierarchical watershed of (G, w).

3. *if* $min(a, b) = min(c, d)$ *and* $max(a, b) < max(c, d)$, *then* $C(a, b) \leq C(c, d)$.

By Property 8, we can derive other combining functions which always lead to flattened hierarchical watersheds, such as $C'(a, b) = min(a^m, b^m)$ and $C''(a, b) = min(a, b)^m$, for any $m \geq 1$, or the function stated in the following property.

Property 9. *Let* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *be two hierarchical watersheds of* (G, w). *Let* C *be a combining function such that:*

$$C(x, y) = \begin{cases} 0 & if\ x{=}0\ and\ y{=}0 \\ \frac{x^m y^m}{x^m + y^m} & otherwise \end{cases} \quad (6)$$

for $m \geq n$. *The combination of* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *with* C *is a flattened hierarchical watershed of* (G, w).

4. Conclusion

In this article, we presented a study of the combinations of hierarchical watersheds assessed in [19]. Through the characterization of hierarchical watersheds introduced in [20], we concluded that combinations of hierarchical watersheds with infimum, supremum, linear combination and concatenation are not hierarchical watersheds in general. However, if the input hierarchical watersheds have saliency maps that are one-side increasing for the same binary partition hierarchy of (G, w), then their combination with infimum is a flattened hierarchical watershed. We also provided a sufficient condition for a combining function to always output flattened hierarchical watersheds. Among the four combining functions aforementioned, this condition holds true only for infimum.

5. Acknowledgements

We would like to thank Dr. Gia-Thuy Pham for his assistance with the proof of Property 9.

References

[1] Angulo, J., Velasco-Forero, S., Chanussot, J., 2009. Multiscale stochastic watershed for unsupervised hyperspectral image segmentation, in: IEEE International Geoscience and Remote Sensing Symposium, pp. III–93.

[2] Arbelaez, P., Maire, M., Fowlkes, C., Malik, J., 2011. Contour detection and hierarchical image segmentation. IEEE PAMI 33, 898–916.

[3] Arbeláez, P., Pont-Tuset, J., Barron, J.T., Marques, F., Malik, J., 2014. Multiscale combinatorial grouping, in: Proceedings IEEE CVPR, pp. 328–335.

[4] Beucher, S., 1979. Use of watersheds in contour detection, in: Proceedings of the International Workshop on Image Processing, CCETT.

[5] Beucher, S., 1994. Watershed, hierarchical segmentation and waterfall algorithm, in: Mathematical morphology and its applications to image processing. Springer, pp. 69–76.

[6] Beucher, S., Meyer, F., 1993. The morphological approach to segmentation: the watershed transformation. Mathematical morphology in image processing 34, 433–481.

[7] Bosilj, P., Kijak, E., Lefèvre, S., 2018. Partition and inclusion hierarchies of images: A comprehensive survey. Journal of Imaging 4, 33.

[8] Couprie, C., Grady, L., Najman, L., Talbot, H., 2010. Power watershed: A unifying graph-based optimization framework. IEEE transactions on pattern analysis and machine intelligence 33, 1384–1399.

[9] Cousty, J., Bertrand, G., Najman, L., Couprie, M., 2009a. Watershed cuts: Minimum spanning forests and the drop of water principle. IEEE PAMI 31, 1362–1374.

[10] Cousty, J., Bertrand, G., Najman, L., Couprie, M., 2009b. Watershed cuts: Thinnings, shortest path forests, and topological watersheds. IEEE PAMI 32, 925–939.

[11] Cousty, J., Najman, L., 2011. Incremental algorithm for hierarchical minimum spanning forests and saliency of watershed cuts, in: ISMM, Springer. pp. 272–283.

[12] Cousty, J., Najman, L., Kenmochi, Y., Guimarães, S., 2018. Hierarchical segmentations with graphs: Quasi-flat zones, minimum spanning trees, and saliency maps. JMIV 60, 479–502.

[13] Cousty, J., Najman, L., Perret, B., 2013. Constructive links between some morphological hierarchies on edge-weighted graphs, in: ISMM, Springer. pp. 86–97.

[14] Cousty, J., Najman, L., Serra, J., 2008. Raising in watershed lattices, in: 2008 15th IEEE ICIP, IEEE. pp. 2196–2199.

[15] Fehri, A., Velasco-Forero, S., Meyer, F., 2016. Automatic selection of stochastic watershed hierarchies, in: EUSIPCO, IEEE. pp. 1877–1881.

[16] Gauch, J.M., 1999. Image segmentation and analysis via multiscale gradient watershed hierarchies. IEEE TIP 8, 69–79.

[17] Grau, V., Mewes, A., Alcaniz, M., Kikinis, R., Warfield, S.K., 2004. Improved watershed transform for medical image segmentation using prior information. IEEE transactions on medical imaging 23, 447–458.

[18] Kiran, B.R., Serra, J., 2014. Fusion of ground truths and hierarchies of segmentations. PRL 47, 63–71. doi:10.1016/j.patrec.2014.04.019.

[19] Maia, D.S., de Albuquerque Araujo, A., Cousty, J., Najman, L., Perret, B., Talbot, H., 2017. Evaluation of combinations of watershed hierarchies, in: ISMM, Springer. pp. 133–145.

[20] Maia, D.S., Cousty, J., Najman, L., Perret, B., 2019. Recognizing hierarchical watersheds, in: DGCI, Springer. pp. 300–313.

[21] Maninis, K.K., Pont-Tuset, J., Arbeláez, P., Van Gool, L., 2016. Convolutional oriented boundaries, in: ECCV, Springer. pp. 580–596.

[22] Marcotegui, B., 2014. Residual approach on a hierarchical segmentation, in: ICIP, IEEE. pp. 4353–4357.

[23] Marcotegui, B., Zanoguera, F., Correia, P., Rosa, R., Marqués, F., Mech, R., Wollborn, M., 1999. A video object generation tool allowing friendly user interaction, in: ICIP, IEEE. pp. 391–395.

[24] Meyer, F., Maragos, P., 1999. Morphological scale-space representation with levelings, in: International Conference on Scale-Space Theories in Computer Vision, Springer, Berlin.

[25] Meyer, F., Najman, L., 2013. Segmentation, minimum spanning tree and hierarchies. Mathematical morphology: from theory to applications , 229–261.

[26] Meyer, F., Vachier, C., Oliveras, A., Salembier, P., 1997. Morphological tools for segmentation: Connected filters and watersheds, in: Annales des télécommunications, Springer. pp. 367–379.

[27] Nagao, M., Matsuyama, T., Ikeda, Y., 1979. Region extraction and shape analysis in aerial photographs. CGIP 10, 195–223.

[28] Najman, L., Cousty, J., Perret, B., 2013. Playing with kruskal: algorithms for morphological trees in edge-weighted graphs, in: ISMM, Springer. pp. 135–146.

[29] Najman, L., Schmitt, M., 1996. Geodesic saliency of watershed contours and hierarchical segmentation. IEEE PAMI 18, 1163–1173.

[30] Okubadejo, O., Andò, E., Bonnaud, L., Viggiani, G.C., Dalla Mura, M., 2019. Contact based hierarchical segmentation for granular materials, in: ISMM, Saarbrücken, Germany. URL: `https://hal.archives-ouvertes.fr/hal-02097130`.

[31] Pavlidis, T., 1977. Structural pattern recognition. volume 2. Springer.

[32] Perret, B., Cousty, J., Guimaraes, S.J.F., Maia, D.S., 2018. Evaluation of hierarchical watersheds. IEEE TIP 27, 1676–1688.

[33] Salembier, P., Garrido, L., 2000. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. IEEE TIP 9, 561–576.

[34] Soille, P., 2008. Constrained connectivity for hierarchical image partitioning and simplification. IEEE PAMI 30, 1132–1145.

[35] Stoer, M., Wagner, F., 1997. A simple min-cut algorithm. Journal of the ACM (JACM) 44, 585–591.

[36] Vachier, C., Meyer, F., 1995. Extinction value: a new measurement of persistence, in: IEEE Workshop on nonlinear signal and image processing, pp. 254–257.

[37] Wang, S., Siskind, J.M., 2001. Image segmentation with minimum mean cut, in: ICCV, IEEE. pp. 517–524.

Appendix A. Marker-based segmentation on graphs

The hierarchical watersheds [11, 13, 28] are the solutions to a multiscale optimization problem, namely each partition of a hierarchical watershed optimizes a simple cost function. More precisely, each partition of a hierarchical watershed of (G, w) is induced by a solution to the problem of finding a MSF of (G, w) rooted in a certain subset of the minima of w (see Definition 1). Moreover, the (optimal) partitions of a hierarchical watershed satisfy a hierarchical or scale consistency property formalized below in the context of marker-based segmentation. However, as we will see in this section, this property is not satisfied by some of the most common energy terms used in graph based image segmentation.

A *marker set (of* (G, w)*)* is a set of disjoint subsets of V. We denote by Π_V the set of all partitions of V.

Definition 10 (marker-based segmentation). *Let* $\mathcal{M} = \{M_1, \ldots, M_\ell\}$ *be a marker set. A* marker-based segmentation of (G, w) for $\mathcal{M}$ *is a partition* $\mathbf{P}$ *in* Π_V *such that each region of* $\mathbf{P}$ *includes exactly one element of* $\mathcal{M}$.

Marker-based segmentations can be obtained by watershed [9], min-cut [35], average-cut [37] and shortest path forest [10] algorithms, to name a few. It can be observed that the related optimization problems are ill-posed and do not necessarily have a unique solution. Therefore, those algorithms are not deterministic: they can produce several solutions for a given marker set. Hence, in order to study the "hierarchical behavior" of these algorithms, we start by providing a definition of a non-deterministic marker-based segmentation operator.

Definition 11. *A* (non-deterministic) marker-based segmentation operator σ *is a map from the set of all marker sets into the set of all subsets of* Π_V *such that, for any marker set* $\mathcal{M}$*, any partition in* $\sigma(\mathcal{M})$ *is a marker-based segmentation for* $\mathcal{M}$.

