

## A. Supplementary Material

### A.1. Proof of Theorem 1

In order to show that  $\Psi$  is a well-defined map, we must show that  $\Psi(\tilde{D}_X)$  is a quasi-ultrametric network for every quasi-dendrogram  $\tilde{D}_X$ . Given an arbitrary quasi-dendrogram  $\tilde{D}_X = (D_X, E_X)$ , for a particular  $\delta' \geq 0$  consider the quasi-partition  $\tilde{D}_X(\delta')$ . Consider the range of resolutions  $\delta$  associated with such quasi-partition. I.e.,

$$\{\delta \geq 0 \mid \tilde{D}_X(\delta) = \tilde{D}_X(\delta')\}. \quad (17)$$

Right continuity ( $\tilde{D}4$ ) of  $\tilde{D}_X$  ensures that the minimum of the set in (17) is well-defined and hence definition (8) is valid. To prove that  $\tilde{u}_X$  in (8) is a quasi-ultrametric we need to show that it attains non-negative values as well as the identity and strong triangle inequality properties. That  $\tilde{u}_X$  attains non-negative values is clear from the definition (8). The identity property is implied by the first boundary condition in ( $\tilde{D}1$ ). Since  $[x]_0 = [x]_0$  for all  $x \in X$ , we must have  $\tilde{u}_X(x, x) = 0$ . Conversely, since for all  $x \neq x' \in X$ ,  $([x]_0, [x']_0) \notin E_X(0)$  and  $[x]_0 \neq [x']_0$  we must have that  $\tilde{u}_X(x, x') > 0$  for  $x \neq x'$  and the identity property is satisfied. To see that  $\tilde{u}_X$  satisfies the strong triangle inequality in (7), consider nodes  $x, x'$ , and  $x''$  such that the lowest resolution for which  $[x]_\delta = [x']_\delta$  or  $([x]_\delta, [x'']_\delta) \in E_X(\delta)$  is  $\delta_1$  and the lowest resolution for which  $[x'']_\delta = [x']_\delta$  or  $([x'']_\delta, [x']_\delta) \in E_X(\delta)$  is  $\delta_2$ . Right continuity ( $\tilde{D}4$ ) ensures that these lowest resolutions are well-defined. According to (8) we then have

$$\begin{aligned} \tilde{u}_X(x, x'') &= \delta_1, \\ \tilde{u}_X(x'', x') &= \delta_2. \end{aligned} \quad (18)$$

Denote by  $\delta_0 := \max(\delta_1, \delta_2)$ . From the equivalence hierarchy ( $\tilde{D}2$ ) and influence hierarchy ( $\tilde{D}3$ ) properties, it follows that  $[x]_{\delta_0} = [x'']_{\delta_0}$  or  $([x]_{\delta_0}, [x'']_{\delta_0}) \in E_X(\delta_0)$  and  $[x'']_{\delta_0} = [x']_{\delta_0}$  or  $([x'']_{\delta_0}, [x']_{\delta_0}) \in E_X(\delta_0)$ . Furthermore, from transitivity (QP2) of the quasi-partition  $\tilde{D}_X(\delta_0)$ , it follows that  $[x]_{\delta_0} = [x']_{\delta_0}$  or  $([x]_{\delta_0}, [x']_{\delta_0}) \in E_X(\delta_0)$ . Using the definition in (8) for  $x, x'$  we conclude that

$$\tilde{u}_X(x, x') \leq \delta_0. \quad (19)$$

By definition  $\delta_0 := \max(\delta_1, \delta_2)$ , hence we substitute this expression in (19) and compare with (18) to obtain

$$\tilde{u}_X(x, x') \leq \max(\delta_1, \delta_2) = \max(\tilde{u}_X(x, x''), \tilde{u}_X(x'', x')). \quad (20)$$

Consequently,  $\tilde{u}_X$  satisfies the strong triangle inequality and is therefore a quasi-ultrametric, proving that the map  $\Psi$  is well-defined.

For the converse result, we need to show that  $\Upsilon$  is a well-defined map. Given a quasi-ultrametric  $\tilde{u}_X$  on a node set

$X$  and a resolution  $\delta \geq 0$ , we first define the relation

$$x \rightsquigarrow_{\tilde{u}_X(\delta)} x' \iff \tilde{u}_X(x, x') \leq \delta, \quad (21)$$

for all  $x, x' \in X$ . Notice that  $\rightsquigarrow_{\tilde{u}_X(\delta)}$  is a quasi-equivalence relation as defined in Definition 1 for all  $\delta \geq 0$ . The reflexivity property is implied by the identity property of the quasi-ultrametric  $\tilde{u}_X$  and transitivity is implied by the fact that  $\tilde{u}_X$  satisfies the strong triangle inequality. Furthermore, definitions (9) and (10) are just reformulations of (2) and (3) respectively, for the special case of the quasi-equivalence defined in (21). Hence, Proposition 1 guarantees that  $\Upsilon(X, \tilde{u}_X) = \tilde{D}_X(\delta) = (D_X(\delta), E_X(\delta))$  is a quasi-partition for every resolution  $\delta \geq 0$ . In order to show that  $\Upsilon$  is well-defined, we need to show that these quasi-partitions are nested, i.e. that  $\tilde{D}_X$  satisfies ( $\tilde{D}1$ )-( $\tilde{D}4$ ).

The first boundary condition in ( $\tilde{D}1$ ) is implied by (9) and the identity property of  $\tilde{u}_X$ . The second boundary condition in ( $\tilde{D}1$ ) is implied by the fact that  $\tilde{u}_X$  takes finite real values on a finite domain since the node set  $X$  is finite. Hence, any  $\delta_0$  satisfying

$$\delta_0 \geq \max_{x, x' \in X} \tilde{u}_X(x, x'), \quad (22)$$

is a valid candidate to show fulfillment of ( $\tilde{D}1$ ).

To see that  $\tilde{D}_X$  satisfies ( $\tilde{D}2$ ) assume that for a resolution  $\delta_1$  we have two nodes  $x, x' \in X$  such that  $x \sim_{\tilde{u}_X(\delta_1)} x'$  as in (9), then it follows that  $\max(\tilde{u}_X(x, x'), \tilde{u}_X(x', x)) \leq \delta_1$ . Thus, if we pick any  $\delta_2 > \delta_1$  it is immediate that  $\max(\tilde{u}_X(x, x'), \tilde{u}_X(x', x)) \leq \delta_2$  which by (9) implies that  $x \sim_{\tilde{u}_X(\delta_2)} x'$ .

Fulfillment of ( $\tilde{D}3$ ) can be shown in a similar way as fulfillment of ( $\tilde{D}2$ ). Given a scalar  $\delta_1 \geq 0$  and  $x, x' \in X$ , if  $([x]_{\delta_1}, [x']_{\delta_1}) \in E_X(\delta_1)$  then by (10) we have that

$$\min_{x_1 \in [x]_{\delta_1}, x_2 \in [x']_{\delta_1}} \tilde{u}_X(x_1, x_2) \leq \delta_1. \quad (23)$$

From property ( $\tilde{D}2$ ), we know that for all  $x \in X$ ,  $[x]_{\delta_1} \subset [x]_{\delta_2}$  for all  $\delta_2 > \delta_1$ . Hence, two things might happen. Either  $\max(\tilde{u}_X(x, x'), \tilde{u}_X(x', x)) \leq \delta_2$  in which case  $[x]_{\delta_2} = [x']_{\delta_2}$  or it might be that  $[x]_{\delta_2} \neq [x']_{\delta_2}$  but

$$\min_{x_1 \in [x]_{\delta_2}, x_2 \in [x']_{\delta_2}} \tilde{u}_X(x_1, x_2) \leq \delta_1 < \delta_2, \quad (24)$$

which implies that  $([x]_{\delta_2}, [x']_{\delta_2}) \in E_X(\delta_2)$ . Consequently, ( $\tilde{D}3$ ) is satisfied.