Definition 12. *Let* σ *be a marker-based segmentation operator. We say that* σ is hierarchical *if, for any two marker sets* $\mathcal{M}$ *and* $\mathcal{M}'$ *such that* $\mathcal{M}'$ *is a subset of* $\mathcal{M}$*, there exists a pair* $(\mathbf{P}, \mathbf{P}')$ *of partitions such that* $\mathbf{P}$ *and* $\mathbf{P}'$ *belongs to respectively* $\sigma(\mathcal{M})$ *and* $\sigma(\mathcal{M}')$*, and such that* $\mathbf{P}$ *is a refinement of* $\mathbf{P}'$.

Let L be a subset of E. We say that L *is a cut if*, for any edge $u = \{x, y\}$ in L, x and y belong to distinct connected components of $(V, E \setminus L)$. We denote the set of connected components of $(V, E \setminus L)$ by the *partition induced by* L. By abuse of notation, given a spanning forest G' of G, we also denote the set of connected components of G' by the *partition induced by* G'.

The following property asserts that the MSF operator is indeed hierarchical.

Property 13. *Let* σ *be the operator that maps any marker set* $\mathcal{M}$ *into the set of partitions induced by each of the MSFs rooted in* $\mathcal{M}$*. Then, the operator* σ *is hierarchical.*

Property 13 is the basis of hierarchical watersheds (Definition 1). In the following, we show that the operators that produce min-cuts, average-cuts and shortest path forest cuts are not hierarchical.

Definition 14 (min-cuts). *Let* $\mathcal{M}$ *be a marker set. A* min-cut of (G, w) for $\mathcal{M}$ *is a subset* L *of* E *such that:*

1. *the set of connected components of* $(V, E \setminus L)$ *is a marker-based segmentation of* (G, w) *for* $\mathcal{M}$*; and*
2. *the sum* $\sum_{u \in L} w(u)$ *is minimal for all subsets of* E *for which statement 1 holds true.*

Property 15. *Let* σ_{min} *be the operator that maps any marker set* $\mathcal{M}$ *into the set of partitions induced by each of the min-cuts of* (G, w) *for* $\mathcal{M}$*. The operator* σ_{min} *is not hierarchical.*

Proof 2. *Let* (G, w) *be the weighted graph of Figure A.10(a). Let* $\mathcal{M} = \{\{a\}, \{c\}, \{f\}\}$ *be a marker set of* (G, w)*. In Figure A.10(b), we show the unique min-cut* L *(dashed edges) of* (G, w) *for* $\mathcal{M}$*. Hence, we have* $\sigma_{min}(\mathcal{M}) = \{\{\{a, b, d\}, \{c\}, \{e, f\}\}\}$*. In Figure A.10(c), we shown the unique min-cut* L' *of* (G, w) *for the subset* $\mathcal{M}' = \{\{a\}, \{f\}\}$ *of* $\mathcal{M}$*. Therefore, we have* $\sigma_{min}(\mathcal{M}') = \{\{\{a, b\}, \{c, d, e, f\}\}\}$*. We can observe that the unique partition in* $\sigma_{min}(\mathcal{M})$ *is not a refinement of the unique partition in* $\sigma_{min}(\mathcal{M}')$*. Thus,* σ_{min} *is not hierarchical.* □

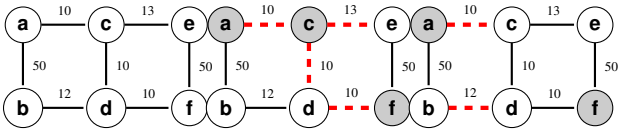


Figure A.10: From left to right: a graph (G, w), the min-cut L (dashed edges) of (G, w) for the set of markers $\mathcal{M} = \{\{a\}, \{c\}, \{f\}\}$, and the min-cut L' of (G, w) for the set of markers $\mathcal{M}' = \{\{a\}, \{f\}\}$. The partition induced by L is not a refinement of the partition induced by L'.

Definition 16 (average-cuts). *Let $\mathcal{M}$ be a marker set. An* average-cut of (G, w) for $\mathcal{M}$ *is a subset L of E such that:*

1. *the set of connected components of $(V, E \setminus L)$ is a marker-based segmentation of (G, w) for $\mathcal{M}$; and*
2. *the value $\frac{\sum_{u \in L} w(u)}{|L|}$ is minimal for all subsets of E for which statement 1 holds true.*

Property 17. *Let σ_{avg} be the operator that maps any marker set $\mathcal{M}$ into the set of partitions induced by each of the average-cuts of (G, w) for $\mathcal{M}$. The operator σ_{avg} is not hierarchical.*

Proof 3. *Let (G, w) be the weighted graph of Figure A.11(a). Let $\mathcal{M} = \{\{a\}, \{c\}, \{d\}\}$ be a marker set of (G, w). In Figure A.11(b), we show the unique average-cut L (dashed edges) of (G, w) for $\mathcal{M}$. Hence, we have $\sigma_{avg}(\mathcal{M}) = \{\{\{a\}, \{b, d\}, \{c\}\}\}$. In Figure A.11(c), we shown the unique average-cut L' of (G, w) for the subset $\mathcal{M}' = \{\{a\}, \{d\}\}$ of $\mathcal{M}$. Therefore, we have $\sigma_{avg}(\mathcal{M}') = \{\{\{a, b, c\}, \{d\}\}\}$. We can observe that the unique partition in $\sigma_{avg}(\mathcal{M})$ is not a refinement of the unique partition in $\sigma_{avg}(\mathcal{M}')$. Thus, σ_{avg} is not hierarchical.* □

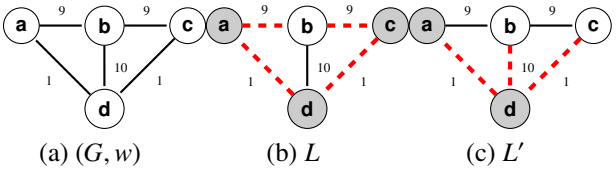


Figure A.11: From left to right: a graph (G, w), the average-cut L (dashed edges) of (G, w) for the set of markers $\mathcal{M} = \{\{a\}, \{c\}, \{d\}\}$, and the average-cut L' of (G, w) for the set of markers $\mathcal{M}' = \{\{a\}, \{d\}\}$. The partition induced by L is not a refinement of the partition induced by L'.

Let d be a distance on V, *i.e.*, a map from $V \times V$ to $\mathbb{R}^+$ such that:

- for any two vertices x and y in V, $d(x, y) = d(y, x)$;
- for any two vertices x and y in V, $d(x, y) = 0$ if and only if $x = y$; and
- for any three vertices x, y and z, we have $d(x, y) \geq d(x, z) + d(z, y)$.

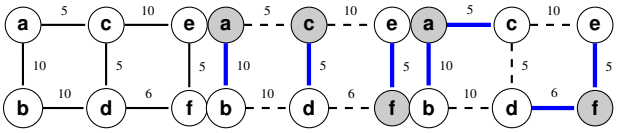
Let $\mathcal{M}$ be a marker set and let x be a vertex in V. Let π be a path from x to y such that y belongs to an element of $\mathcal{M}$. We say that π *is a d-shortest path from x to $\mathcal{M}$* if the distance $d(x, y)$ is less than the distance $d(x, z)$ for any other vertex z such that z belongs to an element of $\mathcal{M}$.

Definition 18 (shortest path forests). *Let $\mathcal{M}$ be a marker set and let d be a distance on V. Let G' be a forest of (G, w) rooted in $\mathcal{M}$. The graph G'* is a d-shortest path forest of (G, w) for $\mathcal{M}$ *if, for each vertex x in V, there is a d-shortest path π from x to $\mathcal{M}$ in G such that π is also a d-shortest path from x to $\mathcal{M}$ in G'.*

Let x and y be two vertices in V and let $\pi = (z_1, \ldots, z_\ell)$ be a path from x to y. We denote the sum $\sum_{i=1}^{\ell-1} w(\{z_i, z_{i+1}\})$ by the *weight of π*.

Property 19. *Let d be a distance on V. Let σ_s be the operator that maps any marker set $\mathcal{M}$ into the set of partitions induced by each of the d-shortest path forests of (G, w) for $\mathcal{M}$. The operator σ_s is not hierarchical.*

Proof 4. *Let (G, w) be the graph of Figure A.12(a) and let d be a distance on $V(G)$ such that, for any two vertices x and y in $V(G)$, we have $d(x, y)$ equal to the minimum among the weights of all paths from x to y. Let $\mathcal{M} = \{\{a\}, \{c\}, \{f\}\}$ and $\mathcal{M}' = \{\{a\}, \{f\}\}$ be two marker sets of (G, w). In Figures A.12(b) and (c), we show the unique d-shortest path forests of (G, w) for $\mathcal{M}$ and $\mathcal{M}'$, respectively. Hence, we have $\sigma_s(\mathcal{M}) = \{\{\{a, b\}, \{c, d\}, \{e, f\}\}\}$ and $\sigma_s(\mathcal{M}') = \{\{\{a, b, c\}, \{d, e, f\}\}\}$. We can observe that the unique partition in $\sigma_s(\mathcal{M})$ is not a refinement of the unique partition in $\sigma_s(\mathcal{M}')$. Thus, σ_s is not hierarchical.* □



(a) (G, w) (b) G' (c) G''

Figure A.12: From left to right: a graph (G, w), the d-shortest path forest G' (all vertices plus blue edges) of (G, w) for the marker set $\mathcal{M} = \{\{a\}, \{c\}, \{f\}\}$, and the d-shortest path forest G'' of (G, w) for the marker set $\mathcal{M}' = \{\{a\}, \{f\}\}$. The partition induced by G' is not a refinement of the partition induced by G''.

Appendix B. Proofs of theorems and properties

Appendix B.1. Proof of Theorem 5

In order to prove Theorem 5, we will use an equivalent formulation provided in Property 12 of [13], which requires to introduce the following new definitions.

Let $\prec$ be an altitude ordering for w, let $\mathcal{B}_\prec$ be the binary partition hierarchy by $\prec$ and let $\mathcal{S} = (M_1, \ldots, M_n)$ be a sequence of minima of (G, w). Let X be a region of $\mathcal{B}_\prec$. Following the terminology of [13], the *extinction value of X for $\mathcal{S}$* is zero if there is no minimum M of w such that M is a subset of X and, otherwise, it is the highest index k such that M_k is a subset of X. Let ϵ be the map from the regions of $\mathcal{B}_\prec$ into $\mathbb{R}$ such that, for any region Y of $\mathcal{B}_\prec$, $\epsilon(Y)$ is the extinction map of Y for $\mathcal{S}$. We say that *ϵ is the extinction map (for $\prec$ and $\mathcal{S}$)* or that *ϵ is an extinction map* (for $\prec$).