Finally, to see that  $\tilde{D}_X$  satisfies the right continuity condition ( $\tilde{D}4$ ), for each  $\delta \geq 0$  such that  $\tilde{D}_X(\delta) \neq (\{X\}, \emptyset)$  we may define  $\epsilon(\delta)$  as any positive scalar satisfying

$$0 < \epsilon(\delta) < \min_{\substack{x, x' \in X \\ \text{s.t. } \tilde{u}_X(x, x') > \delta}} \tilde{u}_X(x, x') - \delta, \quad (25)$$

where the finiteness of  $X$  ensures that  $\epsilon(\delta)$  is well-defined. Hence, (9) and (10) guarantee that  $\tilde{D}_X(\delta) = \tilde{D}_X(\delta')$  for  $\delta' \in [\delta, \delta + \epsilon(\delta)]$ . For all other resolutions  $\delta$  such that  $\tilde{D}_X(\delta) = (\{X\}, \emptyset)$ , right continuity is trivially satisfied since the quasi-dendrogram remains unchanged for increasing resolutions. Consequently,  $\Upsilon(X, \tilde{u}_X)$  is a valid quasi-dendrogram for every quasi-ultrametric network  $(X, \tilde{u}_X)$ , proving that  $\Upsilon$  is well-defined.

In order to conclude the proof, we need to show that  $\Psi \circ \Upsilon$  and  $\Upsilon \circ \Psi$  are the identities on  $\tilde{U}$  and  $\tilde{D}$ , respectively. To see why the former is true, pick any quasi-ultrametric network  $(X, \tilde{u}_X)$  and consider an arbitrary pair of nodes  $x, x' \in X$  such that  $\tilde{u}_X(x, x') = \delta_0$ . Also, consider the ultrametric network  $\Psi \circ \Upsilon(X, \tilde{u}_X) := (X, \tilde{u}_X^*)$ . From (9) and (10), in the quasi-dendrogram  $\Upsilon(X, \tilde{u}_X)$  there is no influence from  $x$  to  $x'$  for resolutions  $\delta < \delta_0$  and at resolution  $\delta = \delta_0$  either an edge appears from  $[x]_{\delta_0}$  to  $[x']_{\delta_0}$ , or both nodes merge into one single cluster. In any case, when we apply  $\Psi$  to the resulting quasi-dendrogram, we obtain  $\tilde{u}_X^*(x, x') = \delta_0$ . Since  $x, x' \in X$  were chosen arbitrarily, we have that  $\tilde{u}_X = \tilde{u}_X^*$ , showing that  $\Psi \circ \Upsilon$  is the identity on  $\tilde{U}$ . A similar argument shows that  $\Upsilon \circ \Psi$  is the identity on  $\tilde{D}$ .

## A.2. Proof of Proposition 2

For this proof, we introduce the concept of chain concatenation. Given two chains  $C(x, x') = [x = x_0, x_1, \dots, x_l = x']$  and  $C(x', x'') = [x' = x'_0, x'_1, \dots, x'_{l'} = x'']$  such that the end point  $x'$  of the first one coincides with the starting point of the second one, define the *concatenated chain*  $C(x, x') \uplus C(x', x'')$  as

$$\begin{aligned} C(x, x') \uplus C(x', x'') \\ := [x = x_0, \dots, x_l = x' = x'_0, \dots, x'_{l'} = x'']. \end{aligned} \quad (26)$$

For the method  $\tilde{\mathcal{H}}^*$  to be a properly defined hierarchical quasi-clustering method, we need to establish that  $\tilde{u}_X^*$  is a valid ultrametric. To see that  $\tilde{u}_X^*(x, x') = 0$  if and only if  $x = x'$ , notice that when  $x = x'$ , the chain  $C(x, x) = [x, x]$  has null cost and when  $x \neq x'$  any chain must contain at least one link with strictly positive cost. To verify that the strong triangle inequality in (7) holds, let  $C^*(x, x'')$  and  $C^*(x'', x')$  be chains that achieve the minimum in (13) for  $\tilde{u}_X^*(x, x'')$  and  $\tilde{u}_X^*(x'', x')$ , respectively. The maximum cost in the concatenated chain  $C(x, x') = C^*(x, x'') \uplus C^*(x'', x')$  does not exceed the maximum cost in each individual chain. Thus, while the maximum cost may be smaller on a different chain, the chain  $C(x, x')$  suffices to bound  $\tilde{u}_X^*(x, x') \leq \max(\tilde{u}_X^*(x, x''), \tilde{u}_X^*(x'', x'))$  as in (7).

To show fulfillment of Axiom ( $\tilde{A}1$ ), pick an arbitrary two-node network  $\tilde{\Delta}_2(\alpha, \beta) := (\{p, q\}, A_{p,q})$  with

$A_{p,q}(p, q) = \alpha$  and  $A_{p,q}(q, p) = \beta$  for some  $\alpha, \beta > 0$  and denote by  $(\{p, q\}, \tilde{u}_{p,q}^*) = \tilde{\mathcal{H}}^*(\tilde{\Delta}_2(\alpha, \beta))$ . Then, we have  $\tilde{u}_{p,q}^*(p, q) = \alpha$  and  $\tilde{u}_{p,q}^*(q, p) = \beta$  because there is only one possible chain selection in each direction [cf. (13)]. To prove that Axiom ( $\tilde{A}2$ ) is satisfied consider arbitrary points  $x, x' \in X$  and denote by  $C^*(x, x')$  one chain achieving the minimum chain cost in (13),

$$\tilde{u}_X^*(x, x') = \max_{i|x_i \in C^*(x, x')} A(x_i, x_{i+1}). \quad (27)$$

Consider the transformed chain  $C_Y(\phi(x), \phi(x')) = [\phi(x) = \phi(x_0), \dots, \phi(x_l) = \phi(x')]$  in the space  $Y$ . Since the map  $\phi : X \rightarrow Y$  reduces dissimilarities we have that for all links in this chain  $A_Y(\phi(x_i), \phi(x_{i+1})) \leq A_X(x_i, x_{i+1})$ . Consequently,

$$\begin{aligned} \max_{i|\phi(x_i) \in C_Y(\phi(x), \phi(x'))} A_Y(\phi(x_i), \phi(x_{i+1})) \\ \leq \max_{i|x_i \in C^*(x, x')} A_X(x_i, x_{i+1}). \end{aligned} \quad (28)$$

Further note that the minimum chain cost  $\tilde{u}_Y^*(\phi(x), \phi(x'))$  among all chains linking  $\phi(x)$  to  $\phi(x')$  cannot exceed the cost in the given chain  $C_Y(\phi(x), \phi(x'))$ . Combining this observation with the inequality in (28) it follows that

$$\tilde{u}_Y^*(\phi(x), \phi(x')) \leq \max_{i|x_i \in C^*(x, x')} A_X(x_i, x_{i+1}) = \tilde{u}_X^*(x, x'), \quad (29)$$

where we also used (27) to write the equality. Expression (29) ensures fulfillment of Axiom ( $\tilde{A}2$ ), as wanted.

## A.3. Proof of Theorem 2

In proving this theorem, the concept of *separation* of a network is useful. Given an arbitrary network  $(X, A_X)$ , its separation  $\text{sep}(X, A_X)$  is defined as the minimum positive dissimilarity in the network, that is

$$\text{sep}(X, A_X) = \min_{x \neq x'} A_X(x, x'). \quad (30)$$

The following auxiliary result is useful in showing Theorem 2.

**Lemma 1** *A network  $N = (X, A_X)$  and a positive constant  $\delta$  are given. Then, for any pair of nodes  $x, x' \in X$  whose minimum chain cost [cf. (13)] satisfies*

$$\tilde{u}_X^*(x, x') \geq \delta, \quad (31)$$

*there exists a partition  $P_\delta(x, x') = \{B_\delta(x), B_\delta(x')\}$  of the node space  $X$  into blocks  $B_\delta(x)$  and  $B_\delta(x')$  with  $x \in B_\delta(x)$  and  $x' \in B_\delta(x')$  such that for all points  $b \in B_\delta(x)$  and  $b' \in B_\delta(x')$*

$$A_X(b, b') \geq \delta. \quad (32)$$

**Proof:** We prove this result by contradiction. If a partition  $P_\delta(x, x') = \{B_\delta(x), B_\delta(x')\}$  with  $x \in B_\delta(x)$  and  $x' \in B_\delta(x')$  and satisfying (32) does not exist for all pairs of points  $x, x' \in X$  satisfying (31), then there is at least one pair of nodes  $x, x' \in X$  satisfying (31) such that for *all* partitions of  $X$  into two blocks  $P = \{B, B'\}$  with  $x \in B$  and  $x' \in B'$  we can find at least a pair of elements  $b_P \in B$  and  $b'_P \in B'$  for which

$$A_X(b_P, b'_P) < \delta. \quad (33)$$

Begin by considering the partition  $P_1 = \{B_1, B'_1\}$  where  $B_1 = \{x\}$  and  $B'_1 = X \setminus \{x\}$ . Since (33) is true for all partitions having  $x \in B$  and  $x' \in B'$  and  $x$  is the unique element of  $B_1$ , there must exist a node  $b'_{P_1} \in B'_1$  such that

$$A_X(x, b'_{P_1}) < \delta. \quad (34)$$

Hence, the chain  $C(x, b'_{P_1}) = [x, b'_{P_1}]$  composed of these two nodes has cost smaller than  $\delta$ . Moreover, since  $\tilde{u}_X^*(x, b'_{P_1})$  represents the minimum cost among all chains  $C(x, b'_{P_1})$  linking  $x$  to  $b'_{P_1}$ , we can assert that

$$\tilde{u}_X^*(x, b'_{P_1}) \leq A_X(x, b'_{P_1}) < \delta. \quad (35)$$

Consider now the partition  $P_2 = \{B_2, B'_2\}$  where  $B_2 = \{x, b'_{P_1}\}$  and  $B'_2 = X \setminus B_2$ . From (33), there must exist a node  $b'_{P_2} \in B'_2$  that satisfies at least one of the two following conditions

$$A_X(x, b'_{P_2}) < \delta, \quad (36)$$

$$A_X(b'_{P_1}, b'_{P_2}) < \delta. \quad (37)$$

If (36) is true, the chain  $C(x, b'_{P_2}) = [x, b'_{P_2}]$  has cost smaller than  $\delta$ . If (37) is true, we combine the dissimilarity bound with the one in (34) to conclude that the chain  $C(x, b'_{P_2}) = [x, b'_{P_1}, b'_{P_2}]$  has cost smaller than  $\delta$ . In either case we conclude that there exists a chain  $C(x, b'_{P_2})$  linking  $x$  to  $b'_{P_2}$  whose cost is smaller than  $\delta$ . Therefore, the minimum chain cost must satisfy

$$\tilde{u}_X^*(x, b'_{P_2}) < \delta. \quad (38)$$

Repeat the process by considering the partition  $P_3$  with  $B_3 = \{x, b'_{P_1}, b'_{P_2}\}$  and  $B'_3 = X \setminus B_3$ . As we did in arguing (36)-(37) it must follow from (33) that there exists a point  $b'_{P_3}$  such that at least one of the dissimilarities  $A_X(x, b'_{P_3})$ ,  $A_X(b'_{P_1}, b'_{P_3})$ , or  $A_X(b'_{P_2}, b'_{P_3})$  is smaller than  $\delta$ . This observation implies that at least one of the chains  $[x, b'_{P_3}]$ ,  $[x, b'_{P_1}, b'_{P_3}]$ ,  $[x, b'_{P_2}, b'_{P_3}]$ , or  $[x, b'_{P_1}, b'_{P_2}, b'_{P_3}]$  has cost smaller than  $\delta$  from where it follows that

$$\tilde{u}_X^*(x, b'_{P_3}) < \delta. \quad (39)$$

This recursive construction can be repeated  $n - 1$  times to obtain partitions  $P_1, P_2, \dots, P_{n-1}$  and corresponding nodes

$b'_{P_1}, b'_{P_2}, \dots, b'_{P_{n-1}}$  such that the minimum chain cost satisfies

$$\tilde{u}_X^*(x, b'_{P_i}) < \delta, \quad \text{for all } i. \quad (40)$$

Observe now that the nodes  $b'_{P_i}$  are distinct by construction and distinct from  $x$ . Since there are  $n$  nodes in the network it must be that  $x' = b'_{P_k}$  for some  $i \in \{1, \dots, n - 1\}$ . It then follows from (40) that

$$\tilde{u}_X^*(x, x') < \delta. \quad (41)$$

This is a contradiction because  $x, x' \in X$  were assumed to satisfy (31). Thus, the assumption that (33) is true for *all* partitions is incorrect. Hence, the claim that there exists a partition  $P_\delta(x, x') = \{B_\delta(x), B_\delta(x')\}$  satisfying (32) must be true. ■

Returning to the main proof, given an arbitrary network  $N = (X, A_X)$  denote as  $(X, \tilde{u}_X) = \tilde{\mathcal{H}}(X, A_X)$  the output quasi-ultrametric resulting from application of an arbitrary admissible quasi-clustering method  $\tilde{\mathcal{H}}$ . We will show that for all  $x, x' \in X$

$$\tilde{u}_X^*(x, x') \leq \tilde{u}_X(x, x') \leq \tilde{u}_X^*(x, x'). \quad (42)$$

To prove the rightmost inequality in (42) we begin by showing that the dissimilarity function  $A_X$  acts as an upper bound on all admissible quasi-ultrametrics  $\tilde{u}_X$ , i.e.

$$\tilde{u}_X(x, x') \leq A_X(x, x'), \quad (43)$$

for all  $x, x' \in X$ . To see this, suppose  $A_X(x, x') = \alpha$  and  $A_X(x', x) = \beta$ . Define the two-node network  $N_{p,q} = (\{p, q\}, A_{p,q})$  where  $A_{p,q}(p, q) = \alpha$  and  $A_{p,q}(q, p) = \beta$  and denote by  $(\{p, q\}, \tilde{u}_{p,q}) = \tilde{\mathcal{H}}(N_{p,q})$  the output of applying the method  $\tilde{\mathcal{H}}$  to the network  $N_{p,q}$ . From axiom ( $\tilde{\text{A}}1$ ), we have  $\tilde{\mathcal{H}}(N_{p,q}) = N_{p,q}$ , in particular

$$\tilde{u}_{p,q}(p, q) = A_{p,q}(p, q) = A_X(x, x'). \quad (44)$$

Moreover, notice that the map  $\phi : \{p, q\} \rightarrow X$ , where  $\phi(p) = x$  and  $\phi(q) = x'$  is a dissimilarity reducing map, i.e. it does not increase any dissimilarity, from  $N_{p,q}$  to  $N$ . Hence, from axiom ( $\tilde{\text{A}}2$ ), we must have

$$\tilde{u}_{p,q}(p, q) \geq \tilde{u}_X(\phi(p), \phi(q)) = \tilde{u}_X(x, x'). \quad (45)$$

Substituting (44) in (45), we obtain (43).