Let $\prec$ be an altitude ordering for w, let $\mathcal{B}_\prec$ be the binary partition hierarchy by $\prec$ and let $\mathcal{S} = (M_1, \ldots, M_n)$ be a sequence of minima of w. Let u be a building edge of $\mathcal{B}_\prec$ and let X be the region of $\mathcal{B}_\prec$ whose building edge is u. The *persistence value of u (for $\prec$ and $\mathcal{S}$)* is the minimum of the extinction values of the children of X. Let ρ be the map from the building edges of $\mathcal{B}_\prec$ into $\mathbb{R}$ such that, for any building edge u of $\mathcal{B}_\prec$, $\rho(u)$ is the persistence value of u. We say that *ρ is the persistence map (for $\prec$ and $\mathcal{S}$)*. We denote by B_i the set of building edges of $\mathcal{B}_\prec$ whose persistence value is lower than or equal to i.

Definition 20. (hierarchy induced by an altitude ordering and a sequence of minima [13]) *Let $\prec$ be an altitude ordering for w, let $\mathcal{S} = (M_1, \ldots, M_n)$ be a sequence of minima of w and let ρ be the persistence map for $\prec$ and for $\mathcal{S}$. The sequence of partitions $(CC(V, B_0), \ldots, CC(V, B_{n-1}))$ is a hierarchy called* the hierarchy induced by $\prec$ and $\mathcal{S}$.

Recall that Theorem 5 states the equivalence between:

A A hierarchy $\mathcal{H}$ is a hierarchical watershed of (G, w).

B There exists an altitude ordering $\prec$ for w such that the saliency map $\Phi(\mathcal{H})$ is one-side increasing for $\mathcal{B}_\prec$

In order to prove this theorem, we will use another equivalent property:

C There exists an altitude ordering $\prec$ for w and a sequence of minima $\mathcal{S}$ of w such that $\mathcal{H}$ is induced by $\prec$ and $\mathcal{S}$.

Property 12 of [13] established the equivalence between A and C. We will then show the equivalence between B and C. This proof is decomposed in Lemma 25 (C $\Rightarrow$ B) and Lemma 26 (B $\Rightarrow$ C).

We now introduce some new lemmas needed to prove Lemma 25.

Lemma 21. *Let $\prec$ be an altitude ordering for w and let ϵ be an extinction map for $\mathcal{B}_\prec$. Let X and Y be two regions of $\mathcal{B}_\prec$. If $X \subseteq Y$, then $\epsilon(X) \leq \epsilon(Y)$.*

Proof 5. *Since $\mathcal{B}_\prec$ is a hierarchy, we can affirm that, for any two regions Y and Z of $\mathcal{B}_\prec$, if $Y \subseteq Z$, then all minima of w included in Y are also included in Z and, therefore, $\epsilon(Y) \leq \epsilon(Z)$.* □

Lemma 22. *Let $\prec$ be an altitude ordering for w and let ϵ be an extinction map for $\prec$. The range of ϵ is $\{0, \ldots, n\}$.*

Proof 6. *To prove that the range of ϵ is $\{0, \ldots, n\}$, we will prove that $\{0, \ldots, n\} \subseteq range(\epsilon)$ and that $range(\epsilon) \subseteq \{0, \ldots, n\}$.*

1. *$\{0, \ldots, n\} \subseteq range(\epsilon)$. First, we prove that 0 is in $range(\epsilon)$. Let u be the lowest edge of $E(\mathcal{B}_\prec)$ for $\prec$. We can say that u is in a minimum of w. Moreover, the children of R_u are necessarily singletons.*

Hence, the extinction value of both children of u is zero. Now, we will prove for i in $\{1, \ldots, n\}$. By [28] (page 7), any minimum of w is a region of $\mathcal{B}$. Therefore, for any i in $\{1, \ldots, n\}$, there is a region of $\mathcal{B}_{\prec}$ whose extinction value is i.

2. *range$(\epsilon) \subseteq \{0, \ldots, n\}$. For any region X of $\mathcal{B}_{\prec}$, if X contains at least one minimum of w, then its extinction value is in $\{1, \ldots, n\}$. Otherwise, the extinction value of X is zero.* □

Lemma 23. *Let $\prec$ be an altitude ordering on the edges of G for w, let $\mathcal{S} = (M_1, \ldots, M_n)$ be a sequence of minima of w and let ρ be the persistence map for $\prec$ and for $\mathcal{S}$. The range of ρ is $\{0, \ldots, n-1\}$.*

Proof 7. *Let ϵ denote the extinction map for $\prec$ and $\mathcal{S}$. We will prove that (1) for any building edge u of $\mathcal{B}_{\prec}$, $\rho(u)$ is in $\{0, \ldots, n-1\}$, and that, (2) for any i in $\{0, \ldots, n-1\}$, there is a building edge of $\mathcal{B}_{\prec}$ whose persistence value is i.*

1. *$\{0, \ldots, n-1\} \subseteq$ range(ρ). First, we prove that 0 is in range(ρ). By Lemma 22, there is a region X of $\mathcal{B}_{\prec}$ whose extinction value is zero. Therefore, the persistence value of the parent of X is equal to zero. Now, we will prove that any i in $\{1, \ldots, n-1\}$ is in range(ρ). Let i be a value in $\{1, \ldots, n-1\}$. By [28] (page 7), the minimum M_i is a region of $\mathcal{B}_{\prec}$. Then, there is a region of $\mathcal{B}_{\prec}$ whose extinction value is i. Let X be the largest region of $\mathcal{B}_{\prec}$ whose extinction value is i. We can say that $X \neq V$ because M_n is included in V and, therefore, $\epsilon(V) = n$. Let Z be the parent of X. We can infer that the extinction value $\epsilon(Z)$ of Z is strictly greater than i. Therefore, there is a minimum M_j with $j > i$ included in the sibling of X. Hence, the extinction value of sibling(X) is also strictly greater than i. Then, the persistence value of the building edge of Z, being the minimum of the extinction value of its children, is i.*

2. *range$(\rho) \subseteq \{0, \ldots, n-1\}$. Let u be an edge in $E(\mathcal{B}_{\prec})$. By Lemma 22, and as the persistence value of u is equal to the extinction value of a child of R_u, we have that $\rho(u)$ is in $\{0, \ldots, n\}$. Moreover, the persistence value $\rho(u)$ of u is lower than n because, if the extinction value of one child X of R_u is n, then the minimum M_n is included in X and M_n is not included in sibling(X), which implies that the extinction value of sibling(X) is strictly lower than n. Therefore, since $\rho(u) = min\{\epsilon(X), \epsilon(sibling(X))\}$, the persistence value of u is strictly lower than n. Thus, we have that range$(\rho) \subseteq \{0, \ldots, n-1\}$.* □

Lemma 24. *Let $\prec$ be an altitude ordering for w, let $\mathcal{S} = (M_1, \ldots, M_n)$ be a sequence of minima of w and let ρ be the persistence map for $\prec$ and for $\mathcal{S}$. Let $\mathcal{H}$ be the hierarchy induced by $\prec$ and $\mathcal{S}$. For any building edge u of $\mathcal{B}_{\prec}$, we have $\Phi(\mathcal{H})(u) = \rho(u)$.*

Proof 8. *By definition, $\mathcal{H}$ is the sequence $(CC(V, B_0), \ldots, CC(V, B_{n-1}))$ such that, for any i in $\{0, \ldots, n-1\}$, B_i is the set of building edges of $\mathcal{B}_{\prec}$ whose persistence values is lower than or equal to i. Let $u = \{x, y\}$ be a building edge of $\mathcal{B}_{\prec}$ and let i be the persistence value of u. We can say that x and y are in the same region of $CC(V, B_i)$ but in distinct regions of $CC(V, B_{i-1})$ if $i \neq 0$. Therefore, since $CC(V, B_i)$ is the i-th partition of $\mathcal{H}$, by the definition of saliency maps, we have $\Phi(\mathcal{H})(u) = i$.* □

Lemma 25. *Let $\prec$ be an altitude ordering for w, let $\mathcal{S}$ be a sequence of minima of w and let $\mathcal{H}$ be the hierarchy induced by $\prec$ and by $\mathcal{S}$. The saliency map $\Phi(\mathcal{H})$ of $\mathcal{H}$ is one-side increasing for $\mathcal{B}_{\prec}$.*

Proof 9. *In order to prove that $\Phi(\mathcal{H})$ is one-side increasing for $\mathcal{B}_{\prec}$, by Definition 4, we need to prove that the following three statements hold true:*

1. *$\{\Phi(\mathcal{H})(e) \mid e \in E(\mathcal{B}_{\prec})\} = \{0, \ldots, n-1\}$;*
2. *for any edge u in $E(\mathcal{B}_{\prec})$, $\Phi(\mathcal{H})(u) > 0$ if and only if u is a watershed-cut edge for $\mathcal{B}_{\prec}$; and*
3. *for any edge u in $E(\mathcal{B}_{\prec})$, there exists a child R of R_u such that $\Phi(\mathcal{H})(u) \geq \vee\{\Phi(\mathcal{H})(v)$ such that R_v is included in R$\}$, where $\vee\{\} = 0$.*

In the sequel of this proof, let ρ and ϵ be respectively the persistence map and the extinction map for $\prec$ and $\mathcal{S}$.

1. *By Lemma 24, we have $\{\Phi(\mathcal{H})(e) \mid e \in E(\mathcal{B}_{\prec})\} = \{\rho(e) \mid e \in E(\mathcal{B}_{\prec})\}$. Then, as Lemma 23 states that the range of ρ is $\{0, \ldots, n-1\}$, we can conclude that $\{\Phi(\mathcal{H})(e) \mid e \in E(\mathcal{B}_{\prec})\}$ is the set $\{0, \ldots, n-1\}$.*

2. *Let u be a building edge of $\mathcal{B}_\prec$. Given the following propositions:*

   (a) u is a watershed-cut edge for $\mathcal{B}_\prec$

   (b) $\Phi(\mathcal{H})(u) > 0$

   we will prove that (a) implies (b), and that not (b) implies not (a).