Consider now an arbitrary chain  $C(x, x') = [x = x_0, x_1, \dots, x_l = x']$  linking nodes  $x$  and  $x'$ . Since  $\tilde{u}_X$  is a valid quasi-ultrametric, it satisfies the strong triangle inequality (7). Thus, we have that

$$\begin{aligned} \tilde{u}_X(x, x') &\leq \max_{i|x_i \in C(x, x')} \tilde{u}_X(x_i, x_{i+1}) \\ &\leq \max_{i|x_i \in C(x, x')} A_X(x_i, x_{i+1}), \end{aligned} \quad (46)$$

where the last inequality is implied by (43). Since by definition  $C(x, x')$  is an arbitrary chain linking  $x$  to  $x'$ , we can minimize (46) over all such chains maintaining the validity of the inequality,

$$\tilde{u}_X(x, x') \leq \min_{C(x, x')} \max_{i | x_i \in C(x, x')} A_X(x_i, x_{i+1}) = \tilde{u}_X^*(x, x'), \quad (47)$$

where the last equality is given by the definition of the directed minimum chain cost (13). Thus, the rightmost inequality in (42) is proved.

To show the leftmost inequality in (42), consider an arbitrary pair of nodes  $x, x' \in X$  and fix  $\delta = \tilde{u}_X^*(x, x')$ . Then, by Lemma 1, there exists a partition  $P_\delta(x, x') = \{B_\delta(x), B_\delta(x')\}$  of the node space  $X$  into blocks  $B_\delta(x)$  and  $B_\delta(x')$  with  $x \in B_\delta(x)$  and  $x' \in B_\delta(x')$  such that for all points  $b \in B_\delta(x)$  and  $b' \in B_\delta(x')$  we have

$$A_X(b, b') \geq \delta. \quad (48)$$

Focus on a two-node network  $N_{u,v} = (\{u, v\}, A_{u,v})$  with  $A_{u,v}(u, v) = \delta$  and  $A_{u,v}(v, u) = s$  where  $s = \text{sep}(X, A_X)$  as defined in (30). Denote by  $(\{u, v\}, \tilde{u}_{u,v}) = \tilde{\mathcal{H}}(N_{u,v})$  the output of applying the method  $\tilde{\mathcal{H}}$  to the network  $N_{u,v}$ . Notice that the map  $\phi : X \rightarrow \{u, v\}$  such that  $\phi(b) = u$  for all  $b \in B_\delta(x)$  and  $\phi(b') = v$  for all  $b' \in B_\delta(x')$  is dissimilarity reducing because, from (48), dissimilarities mapped to dissimilarities equal to  $\delta$  in  $N_{u,v}$  were originally larger. Moreover, dissimilarities mapped into  $s$  cannot have increased due to the definition of separation of a network (30). From Axiom ( $\tilde{\mathcal{A}}1$ ),

$$\tilde{u}_{u,v}(u, v) = A_{u,v}(u, v) = \delta, \quad (49)$$

since  $N_{u,v}$  is a two-node network. Moreover, since  $\phi$  is dissimilarity reducing, from ( $\tilde{\mathcal{A}}2$ ) we may assert that

$$\tilde{u}_X(x, x') \geq \tilde{u}_{u,v}(\phi(x), \phi(x')) = \delta, \quad (50)$$

where we used (49) for the last equality. Recalling that  $\tilde{u}_X^*(x, x') = \delta$  and substituting in (50) concludes the proof of the leftmost inequality in (42).

Since both inequalities in (42) hold, we must have  $\tilde{u}_X^*(x, x') = \tilde{u}_X(x, x')$  for all  $x, x' \in X$ . Since this is true for any arbitrary network  $N = (X, A_X)$ , it follows that the admissible quasi-clustering method must be  $\tilde{\mathcal{H}} \equiv \tilde{\mathcal{H}}^*$ .

#### A.4. The metric on $\mathcal{N}$

Consider two networks  $N_X, N_Y \in \mathcal{N}$  such that  $N_X = (X, A_X)$  and  $N_Y = (Y, A_Y)$ . A *correspondence* between the sets  $X$  and  $Y$  is any subset  $R \subseteq X \times Y$  such that  $\pi_1(R) = X$  and  $\pi_2(R) = Y$ . Here,  $\pi_1$  and  $\pi_2$  are the usual coordinate-wise projections. The *distortion*  $\text{dis}(R)$  of a correspondence  $R$  between networks  $N_X$  and  $N_Y$  is

defined as

$$\text{dis}(R) := \max_{(x,y), (x',y') \in R} |A_X(x, x') - A_Y(y, y')|.$$

The underlying notion of equality on  $\mathcal{N}$  that we use is the following: we say that networks  $N_X$  and  $N_Y$  are *isomorphic* or indistinguishable if and only if there exists a bijection  $\phi : X \rightarrow Y$  such that  $A_X(x, x') = A_Y(\phi(x), \phi(x'))$  for all  $x, x' \in X$ . Given  $N_X$  and  $N_Y$ , we define the *network distance*  $d_{\mathcal{N}}$  on  $\mathcal{N} \times \mathcal{N}$  as

$$d_{\mathcal{N}}(N_X, N_Y) := \frac{1}{2} \min_R \text{dis}(R), \quad (51)$$

where  $R$  spans all correspondences between  $X$  and  $Y$ . The structure of this distance is similar to that of the Gromov-Hausdorff distance (Gromov, 2007) that is often used in the context of compact metric spaces. In our context, it still provides a legitimate distance on the collection  $\mathcal{N}$  modulo our chosen notion of isomorphism.

**Theorem 4** *The network distance defined in (51) is a legitimate metric on  $\mathcal{N}$  modulo isomorphism of networks.*

**Proof:** That  $d_{\mathcal{N}}$  is symmetric and non-negative is clear. Assume now that  $X$  and  $Y$  are isomorphic and let  $\phi : X \rightarrow Y$  be a bijection providing this isomorphism. Then, consider  $R_\phi = \{(x, \phi(x)), x \in X\}$ . Since  $\phi$  is a bijection, it is easy to check that  $R_\phi$  is a correspondence between  $X$  and  $Y$ . Finally, by definition of  $\phi$ ,  $A_X(x, x') = A_Y(y, y')$  for all  $(x, y), (x', y') \in R_\phi$ . Hence

$$0 \leq d_{\mathcal{N}}(X, Y) \leq \frac{1}{2} \text{dis}(R_\phi) = 0$$

and  $d_{\mathcal{N}}(X, Y) = 0$  follows.

The triangle inequality follows from the following observation: if  $R$  is a correspondence between  $X$  and  $Z$  and  $S$  is a correspondence between  $Z$  and  $Y$ , then

$$T := \{(x, y), |\exists z \in Z \text{ with } (x, z) \in R, (z, y) \in S\} \quad (52)$$

is a correspondence between  $X$  and  $Y$ . To show that  $T$  is in fact a correspondence, we have to prove that for every  $x \in X$  there exists  $y \in Y$  such that  $(x, y) \in T$ . Similarly, we must require that for every  $y \in Y$  there exists  $x \in X$  such that  $(x, y) \in T$ . To see this, pick an arbitrary  $x \in X$ , by definition of  $R$ , there must exist  $z \in Z$  such that  $(x, z) \in R$ . By definition of  $S$ , there must exist  $y \in Y$  such that  $(z, y) \in S$ . Hence, there exists  $(x, y) \in T$  for every  $x \in X$ . Similarly, the result can be proven for every element of the set  $Y$ .

We can prove the triangle inequality in the following way. Consider  $R$  and  $S$  to be the minimizing correspondences

associated with distances  $d_{\mathcal{N}}(X, Z)$  and  $d_{\mathcal{N}}(Z, Y)$  respectively and define  $T$  as given by (52). Note that  $T$  need not be the minimizing correspondence for  $d_{\mathcal{N}}(X, Y)$ . Hence,

$$d_{\mathcal{N}}(X, Y) \leq \frac{1}{2} \text{dis}(T). \quad (53)$$

Furthermore, if we add and subtract  $A_Z(z, z')$  within the absolute value defining the distortion of  $T$  in (53), where  $z$  and  $z'$  are the elements in the definition of  $T$  (52), and we use the fact that the maximum of the absolute value of a sum is less than or equal to the sum of the maximums of absolute values, we obtain

$$\begin{aligned} & d_{\mathcal{N}}(X, Y) \\ & \leq \frac{1}{2} \max_{(x,z), (x',z') \in R} |A_X(x, x') - A_Z(z, z')| \\ & + \frac{1}{2} \max_{(z,y), (z',y') \in S} |A_Z(z, z') - A_Y(y, y')|. \end{aligned} \quad (54)$$

By noting that the expression on the right hand side of (54) is the sum of  $d_{\mathcal{N}}(X, Z)$  and  $d_{\mathcal{N}}(Z, Y)$ , the proof of the triangle inequality is completed.

Finally, the most delicate part of the proof is checking that  $d_{\mathcal{N}}(X, Y) = 0$  implies that  $X$  and  $Y$  are isomorphic. Assume that  $R$  is a correspondence such that  $A_X(x, x') = A_Y(y, y')$  for all  $(x, y)$  and  $(x', y')$  both in  $R$ . Define  $\phi : X \rightarrow Y$  in the following way: for each  $x \in X$  let  $Rx \subseteq Y$  be the set of all  $y$  such that  $(x, y) \in R$ . The fact that  $R$  is a correspondence forces that  $Rx \neq \emptyset$ . Hence, we can choose *any*  $y$  in  $Rx$  and declare  $\phi(x) = y$ .

Define in the same way a function  $\psi : Y \rightarrow X$ . Notice that then we forcibly have that  $A_X(x, x') = A_Y(\phi(x), \phi(x'))$  for all  $x, x' \in X$  and also  $A_X(\psi(y), \psi(y')) = A_Y(y, y')$  for all  $y, y' \in Y$ .

To prove that  $\phi$  is *injective*, assume that  $x \neq x'$  but  $\phi(x) = \phi(x')$ , then  $A_X(x, x') = A_Y(\phi(x), \phi(x')) = 0$ , which contradicts our definition of networks. In a similar manner one checks that  $\psi$  must also be injective.

So we have constructed two injections, one from  $X$  into  $Y$ , and one in the opposite direction. The Cantor-Bernstein-Schroeder theorem now applies and guarantees that there exists a bijection between  $X$  and  $Y$ . This immediately forces  $X$  and  $Y$  to have the same cardinality, and in particular, it forces  $\phi$  (and  $\psi$ ) to be bijections. This concludes the proof.  $\blacksquare$

### A.5. Proof of Theorem 3

Assume  $\eta = d_{\mathcal{N}}(N_X, N_Y)$  and let  $R$  be a correspondence between  $X$  and  $Y$  such that  $\text{dis}(R) = 2\eta$ . Write  $(X, \tilde{u}_X) = \tilde{\mathcal{H}}^*(N_X)$  and  $(Y, \tilde{u}_Y) = \tilde{\mathcal{H}}^*(N_Y)$ . We will prove that  $|\tilde{u}_X(x, x') - \tilde{u}_Y(y, y')| \leq 2\eta$  for all

$(x, y), (x', y') \in R$  which will imply the claim. Fix  $(x, y)$  and  $(x', y')$  in  $R$ . Pick any  $x = x_0, x_1, \dots, x_n = x'$  in  $X$  such that  $\max_i A_X(x_i, x_{i+1}) = \tilde{u}_X(x, x')$ . Choose  $y_0, y_1, \dots, y_n \in Y$  so that  $(x_i, y_i) \in R$  for all  $i = 0, 1, \dots, n$ . Then, by definition of  $\tilde{u}_Y(y, y')$  and the definition of  $\eta$ :

$$\begin{aligned} \tilde{u}_Y(y, y') & \leq \max_i A_Y(y_i, y_{i+1}) \\ & \leq \max_i A_X(x_i, x_{i+1}) + 2\eta \\ & = \tilde{u}_X(x, x') + 2\eta. \end{aligned}$$

By symmetry, one also obtains  $\tilde{u}_X(x, x') \leq \tilde{u}_Y(y, y') + 2\eta$ , and the conclusion follows from the arbitrariness of  $(x, y), (x', y') \in R$  and the definition of  $d_{\mathcal{N}}$ .

### A.6. Proof of Proposition 3

Fix any  $(X, A_X) \in \mathcal{N}$  and write  $\tilde{\mathcal{H}}^*(X, A_X) = (X, \tilde{u}_X)$ . Pick any change of scale function  $\Psi$  and write  $(X, \tilde{u}_X^\Psi) = \tilde{\mathcal{H}}^*(X, \Psi(A_X))$ . We need to prove that  $\tilde{u}_X^\Psi = \Psi(\tilde{u}_X)$ . But this follows directly from the explicit structure given in equation (13) and the fact that  $\Psi$  is non-decreasing.

### A.7. Further invariances: vertex permutations and the metric closure

Note that Theorem 3 implies that DSL behaves well under permutations of the vertices. The distance between a given network and a second one obtained by permuting its nodes is null. Thus, by Theorem 3, the distance between the corresponding output quasi-dendrograms must be null as well. More precisely, if  $(X, A_X) \in \mathcal{N}$ ,  $\tilde{\mathcal{H}}^*(X, A_X) = (X, \tilde{u}_X)$ , and  $\varphi : X \rightarrow X$  is any bijection, then  $\tilde{\mathcal{H}}^*(X, A_X \circ (\varphi, \varphi)) = (X, \tilde{u}_X \circ (\varphi, \varphi))$ . This means that permuting the labels of points before applying DSL yields the same result as permuting the labels a posteriori.

For any  $(X, A_X) \in \mathcal{N}$  let  $\bar{A}_X$  be the *maximal* function satisfying  $\bar{A}_X \leq A_X$  pointwisely which in addition satisfies the *directed triangle inequality*:  $\bar{A}_X(x, x') \leq \bar{A}_X(x, x'') + \bar{A}_X(x'', x')$  for all  $x, x', x'' \in X$ . Then, one can also prove (similar to the proof of Theorem 18 in (Carlsson & Mémoli, 2010)) that  $\tilde{\mathcal{H}}^*(X, A_X) = \tilde{\mathcal{H}}^*(X, \bar{A}_X)$  for all  $X \in \mathcal{N}$ .

### A.8. Proof of Proposition 4

In Ch.6, Section 6.1 of (Gondran & Minoux, 2008) it is shown that if  $A_X$  is a dissimilarity matrix then its quasi inverse  $A_X^*$  in the dioid  $(\mathbb{R}^+ \cup \{+\infty\}, \min, \max)$  contains information about the minimum infinity norm of chains in the network. In fact,  $[A_X^*]_{i,j}$  contains the minimum infinity norm of all the chains connecting node  $i$  with node  $j$ . In (Gondran & Minoux, 2008), the analysis is done for the symmetric case but its extension to the asymmetric case is im-

mediate as we present here,

$$[A_X^*]_{i,j} = \min_{C(x_i, x_j)} \max_{k | x_k \in C(x_i, x_j)} A_X(x_k, x_{k+1}). \quad (55)$$

By comparing (55) with (13), we can state that

$$A_X^* = \tilde{u}_X^*. \quad (56)$$

Hence, if we show that  $A_X^* = A_X^{(n-1)}$ , then (56) implies (16), completing the proof. Recall the quasi inverse  $A_X^*$  definition in the dioid  $(\mathbb{R}^+ \cup \{+\infty\}, \min, \max)$  from Ch. 4, Definition 3.1.2 in (Gondran & Minoux, 2008)

$$A_X^* = \lim_{k \rightarrow \infty} I \oplus A_X \oplus A_X^{(2)} \oplus \dots \oplus A_X^{(k)}, \quad (57)$$

where  $I$  has zeros in the diagonal and  $+\infty$  in the off diagonal elements.