   If u is a watershed-cut edge for $\mathcal{B}_\prec$, then both children of R_u contain at least one minimum of w. Therefore, the extinction value of both children of R_u is non-zero and, consequently, the persistence value $\rho(u)$ of u is non-zero. Moreover, by Lemma 24, in this case we have $\Phi(\mathcal{H})(e) = \rho(e)$ for any building edge e of $\mathcal{B}_\prec$. Thus, $\Phi(\mathcal{H})(u)$, being equal to $\rho(u)$, is non-zero.

   On the other hand, if u is not a watershed-cut edge for $\mathcal{B}_\prec$, then there is a child X of R_u which does not contain any minimum of w. Therefore, the extinction value of X is equal to 0: $\epsilon(X) = 0$. Since, by definition $\rho(u) = min\{\epsilon(X), \epsilon(sibling(X))\}$ and the minimal extinction value is zero, we can say that $\rho(u) = 0$. Again, by Lemma 24, in this case we have $\Phi(\mathcal{H})(e) = \rho(e)$ for any building edge e of $\mathcal{B}_\prec$ and thus, $\Phi(\mathcal{H})(u)$, being equal to $\rho(u)$, is equal to 0.

3. *Let u be a building edge of $\mathcal{B}_\prec$. The persistence value of u is the extinction value of a child X of R_u. Let X be a child of R_u such that $\rho(u)$, the persistence value of u, is equal to $\epsilon(X)$, the extinction value of X. By Lemma 21, for any region Y of $\mathcal{B}_\prec$ such that $Y \subseteq X$, we have $\epsilon(Y) \leq \epsilon(X)$ and, as $X \subseteq R_u$, $\epsilon(Y) \leq \epsilon(R_u)$. Let v be the building edge of a region $Z \subseteq X$. Then, we can say that the extinction value of both children of Z is less than or equal to the extinction value $\epsilon(X)$. Hence, $\rho(v) \leq \epsilon(X)$ and, then, $\rho(v) \leq \rho(u)$. By Lemma 24, we can conclude that $\Phi(\mathcal{H})(v) \leq \Phi(\mathcal{H})(u)$. Hence, $\Phi(\mathcal{H})(u) \geq \vee\{\Phi(\mathcal{H})(v)$ such that R_v is included in $X\}$.* □

Lemma 26. *Let $\prec$ be an altitude ordering of (G, w) and let $\mathcal{H}$ be a hierarchy on V such that $\Phi(\mathcal{H})$ is one-side increasing for $\mathcal{B}_\prec$. Then there exists a sequence of minima $\mathcal{S}$ such that $\mathcal{H}$ is the hierarchy induced by $\prec$ and $\mathcal{S}$.*

In order to prove Lemma 26, we establish the following lemma, which results from the generalization of Lemmas 6 and 7 of [20].

Lemma 27. *Let $\prec$ be an altitude ordering for w and let f be map from E into $\mathbb{R}$ such that f is one-side increasing for $\mathcal{B}_\prec$. Then, there exists an extinction map ϵ for $\prec$ such that, for any building edge u of $\mathcal{B}_\prec$:*

$$f(u) = min\{\epsilon(R) \text{ such that } R \text{ is a child of } R_u\}.$$

In [20], the authors established the result in Lemma 27 in the case where (G, w) is a tree with pairwise distinct edge weights. They introduced estimated (approximated) extinction maps such that, for any map f from E into $\mathbb{R}$ and for any altitude ordering $\prec$ for w, the estimated extinction map for f and $\prec$ is an extinction map for $\prec$ if and only if f is one-side increasing for $\mathcal{B}_\prec$. Then, they proved that, given a map f from E into $\mathbb{R}$ and given an altitude ordering $\prec$ for w such that f is one-side increasing for $\mathcal{B}_\prec$, this implies that, for any building edge u for $\mathcal{B}_\prec$, we have $f(u) = min\{\epsilon(R)$ such that R is a child of $R_u\}$ where ϵ is the estimated extinction map for f and $\prec$. The generalization of this result to the case where (G, w) is not a tree with pairwise distinct edge weights will be presented in an extended version of [20]. The proof of Lemma 27 can be found in a preprint of the extended version of [20] available in `https://hal.archives-ouvertes.fr/hal-02280023/file/characterization.pdf`.

The following lemma, established in [12], links MSTs and QFZ hierarchies.

Lemma 28 (Theorem 4 of [12]). *A subgraph G' of G is a MST of (G, w) if and only if:*

1. *the QFZ hierarchy of G' and G are the same; and*
2. *the graph G' is minimal for statement 1,* i.e.*, for any subgraph G'' of G', if the quasi-flat zone hierarchy of G'' for w is the one of G for w, then we have $G'' = G'$.*

Lemma 29. *Let $\mathcal{B}$ be a binary partition hierarchy of (G, w) and let $\mathcal{H}$ be a hierarchy on V such that $\Phi(\mathcal{H})$ is one-side increasing for $\mathcal{B}$. Then $(V, E(\mathcal{B}))$ is a MST of $(G, \Phi(\mathcal{H}))$.*

Proof 10. *Let α denote the sum of the weight of the edges in $E(\mathcal{B})$ in the map $\Phi(\mathcal{H})$: $\alpha = \sum_{e \in E(\mathcal{B})} \Phi(\mathcal{H})(e)$. As $\Phi(\mathcal{H})$ is one-side increasing for $\mathcal{B}$, by the condition 1 of Definition 4, we can affirm that $\alpha = 0 + 1 + \cdots + n - 1$. In order to prove that $(V, E(\mathcal{B}))$ is a MST of $(G, \Phi(\mathcal{H}))$, we will prove that, for any MST G' of $(G, \Phi(\mathcal{H}))$, the sum*

of the weight of the edges in G' is greater than or equal to α. Let G' be a MST of $(G, \Phi(\mathcal{H}))$. As G' is a MST of $(G, \Phi(\mathcal{H}))$, by the condition 1 of Lemma 28, we have that G and G' have the same quasi-flat zones hierarchy: $\mathcal{QFZ}(G, \Phi(\mathcal{H})) = \mathcal{QFZ}(G', \Phi(\mathcal{H}))$. As $\Phi(\mathcal{H})$ is the saliency map of $\mathcal{H}$, we have that $\mathcal{H} = \mathcal{QFZ}(G, \Phi(\mathcal{H}))$. Therefore, $\mathcal{H} = \mathcal{QFZ}(G', \Phi(\mathcal{H}))$. Let i be a value in $\{1, \dots, n-1\}$. By the condition 1 of Definition 4, we can say that $\{1, \dots, n-1\}$ is a subset of the range of $\Phi(\mathcal{H})$. Therefore, $\mathcal{H}$ is composed of at least n distinct partitions. Let $\mathcal{H}$ be the sequence $(\mathbf{P}_0, \dots, \mathbf{P}_{n-1}, \dots)$. Since the partitions $\mathbf{P}_i$ and $\mathbf{P}_{i-1}$ are distinct, then there exists a region in $\mathbf{P}_i$ which is not in $\mathbf{P}_{i-1}$. Therefore, there is a region X of $\mathbf{P}_i$ which is composed of several regions $\{R_1, R_2, \dots\}$ of $\mathbf{P}_{i-1}$. Then, there are two adjacent vertices x and y such that x and y are in distinct regions in $\{R_1, R_2, \dots\}$. Let x and y be two adjacent vertices such that x and y are in distinct regions in $\{R_1, R_2, \dots\}$. Hence, the lowest j such that x and y belong to the same region of $\mathbf{P}_j$ is i. Thus, there exists an edge $u = \{x, y\}$ in $E(\mathcal{B})$ such that $\Phi(\mathcal{H})(u) = i$. Hence, the sum of the weight of the edges of G' is at least $1 + \dots + n - 1$, which is equal to α. Therefore, the graph $(V, E(\mathcal{B}))$ is a MST of $(G, \Phi(\mathcal{H}))$. □

Proof 11 (of Lemma 26). *As $\Phi(\mathcal{H})$ is one-side increasing for $\mathcal{B}_\prec$, then, by Lemma 27, there is an extinction map ϵ such that, for any building edge u of $\mathcal{B}_\prec$, $\Phi(\mathcal{H})(u) = min\{\epsilon(R)$ such that R is a child of $R_u\}$. Since ϵ is an extinction map, then there is a sequence $\mathcal{S} = (M_1, \dots, M_n)$ of minima of w such that ϵ is the extinction map for $\prec$ and $\mathcal{S}$, and such that, for any region R of $\mathcal{B}_\prec$, $\epsilon(R) = \vee\{i \mid M_i \subseteq R\}$. Let G' denote the graph $(V, E(\mathcal{B}_\prec))$. By Lemma 29, G' is a MST of $(G, \Phi(\mathcal{H}))$ and, consequently, by Lemma 28, $\mathcal{H} = \mathcal{QFZ}(G', \Phi(\mathcal{H}))$. Let ρ denote the persistence map for $\prec$ and for $\mathcal{S}$. Since $\Phi(\mathcal{H})(u) = min\{\epsilon(R)$ such that R is a child of $R_u\}$, we have that, for any building edge u, $\Phi(\mathcal{H})(u)$ is the persistence value $\rho(u)$ of u. Then, $\mathcal{QFZ}(G', \Phi(\mathcal{H})) = \mathcal{QFZ}(G', \rho)$. By definition, $\mathcal{QFZ}(G', \rho)$ is precisely the hierarchy induced by $\prec$ and by $\mathcal{S}$.* □

Appendix B.2. Proof of Property 7

To prove Property 7, we first provide in Property 30 a sufficient condition for a hierarchy to be a flattened hierarchical watershed of (G, w).

Property 30. *Let $\mathcal{H}$ be a hierarchy on V and let $\mathcal{B}$ be a binary partition hierarchy of (G, w) such that:*

1. *$(V, E(\mathcal{B}))$ is a MST of $(G, \Phi(\mathcal{H}))$; and*
2. *for any edge u in $E(\mathcal{B})$, if u is not a watershed-cut edge for $\mathcal{B}$, then $\Phi(\mathcal{H})(u) = 0$; and*
3. *for any edge u in $E(\mathcal{B})$, there exists a child R of R_u such that $\Phi(\mathcal{H})(u) \geq \vee\{\Phi(\mathcal{H})(v)$ such that R_v is included in $R\}$, where $\vee\{\} = 0$.*

Then $\mathcal{H}$ is a flattened hierarchical watershed of (G, w).