However, in our dioid algebra where the  $\oplus$  operation is idempotent, i.e.  $a \oplus a = a$  for all  $a$ , it can be shown as in Ch. 4, Proposition 3.1.1 in (Gondran & Minoux, 2008) that

$$I \oplus A_X \oplus A_X^{(2)} \oplus \dots \oplus A_X^{(k)} = (I \oplus A_X)^{(k)}. \quad (58)$$

In our case, it is immediate that  $I \oplus A_X = A_X$ , since diagonal elements are null in both matrices and the off diagonal elements in  $I$  are  $+\infty$ . Hence, the minimization operation  $\oplus$  preserves  $A_X$ . Consequently, (58) becomes

$$I \oplus A_X \oplus A_X^{(2)} \oplus \dots \oplus A_X^{(k)} = A_X^{(k)}. \quad (59)$$

Taking the limit to infinity in both sides of equality (59) and using the quasi inverse definition (57), we get

$$A_X^* = \lim_{k \rightarrow \infty} A_X^{(k)}. \quad (60)$$

Finally, it can be shown as in Theorem 1 of Ch.4, Section 3.3 in (Gondran & Minoux, 2008) that  $A_X^{(n-1)} = A_X^{(n)}$ , proving that the limit in (60) is well defined and, more importantly, that  $A_X^* = A_X^{(n-1)}$ , as wanted.

## A.9. Applications

The dissimilarity function  $A_S$  of the migration network  $N_S$  used in Section 4 of the paper is computed as follows. Denote by  $M : S \times S \rightarrow \mathbb{R}_+$  the migration flow function given by the U.S. census bureau in which  $M(s, s')$  is the number of individuals that migrated from state  $s$  to  $s'$  in year 2011 and  $M(s, s) = 0$  for all  $s, s' \in S$ . We then construct the asymmetric network  $N_S = (S, A_S)$  with node set  $S$  and dissimilarities  $A_S$  such that  $A_S(s, s) = 0$  for all  $s \in S$  and

$$A_S(s, s') = f \left( \frac{M(s, s')}{\sum_i M(s_i, s')} \right), \quad (61)$$

for all  $s \neq s' \in S$  where  $f : [0, 1) \rightarrow \mathbb{R}_{++}$  is a given decreasing function. The normalization  $M(s, s') / \sum_i M(s_i, s')$  in (61) can be interpreted as the probability that an immigrant to state  $s'$  comes from state  $s$ . The role of the decreasing function  $f$  is to transform the similarities  $M(s, s') / \sum_i M(s_i, s')$  into corresponding dissimilarities. For the experiments here we use  $f(x) = 1 - x$ . However, due to the scale invariance property of DSL [cf. Proposition 3], the particular form of  $f$  is of little consequence to our analysis. Indeed, the influence structure between blocks of states obtained when quasi-clustering the network  $N_S$  is independent of the particular choice of the decreasing function  $f$ .

In Fig. 4 we present the dendrogram component  $D_S^*$  of the quasi-dendrogram  $\tilde{D}_S^* = (D_S^*, E_S^*)$  analyzed in Section 4. Some identifiable clusters are highlighted in color to illustrate the influence of geographical proximity in migrational preference. E.g., the blue cluster corresponds to the six states in the region of New England, the red cluster contains the remaining East Coast states with the exception of Delaware, and the green cluster corresponds to states in an extended West Coast plus Texas.

As a second illustrative example of the DSL method, we quasi-cluster a network that records interactions between sectors of the economy. The Bureau of Economic Analysis of the U.S. Department of Commerce publishes a yearly table of inputs and outputs organized by economic sectors (Bureau of Economic Analysis, 2011). This table records how economic sectors interact to generate gross domestic product. We focus on the section of uses of this table which shows the inputs to production. More precisely, we are given a set  $I$  of 61 industrial sectors as defined by the North American Industry Classification System (NAICS) and a function  $U : I \times I \rightarrow \mathbb{R}_+$  where  $U(i, i')$  for all  $i, i' \in I$  represents how much of the production of sector  $i$ , expressed in dollars, is used as an input of sector  $i'$ . The function  $U$  should be interpreted as a measure of directed closeness between two sectors. Thus, we define the network of uses  $N_I = (I, A_I)$  where the dissimilarity function  $A_I$  satisfies  $A_I(i, i) = 0$  and, for  $i \neq i' \in I$ , is given by

$$A_I(i, i') = f \left( \frac{U(i, i')}{\sum_k U(i_k, i')} \right), \quad (62)$$

where  $f : [0, 1) \rightarrow \mathbb{R}_{++}$  is a given decreasing function. The normalization  $U(i, i') / \sum_k U(i_k, i')$  in (62) can be interpreted as the probability that an input dollar to productive sector  $i'$  comes from sector  $i$ . In this way, we focus on the combination of inputs of a sector rather than the size of the economic sector itself. That is, a small dissimilarity from sector  $i$  to sector  $i'$  implies that sector  $i'$  highly relies on the use of sector  $i$  output as an input for its own production. Notice that  $U(i, i)$  for  $i \in I$  is generally positive, i.e., a sector uses outputs of its own production as inputs

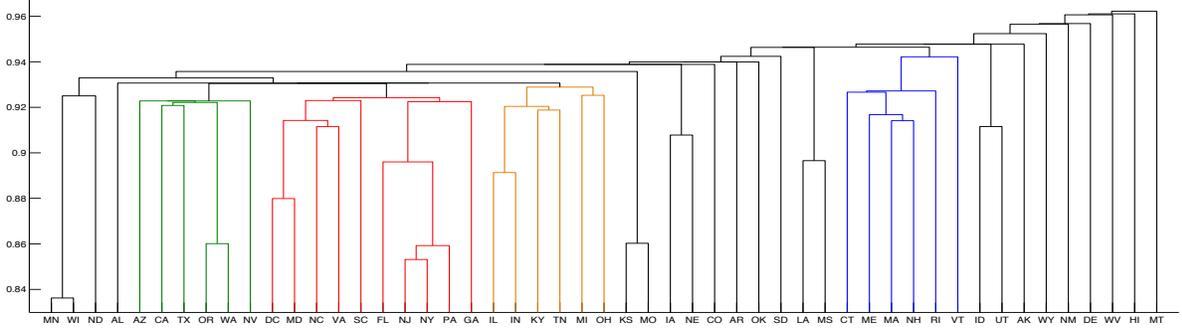


Figure 4. Dendrogram component  $D_S^*$  of the quasi-dendrogram  $\tilde{D}_S^* = (D_S^*, E_S^*)$ . The clustering of states is highly influenced by geographical proximity.

Table 1. Code and description of industrial sectors

Code	Industrial Sector
OG	Oil and gas extraction
CO	Construction
PC	Petroleum and coal products
WH	Wholesale trade
FR	Federal Reserve banks and credit intermediation
SC	Securities, commodity contracts, and investments
RA	Real estate
RL	Rental and leasing serv. and lessors of intang. assets
MP	Misc. professional, scientific, and technical services
AS	Administrative and support services

in other processes. Consequently, if for a given sector we sum the input proportion from every other sector, we obtain a number less than 1. The role of the decreasing function  $f$  is to transform the similarities  $U(i, i') / \sum_k U(i_k, i')$  into corresponding dissimilarities. As in the previous application, we use  $f(x) = 1 - x$ , though the particular form of  $f$  is of little consequence to the analysis since DSL is scale invariant [cf. Proposition 3].