The reader can observe that the statement 3 of the above property is precisely the statement 3 of the definition of one-side increasing maps (Definition 4), and that the statement 2 is an implication of the statement 2 of Definition 4. The statement 1 of the above property corresponds to a property of one-side increasing maps established in Lemma 29.

In order to prove Property 30, we first state two auxiliary lemmas. From Property 10 of [13], we can deduce the following lemma linking binary partition hierarchies and MSTs.

Lemma 31. *Let $\mathcal{B}$ be a binary partition hierarchy of (G, w). The graph $(V, E(\mathcal{B}))$ is a MST of (G, w).*

By Property 12 of [13] linking hierarchical watersheds and hierarchies induced by an altitude ordering and a sequence of minima, and by Lemma 28, we infer the following lemma.

Lemma 32. *Let G' be a MST of (G, w) and let $\mathcal{H}$ be a hierarchical watershed of (G', w). Then $\mathcal{H}$ is also a hierarchical watershed of (G, w).*

Proof 12 (of Property 30). *Let $\mathcal{H}$ be a hierarchy on V such that there is a binary partition hierarchy $\mathcal{B}$ of (G, w) such that:*

1. *$(V, E(\mathcal{B}))$ is a MST of $(G, \Phi(\mathcal{H}))$; and*
2. *for edge u in $E(\mathcal{B})$, if u is not a watershed-cut edge for $\mathcal{B}$, then $\Phi(\mathcal{H})(u) = 0$; and*
3. *for edge u in $E(\mathcal{B})$, there exists a child R of R_u such that $\Phi(\mathcal{H})(u) \geq \vee\{\Phi(\mathcal{H})(v)$ such that R_v is included in $R\}$, where $\vee\{\} = 0$.*

We will prove that $\mathcal{H}$ is a flattened hierarchical watershed of (G, w). To this end, we will prove that there is a hierarchical watershed $\mathcal{H}_w$ of (G, w) such that any partition of $\mathcal{H}$ is also a partition of $\mathcal{H}_w$. Let G' denote the graph $(V, E(\mathcal{B}))$. By Lemma 31, G' is a MST of (G, w). Moreover, by Lemma 32, given a hierarchical watershed $\mathcal{H}_w$ of a MST of (G, w), we can say that $\mathcal{H}_w$ is also a hierarchical watershed of (G, w). Hence, we can simply prove that there is a hierarchical watershed $\mathcal{H}_w$ of (G', w) such that any partition of $\mathcal{H}$ is also a partition of $\mathcal{H}_w$.

To define the hierarchy $\mathcal{H}_w$, we first define a map f from $E(\mathcal{B})$ into $\mathbb{R}$ such that f is one-side increasing for $\mathcal{B}$. Since G' is a tree, by the definition of saliency maps, we can say that f is the saliency map of the hierarchy $\mathcal{QFZ}(G', f)$. By Theorem 5, as f is one-side increasing for $\mathcal{B}$, we may say that $\mathcal{QFZ}(G', f)$ is a hierarchical watershed of (G', w).

In the map f, the edges which are not watershed-cut edges for $\mathcal{B}$ are assigned to zero, and the watershed-cut edges for $\mathcal{B}$ are ranked according to their weights in w and in $\Phi(\mathcal{H})$. Let $\prec$ be an altitude ordering for w such that $\mathcal{B}$ is the binary partition hierarchy for $\prec$. Let $\prec_2$ be a total ordering on the set $\{u \text{ is a watershed-cut edge for } \mathcal{B}_\prec\}$ such that, for any two watershed-cut edges u and v for $\mathcal{B}_\prec$, we have $u \prec_2 v$ if and only if $\Phi(\mathcal{H})(u) < \Phi(\mathcal{H})(v)$ or if $\Phi(\mathcal{H})(u) = \Phi(\mathcal{H})(v)$ and $u \prec v$. The map f is defined as follows:

$$f(u) = \begin{cases} 0 & \text{if } u \text{ is not a watershed-cut edge for } \mathcal{B} \\ \text{rank of } u \text{ for } \prec_2 & \text{otherwise} \end{cases} \tag{B.1}$$

We first demonstrate that f is one-side increasing for $\mathcal{B}$.

1. *By the definition of f, as there are $n-1$ watershed-cut edges for $\mathcal{B}$, we can say that, for any i in $\{1, \ldots, n-1\}$, there is a watershed-cut edge u for $\mathcal{B}$ such that the rank of u for $\prec_2$ is i and, consequently, $f(u) = i$. On the other hand, as w has at least one minimum, there is at least one edge e in $E(\mathcal{B})$ such that e is not a watershed-cut edge for $\mathcal{B}$ and such that $f(e) = 0$. Hence, we have $\{f(e) \mid u \in E(\mathcal{B})\} = \{0, \ldots, n-1\}$. Therefore, the statement 1 of Definition 4 holds true for f.*

2. *For any edge u, by the definition of f, $f(u)$ is non-zero if and only if u is not a watershed-cut edge for $\mathcal{B}$, so the statement 2 of Definition 4 holds true for f.*

3. *Let u be a building edge for $\mathcal{B}$. If u is not a watershed-cut edge for $\mathcal{B}$, then there is a child X of R_u such that there is no minimum of w included in X. Hence, none of the building edges of the descendants of X is a watershed-cut edge for $\mathcal{B}$. By the definition of f, we have $f(u) = 0$ and, for any edge v such that $R_v \subseteq X$, we have $f(v) = 0$. Hence, $f(u) \geq \vee\{f(v)$ such that R_v is included in $X\}$. Otherwise, let us assume that u is a watershed-cut edge for $\mathcal{B}$. Then there is at least one minimum of w included in each child of R_u. By the hypothesis 3, there is a child X of R_u such that $\Phi(\mathcal{H})(u) \geq \vee\{\Phi(\mathcal{H})(v)$ such that R_v is included in $X\}$. Let X be the child of R_u such that $\Phi(\mathcal{H})(u) \geq \vee\{\Phi(\mathcal{H})(v)$ such that R_v is included in $X\}$. Let e be a building edge of $\mathcal{B}$ such that $R_e \subseteq X$. If e is not a watershed-cut edge for $\mathcal{B}$, then $f(e) = 0$ and, consequently, $f(u) > f(e)$. Otherwise, if e is a watershed-cut edge for $\mathcal{B}$, then we have $\Phi(\mathcal{H})(u) \geq \Phi(\mathcal{H})(e)$ and $e \prec u$, which implies that $e \prec_2 u$. Consequently, by the definition of f, we have $f(u) > f(e)$. Therefore, $f(u) \geq \vee\{f(v)$ such that R_v is included in $X\}$. Then, the third condition of Definition 4 holds true for f.*

Hence, f is one-side increasing for $\mathcal{B}$ and, as stated previously, $\mathcal{QFZ}(G', f)$ is a hierarchical watershed of (G', w) (resp. (G, w)). Now, we only need to prove that any partition of $\mathcal{H}$ is a partition of $\mathcal{QFZ}(G', f)$. By the hypothesis 1, G' is a MST of $(G, \Phi(\mathcal{H}))$. Then, by Lemma 28, we can say that $\mathcal{H}$ is the QFZ hierarchy of $(G', \Phi(\mathcal{H}))$. We will prove that any partition of $\mathcal{QFZ}(G', \Phi(\mathcal{H}))$ is also a partition of $\mathcal{QFZ}(G', f)$.

Let the range of $\Phi(\mathcal{H})$ be the set $\{0, \ldots, \ell\}$: $\{\Phi(\mathcal{H})(u) \mid u \in E(\mathcal{B})\} = \{0, \ldots, \ell\}$. Let λ be a value in $\{0, \ldots, \ell\}$. Let $G'_{\lambda,\Phi(\mathcal{H})}$ be the λ-level set of $(G', \Phi(\mathcal{H}))$. Let α be the greatest value in $\{f(u) \mid u \in E(G'_{\lambda,\Phi(\mathcal{H})})\}$. We will prove that the α-level set of (G', f) is equal to the λ-level set of $(G', \Phi(\mathcal{H}))$. Since α is the greatest value in the set $\{f(u) \mid u \in E(G'_{\lambda,\Phi(\mathcal{H})})\}$, we can see that any edge v in the λ-level set of $(G', \Phi(\mathcal{H}))$ also belongs to the α-level set of (G', f). Now, we also need to prove that there is no edge u in the α-level set of (G', f) such that u is not in the λ-level

set of $(G', \Phi(\mathcal{H}))$.

Let u *be an edge which is not in the* λ*-level set of* $(G', \Phi(\mathcal{H}))$*. Then,* $\Phi(\mathcal{H})(u) > \lambda$ *and, for any edge* v *in the* λ*-level set of* $(G', \Phi(\mathcal{H}))$*, we have* $\Phi(\mathcal{H})(u) > \Phi(\mathcal{H})(v)$*. Since the minimum value of* λ *is zero, we can say that* $\Phi(\mathcal{H})(u) > 0$ *and, by the hypothesis 2,* u *is a watershed-cut edge for* $\mathcal{B}$*. Let* v *be an edge in the* λ*-level set of* $(G', \Phi(\mathcal{H}))$*. Since* $\Phi(\mathcal{H})(u) > \Phi(\mathcal{H})(v)$*, if* v *is a watershed-cut edge for* $\mathcal{B}$*, then* $v \prec_2 u$ *and* $f(u) > f(v)$*. Otherwise, if* v *is not a watershed-cut edge for* $\mathcal{B}$*, by the definition of* f*, we have* $f(v) = 0$ *and* $f(u) > f(v)$*. Thus, for any edge* v *in the* λ*-level set of* $(G', \Phi(\mathcal{H}))$*, we have* $f(u) > f(v)$ *and, therefore,* $f(u) > \alpha$*. Then,* u *is not in the* α*-level set of* (G', f)*.*

Therefore, we can conclude that the α*-level set of* (G', f) *is equal to the* λ*-level set of* $(G', \Phi(\mathcal{H}))$*. As the partitions of* $\mathcal{H}$ *are given by the set of connected components of the level sets of* $(G', \Phi(\mathcal{H}))$*, we can affirm that any partition of* $\mathcal{H}$ *is also a partition of* $\mathcal{QFZ}(G', f)$*. Therefore, there is a hierarchical watershed* $\mathcal{H}_w = \mathcal{QFZ}(G', f)$ *of* (G', w) *(resp.* (G, w)*) such that any partition of* $\mathcal{H}$ *is also a partition of* $\mathcal{H}_w$*. Then,* $\mathcal{H}$ *is a flattened hierarchical watershed of* (G', w) *(resp.* (G, w)*).* □

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hierarchical watersheds of (G, w) and let $\mathcal{B}$ be a binary partition hierarchy of (G, w) such that both $\Phi(\mathcal{H}_1)$ and $\Phi(\mathcal{H}_2)$ are one-side increasing for $\mathcal{B}$. Let f_3 denote the map $\curlywedge(\Phi(\mathcal{H}_1), \Phi(\mathcal{H}_2))$. We will prove that the hierarchy $\mathcal{QFZ}(G, f_3)$ is a flattened hierarchical watershed of (G, w). To this end, by Property 30, we will prove that the following statements hold true:

1. $(V, E(\mathcal{B}))$ is a MST of (G, f_3); and
2. for any edge u in $E(\mathcal{B})$, if u is not a watershed-cut edge for $\mathcal{B}$, then $f_3(u) = 0$; and
3. for any edge u in $E(\mathcal{B})$, there exists a child R of R_u such that $f_3(u) \geq \vee\{f_3(v)$ such that R_v is included in $R\}$, where $\vee\{\} = 0$.

The following Lemmas 33, 36 and 37 prove respectively that the conditions 1, 2 and 3 for $\mathcal{QFZ}(G, f_3)$ to be a flattened hierarchical watershed of (G, w) hold true.

Lemma 33. *Let* f_1 *and* f_2 *be two maps from* E *into* $\mathbb{R}$ *and let* G' *be a subgraph of* G *such that* G' *is a MST of both* (G, f_1) *and* (G, f_2)*. Then* G' *is also a MST of* $(G, \curlywedge(f_1, f_2))$*.*

In order to prove Lemma 33, we define cycles in the context of graphs and we state two well-known properties of spanning trees in Lemmas 34 and 35.

Let x and y be two vertices in V and let $\pi = (x_0, \ldots, x_p)$ be a path from x to y. For any edge $u = \{x_{i-1}, x_i\}$ for i in $\{1, \ldots, p\}$, we say that u *is in* π or that π *includes* u. We say that π *is a cycle* if $x_0 = x_p$ and $p > 1$.

Lemma 34. *Let* G' *be a spanning tree of* (G, w) *and let* u *be an edge in* $E \setminus E(G')$*. Then* $(V, E(G') \cup \{u\})$ *contains a cycle* π *that includes* u*.*

Lemma 35. *Let* G' *be a spanning tree of a weighted graph* (G, f)*. Let* u *be an edge in* $E \setminus E(G')$ *and let* π *be the cycle of* $(V, E(G') \cup \{u\})$ *which includes* u*. The graph* G' *is a MST of* (G, f) *if and only if* $f(u) \geq f(v)$ *for any edge* v *in* π*.*

Proof 13 (of Lemma 33). *Let* f_3 *denote the map* $\curlywedge(f_1, f_2)$*. Let* u *be an edge in* $E \setminus E(G')$*. As* G' *is a spanning tree, by Lemma 34, the graph* $(V, E(G') \cup \{u\})$ *contains a cycle* π *which includes the edge* u*. Since* G' *is a MST for* (G, f_1) *and for* (G, f_2)*, by the forward implication of Lemma 35, for any edge* v *in the cycle* π*, we have* $f_1(v) \leq f_1(u)$ *and* $f_2(v) \leq f_2(u)$*. Therefore, for any edge* v *in the cycle* π*, we have* $min(f_1(v), f_2(v)) \leq min(f_1(u), f_2(u))$ *and, consequently,* $f_3(v) \leq f_3(u)$*. Hence, for any edge* v *in* π*, we have* $f_3(u) \geq f_3(v)$*. Thus, by the backward implication of Lemma 35,* G' *is a MST of* (G, f_3)*.* □

The following lemma proves that the condition 2 for $\mathcal{QFZ}(G, f_3)$ to be a flattened hierarchical watershed hold true.

Lemma 36. *Let* f_1 *and* f_2 *be two maps from* E *into* $\mathbb{R}$ *and let* $\mathcal{B}$ *be a binary partition hierarchy of* (G, w) *such that* f_1 *and* f_2 *are one-side increasing for* $\mathcal{B}$*. Let* f_3 *denote the map* $\curlywedge(f_1, f_2)$*. Then for any edge* u *in* $E(\mathcal{B})$*, if* u *is not a watershed-cut edge for* $\mathcal{B}$*, then* $f_3(u) = 0$*.*

Proof 14. *Let* u *be an edge in* $E(\mathcal{B})$*. If* u *is not a watershed-cut edge for* $\mathcal{B}$*, then, by the statement 2 of Definition 4, we have* $f_1(u) = 0$ *and* $f_2(u) = 0$*. Therefore,* $f_3(u) = min(0, 0) = 0$*.* □

The following lemma proves that the condition 3 for $\mathcal{QFZ}(G, f_3)$ to be a flattened hierarchical watershed holds true.

Lemma 37. *Let f_1 and f_2 be two maps from E into $\mathbb{R}$ and let $\mathcal{B}$ be a binary partition hierarchy of (G, w) such that f_1 and f_2 are one-side increasing for $\mathcal{B}$. Let f_3 denote $\curlywedge(f_1, f_2)$. Then, for any building edge u of $\mathcal{B}$, there exists a child R of R_u such that $f_3(u) \geq \vee\{f_3(v)$ such that $R_v \subseteq R\}$.*

Proof 15. *Since f_1 (resp. f_2) is one-side increasing for $\mathcal{B}$, by the statement 3 of Definition 4, we have that, for any building edge u of $\mathcal{B}$, $f_1(u) \geq \vee\{f_1(v) \mid R_v \subseteq X\}$ (resp. $f_2(u) \geq \vee\{f_2(v) \mid R_v \subseteq X\}$) for a child X of R_u. We need to prove that, for any building edge u of $\mathcal{B}$, $f_3(u) \geq \vee\{f_3(v) \mid R_v \subseteq X\}$ for a child X of R_u. Let u be a building edge of $\mathcal{B}$. As $f_3(u) = min(f_1(u), f_2(u))$, we should consider the following cases: (1) $f_3(u) = f_1(u)$; and (2) $f_3(u) = f_2(u)$.*

1. *Let us assume that $f_3(u) = f_1(u)$. Let X and Y be the children of R_u. If $f_1(u) \geq \vee\{f_1(v) \mid R_v \subseteq X\}$ (resp. $f_1(u) \geq \vee\{f_1(v) \mid R_v \subseteq Y\}$), we can affirm that $f_3(u) \geq \vee\{f_1(v) \mid R_v \subseteq X\}$ (resp. $f_3(u) \geq \vee\{f_1(v) \mid R_v \subseteq Y\}$) as well. Since $f_3(e) = min(f_1(e), f_2(e))$ for any edge e in E, we can affirm that $f_3(e) \leq f_1(e)$ for any edge e in E and, therefore, $f_3(u) \geq \vee\{f_3(v) \mid R_v \subseteq X\}$ (resp. $f_3(u) \geq \vee\{f_3(v) \mid R_v \subseteq Y\}$). Therefore, this condition holds true for the child X (resp. Y) of R_u.*

2. *Let us assume that $f_3(u) = f_2(u)$. The same reasoning of (1) can be applied in this case.*

We can conclude that, for any building edge u of $\mathcal{B}$, we have $f_3(u) \geq \vee\{f_3(v) \mid R_v \subseteq R\}$ for a child R of R_u. □

Proof 16 (of Property 7). *By Lemma 29, we can affirm that $(V, E(\mathcal{B}))$ is a MST of both $(G, \Phi(\mathcal{H}_1))$ and $(G, \Phi(\mathcal{H}_2))$. Let f_3 denote the map $\curlywedge(\Phi(\mathcal{H}_1), \Phi(\mathcal{H}_2))$. By Lemma 33, $(V, E(\mathcal{B}))$ is a MST of (G, f_3) as well, which proves the first condition for $\mathcal{QFZ}(G, f_3)$ to be a flattened hierarchical watershed of (G, w). By Lemmas 36 and 37, the second and third conditions for $\mathcal{QFZ}(G, f_3)$ to be a flattened hierarchical watershed of (G, w) hold true. Therefore, $\mathcal{QFZ}(G, f_3)$ is a flattened hierarchical watershed of (G, w).* □

Appendix B.3. Proof of Property 8

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hierarchical watersheds of (G, w) and let $\mathcal{B}$ be a binary partition hierarchy of (G, w) such that both $\Phi(\mathcal{H}_1)$ and $\Phi(\mathcal{H}_2)$ are one-side increasing for $\mathcal{B}$. Let C be a positive function from $\mathbb{R}^2$ into $\mathbb{R}$ such that, for any a, b, c and d in $\{0, \ldots, n-1\}$, we have:

1. $C(0, 0) = 0$; and
2. $C(a, b) = C(b, a)$; and
3. if $min(a, b) = min(c, d)$ and $max(a, b) < max(c, d)$ then $C(a, b) \leq C(c, d)$; and
4. if $min(a, b) < min(c, d)$ then $C(a, b) < C(c, d)$.

Let f_3 denote the map $C(\Phi(\mathcal{H}_1), \Phi(\mathcal{H}_2))$. We want to prove that the hierarchy $\mathcal{QFZ}(G, f_3)$ is a flattened hierarchical watershed of (G, w). By Property 30, we need to prove that there exists a binary partition hierarchy $\mathcal{B}'$ of (G, w) such that the following statements hold true:

1. $(V, E(\mathcal{B}'))$ is a MST of (G, f_3); and
2. for any edge u in $E(\mathcal{B}')$, if u is not a watershed-cut edge for $\mathcal{B}'$, then $f_3(u) = 0$; and
3. for any edge u in $E(\mathcal{B}')$, there exists a child R of R_u such that $f_3(u) \geq \vee\{f_3(v)$ such that R_v is included in $R\}$, where $\vee\{\} = 0$.

The proof of this property follows the same idea of the proof of Property 7. To prove Property 8, we establish the following auxiliary lemma.