The outcome of applying the DSL quasi-clustering method  $\tilde{\mathcal{H}}^*$  with output quasi-ultrametrics defined in (13) to the network  $N_I$  is computed with the algorithmic formula in (16). As we did with the migration network, in order to facilitate understanding we present quasi-partitions obtained by restricting the output quasi-ultrametric to a subset of nodes. In Fig. 5 we present four quasi-partitions focusing on ten economic sectors; see Table 1. We present quasi-partitions  $\tilde{D}_I^*(\delta)$  for four different resolutions  $\delta_1^* = 0.884$ ,  $\delta_2^* = 0.886$ ,  $\delta_3^* = 0.894$ , and  $\delta_4^* = 0.899$ .

The edge component  $E_I^*$  of the quasi-dendrogram  $\tilde{D}_I^*$  captures the asymmetric influence between clusters. E.g. in the quasi-partition in Fig. 5 for resolution  $\delta_1^* = 0.884$  every cluster is a singleton since the resolution is smaller than that of the first merging. However, the influence structure

reveals an asymmetry in the dependence between the economic sectors. At this resolution the professional service sector MP has influence over every other sector except for the rental services RL as depicted by the eight arrows leaving the MP sector. No sector has influence over MP at this resolution since this would imply, except for RL, the formation of a non-singleton cluster. The influence of MP reaches primary sectors as OG, secondary sectors as PC and tertiary sectors as AS or SC. The versatility of MP's influence can be explained by the diversity of services condensed in this economic sector, e.g. civil engineering and architectural services are demanded by CO, production engineering by PC and financial consulting by SC. For the rest of the influence pattern, we can observe an influence of CO over OG mainly due to the construction and maintenance of pipelines, which in turn influences PC due to the provision of crude oil for refining. Thus, from the transitivity (QP2) property of quasi-partitions we have an influence edge from CO to PC. The sectors CO, PC and OG influence the support service sector AS. Moreover, the service sectors RA, SC and FR have a totally hierarchical influence structure where SC has influence over the other two and FR has influence over RA. Since these three nodes remain as singleton clusters for the resolutions studied, the influence structure described is preserved for higher resolutions as it should be from the influence hierarchy property of the edge set  $E_S^*(\delta)$  stated in condition ( $\tilde{D}3$ ) in the definition of quasi-dendrogram in Section 3.1.

At resolution  $\delta_2^* = 0.886$ , we see that the sectors OG-PC-CO have formed a three-node cluster depicted in red that influences AS. At this resolution, the influence edge from MP to RL appears and, thus, MP gains influence over every other cluster in the quasi-partition including the three-node cluster. At resolution  $\delta = 0.887$  the service sectors AS and MP join the cluster OG-PC-CO and for  $\delta_3^* = 0.894$  we have this five-node cluster influencing the other five singleton clusters plus the mentioned hierarchical structure

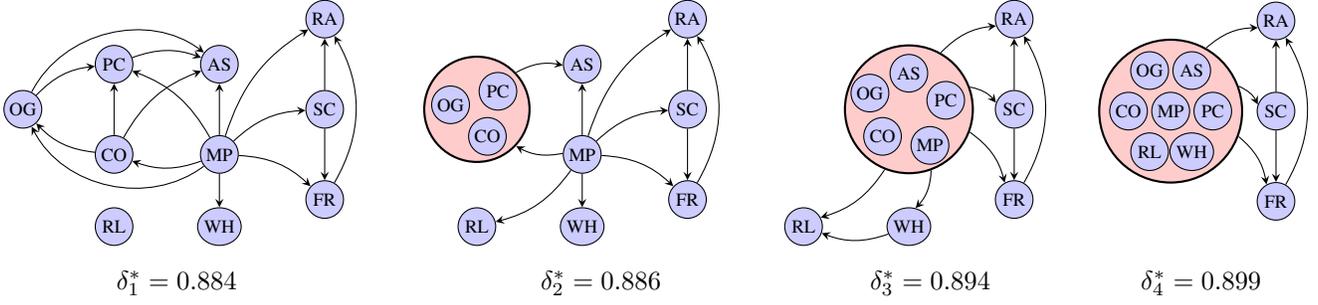


Figure 5. Directed single linkage quasi-clustering method applied to a portion of the sectors of the economy. The edges define a partial order among the blocks of every quasi-partition.

among SC, FR, and RA and an influence edge from WH to RL. When we increase the resolution to  $\delta_4^* = 0.899$  we see that RL and WH have joined the main cluster that influences the other three singleton clusters. If we keep increasing the resolution, we would see at resolution  $\delta = 0.900$  the sectors SC and FR joining the main cluster which would have influence over RA the only other cluster in the quasi-partition. Finally, at resolution  $\delta = 0.909$ , RA joins the main cluster and the quasi-partition contains only one block.

The influence structure between clusters at any given resolution defines a partial order. More precisely, for every resolution  $\delta$ , the edge set  $E_I^*(\delta)$  defines a partial order between the blocks given by the partition  $D_I^*(\delta)$ . We can use this partial order to evaluate the relative importance of different clusters by stating that more important sectors have influence over less important ones. E.g., at resolution  $\delta_1^* = 0.884$  we have that MP is more important than every other sector except for RL, which is incomparable at this resolution. There are three totally ordered chains that have MP as the most important sector at this resolution. The first one contains five sectors which are, in decreasing order of importance, MP, CO, OG, PC, and AS. The second one is comprised of MP, SC, FR, and RA and the last one only contains MP and WH. At resolution  $\delta_2^* = 0.886$  we observe that the three-node cluster OG-PC-CO, although it contains more nodes than any other cluster, it is not the most important of the quasi-partition. Instead, the singleton cluster MP has influence over the three-node cluster and, on top of that, is comparable with every other cluster in the quasi-partition. From resolution  $\delta_3^* = 0.894$  onwards, after MP joins the red cluster, the cluster with the largest number of nodes coincides with the most important of the quasi-partition. At resolution  $\delta_4^* = 0.899$  we have a total order among the four clusters of the quasi-partition. This is not true for the other three depicted quasi-partitions.

As a further illustration of the quasi-clustering method  $\tilde{\mathcal{H}}^*$ , we apply this method to the network  $N_C = (C, A_C)$  of consolidated industrial sectors (Bureau of Economic Analy-

Table 2. Code and description of consolidated industrial sectors

Code	Consolidated Industrial Sector
AGR	Agriculture, forestry, fishing, and hunting
MIN	Mining
UTI	Utilities
CON	Construction
MAN	Manufacturing
WHO	Wholesale trade
RET	Retail trade
TRA	Transportation and warehousing
INF	Information
FIR	Finance, insurance, real estate, rental, and leasing
PRO	Professional and business services
EHS	Educational services, health care, and social assistance
AER	Arts, entertain., recreation, accomm., and food serv.
OSE	Other services, except government

sis, 2011) where  $|C| = 14$  – see Table 2 – instead of the original 61 sectors. Dissimilarity function  $A_C$  is analogous to  $A_I$  but computed for the consolidated sectors. Of the output quasi-dendrogram  $\tilde{D}_C^* = (D_C^*, E_C^*)$ , in Fig. 6-(a) we show the dendrogram component  $D_C^*$  and in Fig. 6-(b) we depict the quasi-partitions  $\tilde{D}_C^*(\delta_i^{**})$  for  $\delta_1^{**} = 0.787$ ,  $\delta_2^{**} = 0.845$ ,  $\delta_3^{**} = 0.868$ ,  $\delta_4^{**} = 0.929$ , and  $\delta_5^{**} = 0.933$ . The reason we use the consolidated network  $N_C$  is to facilitate the visualization of quasi-partitions that capture every sector of the economy instead of only ten particular sectors as in the previous application.