Lemma 38. *Let C be a function from $\mathbb{R}^2$ into $\mathbb{R}$ such that, for any two real values x and y, we have $C(x, y) = C(y, x)$. Let a, b, c and d be four real values. If $min(a, b) = min(c, d)$ and $max(a, b) = max(c, d)$, then $C(a, b) = C(c, d)$.*

Proof 17. *As $min(a, b) = min(c, d)$ and $max(a, b) = max(c, d)$, then either we have (i) $a = c$ and $b = d$ which implies that $C(a, b) = C(c, d)$; or (ii) $c = b$ and $d = a$ which implies that $C(c, d) = C(b, a)$, which, by our hypothesis on C, is equal to $C(a, b)$. Hence, we have $C(a, b) = C(c, d)$.* □

The following three lemmas prove that the conditions 1, 2 and 3 for $\mathcal{QFZ}(G, f_3)$ to be a flattened hierarchical watershed of (G, w) hold true.

Lemma 39. *Let C be a positive function such that, for any a, b, c and d in $\{0, \ldots, n-1\}$, we have:*

1. *$C(0,0) = 0$; and*
2. *$C(a,b) = C(b,a)$; and*
3. *if $min(a,b) = min(c,d)$ and $max(a,b) < max(c,d)$ then $C(a,b) \leq C(c,d)$; and*
4. *if $min(a,b) < min(c,d)$ then $C(a,b) < C(c,d)$.*

Let f_1 and f_2 be the saliency maps of two hierarchies on V and let G' be a subgraph of G such that G' is a MST of both (G, f_1) and (G, f_2). Then G' is also a MST of $(G, C(f_1, f_2))$.

Proof 18. *Let u be an edge in $E \setminus E(G')$. Let f_3 denote the map $C(f_1, f_2)$. Since G' is a spanning tree, by Lemma 34, the graph $(V, E(G') \cup \{u\})$ contains a cycle π which includes the edge u. Let π be the cycle of $(V, E(G') \cup \{u\})$ which includes the edge u. As G' is a MST of (G, f_1) and of (G, f_2), by Lemma 35, for any edge v in the cycle π, we have $f_1(v) \leq f_1(u)$ and $f_2(v) \leq f_2(u)$. Therefore, for any edge v in the cycle π, we have $min(f_1(v), f_2(v)) \leq min(f_1(u), f_2(u))$ and $max(f_1(v), f_2(v)) \leq max(f_1(u), f_2(u))$. Then, we should consider the three following cases:*

1. *If $min(f_1(v), f_2(v)) < min(f_1(u), f_2(u))$, then, by the hypothesis 4 on C, we have $C(f_1(v), f_2(v)) < C(f_1(u), f_2(u))$.*
2. *If $min(f_1(v), f_2(v)) = min(f_1(u), f_2(u))$ and $max(f_1(v), f_2(v)) < max(f_1(u), f_2(u))$, then, by the hypothesis 3 on C, we have $C(f_1(v), f_2(v)) \leq C(f_1(u), f_2(u))$.*
3. *If $min(f_1(v), f_2(v)) = min(f_1(u), f_2(u))$ and $max(f_1(v), f_2(v)) = max(f_1(u), f_2(u))$, then, by Lemma 38, we have $C(f_1(v), f_2(v)) = C(f_1(u), f_2(u))$.*

Consequently, $C(f_1(v), f_2(v)) = f_3(v) \leq C(f_1(u), f_2(u)) = f_3(u)$. Hence, for any edge v in the cycle π, we have $f_3(v) \leq f_3(u)$. Thus, by Lemma 35, G' is a MST of (G, f_3). □

Lemma 40. *Let C be a positive function such that, for any a, b, c and d in $\{0, \ldots, n-1\}$, we have:*

1. *$C(0,0) = 0$; and*
2. *$C(a,b) = C(b,a)$; and*
3. *if $min(a,b) = min(c,d)$ and $max(a,b) < max(c,d)$ then $C(a,b) \leq C(c,d)$; and*
4. *if $min(a,b) < min(c,d)$ then $C(a,b) < C(c,d)$.*

Let f_1 and f_2 be the saliency maps of two hierarchies on V and let $\mathcal{B}$ be a binary partition hierarchy of (G, w) such that both f_1 and f_2 are one-side increasing for $\mathcal{B}$. Then for any u in $E(\mathcal{B})$, if u is not a watershed-cut edge for $\mathcal{B}$, then $C(f_1, f_2)(u) = 0$.

Proof 19. *Let u be an edge in $E(\mathcal{B})$. If u is not an watershed-cut edge for $\mathcal{B}$, then, by the second condition of Definition 4, we have $f_1(u) = 0$ and $f_2(u) = 0$. Therefore, $C(f_1, f_2)(u) = C(0,0) = 0$.* □

Lemma 41. *Let C be a positive function such that, for any a, b, c and d in $\{0, \ldots, n-1\}$, we have:*

1. *$C(0,0) = 0$; and*
2. *$C(a,b) = C(b,a)$; and*
3. *if $min(a,b) = min(c,d)$ and $max(a,b) < max(c,d)$ then $C(a,b) \leq C(c,d)$; and*
4. *if $min(a,b) < min(c,d)$ then $C(a,b) < C(c,d)$.*

Let f_1 and f_2 be the saliency maps of two hierarchies on V and let $\mathcal{B}$ be a binary partition hierarchy of (G, w) such that both f_1 and f_2 are one-side increasing for $\mathcal{B}$. Let f_3 denote the map $C(f_1, f_2)$. Then, for any building edge u of $\mathcal{B}$, there exists a child R of R_u such that $f_3(u) \geq \vee\{f_3(v)$ such that R_v is included in $R\}$.

Proof 20. *Since f_1 (resp. f_2) is one-side increasing for $\mathcal{B}$, by the third condition of Definition 4, we have that, for any building edge u of $\mathcal{B}$, $f_1(u) \geq \vee\{f_1(v) \mid R_v \subseteq X\}$ (resp. $f_2(u) \geq \vee\{f_2(v) \mid R_v \subseteq X\}$) for a child X of R_u. We need to prove that, for any building edge u of $\mathcal{B}$, there is a child X of R_u such that $f_3(u) \geq \vee\{f_3(v) \mid R_v \subseteq X\}$. Let u be a building edge of $\mathcal{B}$ and let X and Y be the children of R_u. We should consider the following four cases:*

1. *If* $f_1(u) \geq \vee\{f_1(v) \mid R_v \subseteq X\}$ *and* $f_2(u) \geq \vee\{f_2(v) \mid R_v \subseteq X\}$*, then, for any building edge e such that* $R_e \subseteq X$*, we have* $f_1(u) \geq f_1(e)$ *and* $f_2(u) \geq f_2(e)$*. Let e be an edge edge such that* $R_e \subseteq X$*. Therefore,* $min(f_1(e), f_2(e)) \leq min(f_1(u), f_2(u))$*. If* $min(f_1(e), f_2(e)) < min(f_1(u), f_2(u))$ *then, by the hypothesis 4 on C,* $C(f_1(e), f_2(e)) < C(f_1(u), f_2(u))$*. Otherwise, we have* $min(f_1(e), f_2(e)) = min(f_1(u), f_2(u))$*. As* $f_1(u) \geq f_1(v)$ *and* $f_2(u) \geq f_2(e)$*, we have* $\max(f_1(u), f_2(u)) \geq \max(f_1(e), f_2(e))$*. If* $\max(f_1(u), f_2(u)) = \max(f_1(e), f_2(e))$ *then, by Lemma 38,* $C(f_1(u), f_2(u)) = C(f_1(e), f_2(e))$*. Otherwise, we have* $\max(f_1(u), f_2(u)) > \max(f_1(e), f_2(e))$ *and then by hypothesis 3 on C, we have* $C(f_1(u), f_2(u)) \geq C(f_1(e), f_2(e))$*. Thus in all cases we have* $C(f_1(u), f_2(u)) \geq C(f_1(v), f_2(v))$ *, and by definition of* f_3*:* $f_3(u) \geq f_3(e)$*. Therefore,* $f_3(u) \geq \vee\{f_3(v) \mid R_v \subseteq X\}$*.*

2. *If* $f_1(u) \geq \vee\{f_1(v) \mid R_v \subseteq X\}$ *and* $f_2(u) \geq \vee\{f_2(v) \mid R_v \subseteq Y\}$*, then we have to consider two cases: (i)* $f_1(u) \leq f_2(u)$ *and (ii)* $f_1(u) > f_2(u)$*.*

   (i) *Assume that* $f_1(u) \leq f_2(u)$*. Then* $min(f_1(u), f_2(u)) = f_1(u)$*. Let v be an edge such that* $R_v \subseteq X$*. By our assumption, we have* $f_1(u) \geq f_1(v)$*. Indeed, since f is a one-side increasing map, we can say that either* $f_1(u) = f_1(v) = 0$ *or* $f_1(u) > f_1(v)$ *because only the watershed-cut edges for* $\mathcal{B}$ *have non-zero and pairwise distinct weights. If* $f_1(u) = f_1(v) = 0$*, this implies that neither u nor v are watershed-cut edges for* $\mathcal{B}$ *and therefore* $f_2(u) = f_2(v) = 0$*, which implies that* $f_3(u) = 0 \geq f_3(v) = 0$*. Otherwise, let us assume that* $f_1(u) > f_1(v)$*. In this case, and as* $min(f_1(u), f_2(u)) = f_1(u)$*, we have* $min(f_1(u), f_2(u)) > f_1(v)$*, and thus* $min(f_1(u), f_2(u)) > min(f_1(v), f_2(v))$*. Then by hypothesis 4 on C, we have* $C(f_1(u), f_2(u)) > C(f_1(v), f_2(v)$ *which is equivalent to* $f_3(u) > f_3(v)$*. Therefore, we have* $f_3(u) \geq \vee\{f_3(v) \mid R_v \subseteq X\}$*.*

   (ii) *If* $f_1(u) > f_2(u)$ *then we can apply the same reasoning as in the case where* $f_1(u) \leq f_2(u)$*.*

3. $f_1(u) \geq \vee\{f_1(v) \mid R_v \subseteq Y\}$ *and* $f_2(u) \geq \vee\{f_2(v) \mid R_v \subseteq X\}$*. This case is symmetric to 2.*

4. $f_1(u) \geq \vee\{f_1(v) \mid R_v \subseteq Y\}$ *and* $f_2(u) \geq \vee\{f_2(v) \mid R_v \subseteq Y\}$*. This case is symmetric to 1.*

Thus, we can conclude that, for any building edge u of $\mathcal{B}$*, there exists a child R of* R_u *such that* $f_3(u) \geq \vee\{f_3(v)$ *such that* R_v *is included in R*$\}$*.* □

Proof 21 (of Property 8). *By Lemma 29, we can affirm that* $(V, E(\mathcal{B}))$ *is a MST of both* $(G, \Phi(\mathcal{H}_1))$ *and* $(G, \Phi(\mathcal{H}_2))$*. Therefore, by Lemma 39,* $(V, E(\mathcal{B}))$ *is a MST of* (G, f_3) *as well, which proves that the first condition for* $\mathcal{QFZ}(G, f_3)$ *to be a flattened hierarchical watershed of* (G, w) *holds true. The second and third conditions are the result of Lemmas 40 and 41, respectively. Therefore,* $\mathcal{QFZ}(G, f_3)$ *is a flattened hierarchical watershed of* (G, w)*.* □

Appendix B.4. Proof of Property 9

Let C be the function:

$$C(x, y) = \begin{cases} 0 & if \text{ x=0 } and \text{ y=0} \\ \frac{x^m y^m}{x^m + y^m} & \end{cases} \tag{B.2}$$

where m is equal or greater than the number of minima n of (G, w). We want to prove that, for any a, b, c and d in $\{0, \ldots, n-1\}$:

1. $C(0, 0) = 0$; and
2. $C(a, b) = C(b, a)$; and
3. if $min(a, b) = min(c, d)$ and $max(a, b) < max(c, d)$ then $C(a, b) \leq C(c, d)$; and
4. if $min(a, b) < min(c, d)$ then $C(a, b) < C(c, d)$.

Since $m \geq n$, we can prove that those fours statements hold true for any a, b, c and d in $\{0, \ldots, m-1\}$.

The proof of the first and second statements are trivial. In order to prove the third and fourth statements, we state Lemmas 42 and 43.

Lemma 42. *Let* $C(x, y) = \frac{x^m y^m}{x^m + y^m}$ *and let a, b and d be natural numbers such that* $a \leq b$*,* $a \leq d$ *and* $b < d$*. Then* $C(a, b) \leq C(a, d)$*.*

Proof 22 (of Lemma 42). *If $a = 0$, then $C(a,b) = 0$ which is less than or equal to $C(a,d) = 0$. Otherwise, let us assume that $a > 0$. We will prove that $C(a,b) \leq C(a,d)$ by proving that $C(a,d) - C(a,b)$ is positive.*

$$C(a,d) - C(a,b) \tag{B.3}$$
$$= \frac{a^m d^m}{a^m + d^m} - \frac{a^m b^m}{a^m + b^m} \tag{B.4}$$
$$= \frac{a^{2m} d^m + \cancel{a^m b^m d^m} - a^{2m} b^m - \cancel{a^m b^m d^m}}{(a^m + d^m)(a^m + b^m)} \tag{B.5}$$
$$= \frac{a^{2m} d^m - a^{2m} b^m}{(a^m + d^m)(a^m + b^m)} \tag{B.6}$$
$$= \frac{a^{2m}(d^m - b^m)}{(a^m + d^m)(a^m + b^m)} \tag{B.7}$$

The denominator of the fraction (A.6) is clearly positive and, since $d > b$, we can say that $d^m - b^m$ is positive as well. Therefore, $C(a,b) - C(c,d)$ is positive and, consequently, $C(a,b) \leq C(c,d)$. □

Lemma 43. *Let $C(x,y) = \frac{x^m y^m}{x^m + y^m}$ and let a, b, c and d be natural numbers in $\{0, \ldots, m-1\}$ such that $a \leq b$ and $c \leq d$. If $a < c$ then $C(a,b) < C(c,d)$.*

Proof 23. *Let us define the function $f_a(y) = \frac{a^m y^m}{a^m + y^m}$ where y is a natural number. We will compute the limit of $f_a(y)$ for y tending to infinity in order to find the greatest value $C(a,y)$ for any y.*

$$\lim_{y \to \infty} \frac{a^m y^m}{a^m + y^m} \tag{B.8}$$
$$= \lim_{y \to \infty} \frac{\frac{a^m \cancel{y^m}}{\cancel{y^m}}}{\frac{a^m + y^m}{y^m}} \tag{B.9}$$
$$= \lim_{y \to \infty} \frac{a^m}{\frac{a^m}{y^m} + \frac{y^m}{y^m}} \tag{B.10}$$
$$= \frac{a^m}{0 + 1} \tag{B.11}$$
$$= a^m \tag{B.12}$$

Let c be a value in $\{0, \ldots, m-1\}$ such that $a < c$. We will prove that $C(c,d)$ is greater than $\lim_{y \to \infty} \frac{a^m y^m}{a^m + y^m} = a^m$. Since $a < c$, we have $c \geq a+1$. If we prove that this lemma holds for the case where $c = a + 1$, we can infer by recurrence that it holds for any c greater than a. Therefore, we can simply prove that $C(a+1,d)$ is greater than a^m for any $d > a$ (because $d \geq c$ by hypothesis). By Lemma 42, given any value d' such that $a + 1 \leq d'$, we have that $C(a+1, a+1) \leq C(a+1, d')$. Since $a < d$, the minimal value of d is $a + 1$. Given that $d = a + 1$, we have $C(a+1, d) = C(a+1, a+1) = \frac{(a+1)^{2m}}{2(a+1)^m} = \frac{(a+1)^m}{2}$. Then we only need to prove that $a^m < \frac{(a+1)^m}{2}$ or that $a^m - \frac{(a+1)^m}{2} < 0$.

$$a^m - \frac{(a+1)^m}{2} \tag{B.13}$$
$$= a^m - \left(\frac{a+1}{\sqrt[m]{2}}\right)^m \tag{B.14}$$
$$= \left(a - \frac{a+1}{\sqrt[m]{2}}\right)\left(a^{m-1} + a^{m-2}\left(\frac{a+1}{\sqrt[m]{2}}\right) + \cdots + a\left(\frac{a+1}{\sqrt[m]{2}}\right)^{m-2} + \left(\frac{a+1}{\sqrt[m]{2}}\right)^{m-1}\right) \tag{B.15}$$

The equation (A.15) is obtained by the factorization of the equation (A.14). The sign of equation (A.15) is determined by the first term $\left(a - \frac{a+1}{\sqrt[m]{2}}\right)$ because the other terms are positive since a and m are natural numbers. Thus, in order to prove that (A.13) is negative, we only need to show that $\left(a - \frac{a+1}{\sqrt[m]{2}}\right) < 0$.

$$\left(a - \frac{a+1}{\sqrt[m]{2}}\right) < 0 \tag{B.16}$$
$$\sqrt[m]{2}a - a - 1 < 0 \tag{B.17}$$
$$a(\sqrt[m]{2} - 1) < 1 \tag{B.18}$$
$$a < \frac{1}{\sqrt[m]{2} - 1} \tag{B.19}$$

Hence, we need to demonstrate that $a < \frac{1}{\sqrt[m]{2}-1}$. Since a is in $\{0, \ldots, m-1\}$, we know that $a < m$ and we can simply prove that $m \leq \frac{1}{\sqrt[m]{2}-1}$ or that $\sqrt[m]{2}m - m \leq 1$ or $m - \sqrt[m]{2}m \geq -1$. Let us define the real function h as follows:

$$h(x) = x - \sqrt[x]{2}x \tag{B.20}$$
$$= x(1 - \sqrt[x]{2}) \tag{B.21}$$
$$= x(1 - e^{\frac{\ln 2}{x}}) \tag{B.22}$$

We will show that $h(x) \geq -1$ for any x in $[1, +\infty[$. If this holds true in the continuous case, we can infer that it also holds true in the discrete case. Given that $h(1) = -1$, we can prove that $h(x) \geq -1$ for any x in $[1, +\infty[$ by showing that $h(x)$ is increasing in the interval $[1, +\infty[$. To that end, we will verify that the derivative of $h(x)$ is positive for any x in $[1, +\infty[$.

$$h'(x) = 1 - e^{\frac{\ln 2}{x}} - x\left(e^{\frac{\ln 2}{x}}\left(-\frac{\ln 2}{x^2}\right)\right) \tag{B.23}$$

$$= 1 - e^{\frac{\ln 2}{x}} + e^{\frac{\ln 2}{x}}\frac{\ln 2}{x} \tag{B.24}$$

To verify that $h'(x)$ is positive in the interval $[1, +\infty[$, we compute the limite of $h'(x)$ when x goes to $+\infty$ and its derivative $h''(x)$ of $h'(x)$.

$$\lim_{x \to +\infty} h'(x) = 1 - e^{\frac{\ln 2}{+\infty}} + e^{\frac{\ln 2}{+\infty}}\frac{\ln 2}{+\infty} \tag{B.25}$$

$$= 1 - e^0 + e^0 \times 0 \tag{B.26}$$

$$= 0 \tag{B.27}$$

$$h''(x) = -\cancel{\left(e^{\frac{\ln 2}{x}}\left(-\frac{\ln 2}{x^2}\right)\right)} + \left(-\frac{\ln 2}{x^2}\right)e^{\frac{\ln 2}{x}}\frac{\ln 2}{x} + \cancel{\left(-\frac{\ln 2}{x^2}\right)e^{\frac{\ln 2}{x}}} \tag{B.28}$$

$$= -\frac{(\ln 2)^2}{x^3}e^{\frac{\ln 2}{x}} \tag{B.29}$$

Therefore, we can affirm that $h''(x)$ is negative for any x in $[1, +\infty[$, which implies that $h'(x)$ is decreasing in the interval $[1, +\infty[$. Since $h'(x)$ is decreasing and the limit of $h'(x)$ going to infinity is zero, we can say that $h'(x)$ is positive for any x in $[1, +\infty[$. In addition, as $h(1) = -1$, this implies that $h(x)$ is increasing in the interval $[1, +\infty[$. This implies that $h(x) \geq -1$ and, therefore, $m - \sqrt[m]{2}m \geq -1$. This completes the proof that $C(a, b) < C(c, d)$. □