The quasi-dendrogram  $\tilde{D}_C^*$  captures the asymmetric influences between clusters of industrial sectors at every resolution. E.g., at resolution  $\delta_1^{**} = 0.787$  the dendrogram  $D_C^*$  in Fig. 6-(a) informs us that every industrial sector forms its own singleton cluster. However, this simplistic representation, characteristic of clustering methods, ignores the asymmetric relations between clusters at resolution  $\delta_1^{**}$ . These influence relations are formalized in the quasi-dendrogram  $\tilde{D}_C^*$  with the introduction of the edge set  $E_C^*(\delta)$  for every resolution  $\delta$ . In particular, for  $\delta_1^{**}$  we see in Fig. 6-(b) that the sectors of ‘Finance, insurance, real estate, rental, and leasing’ (FIR) and ‘Manufacturing’

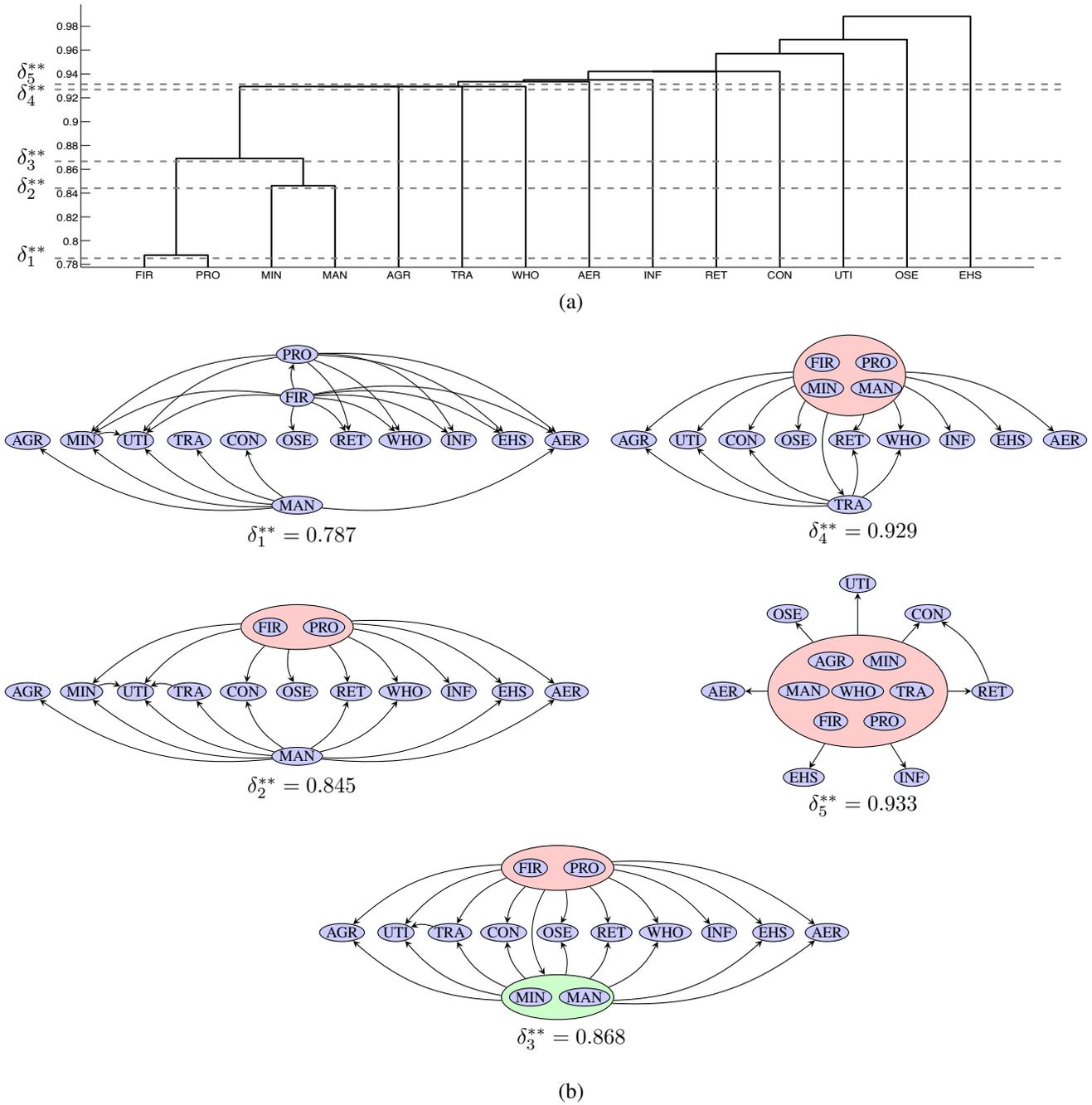


Figure 6. (a) Dendrogram component  $D_C^*$  of the quasi-dendrogram  $\tilde{D}_C^* = (D_C^*, E_C^*)$ . Output of the DSL quasi-clustering method  $\tilde{H}^*$  when applied to the network  $N_C$ . (b) Quasi-partitions. Given by the specification of the quasi-dendrogram  $\tilde{D}_C^*$  at a particular resolution  $\tilde{D}_C^*(\delta_k^{**})$  for  $k = 1, \dots, 5$ .

(MAN) combined have influence over the remaining 12 sectors. More precisely, the influence of FIR is concentrated on the service and commercialization sectors of the economy whereas the influence of MAN is concentrated on primary sectors, transportation, and construction. Furthermore, note that due to the transitivity (QP2) property of quasi-partitions defined in Section 3, the influence of FIR over ‘Professional and business services’ (PRO) implies influence of FIR over every sector influenced by PRO. The influence among the remaining 11 sectors, i.e. excluding MAN, FIR and PRO, is minimal, with the ‘Mining’ (MIN) sector influencing the ‘Utilities’ (UTI) sector. This influence is promoted by the influence of the ‘Oil and gas extraction’ (OG) subsector of MIN over the utilities sector. At resolution  $\delta_2^{**} = 0.845$ , FIR and PRO form one cluster, depicted in red, and they add an influence to the ‘Construction’ (CON) sector apart from the previously formed influences that must persist due to the influence hierarchy property of the edge set  $E_C^*(\delta)$  stated in condition (D3) in the definition of quasi-dendrogram in Section 3.1. The manufacturing sector also intensifies its influences by reaching the commercialization sectors ‘Retail trade’ (RET) and ‘Wholesale trade’ (WHO) and the service sector ‘Educational services, health care, and social assistance’ (EHS). The influence among the rest of the sectors is still scarce with the only addition of the influence of ‘Transportation and warehousing’ (TRA) over UTI. At resolution  $\delta_3^{**} = 0.868$  we see that mining MIN and manufacturing MAN form their own cluster, depicted in green. The previously formed red cluster has influence over every other cluster in the quasi-partition, including the green one. At resolution  $\delta_4^{**} = 0.929$ , the red and green clusters become one, composed of four original sectors. Also, the influence of the transportation TRA sector over the rest is intensified with the appearance of edges to the primary sector ‘Agriculture, forestry, fishing, and hunting’ (AGR), the construction CON sector and the commercialization sectors RET and WHO. Finally, at resolution  $\delta_5^{**} = 0.933$  there is one clear main cluster depicted in red and composed of seven sectors spanning the primary, secondary, and tertiary sectors of the economy. This main cluster influences ev-

ery other singleton cluster. The only other influence in the quasi-partition  $\tilde{D}_C^*(0.933)$  is the one of RET over CON. For increasing resolutions, the singleton clusters join the main red cluster until at resolution  $\delta = 0.988$  the 14 sectors form one single cluster.

The influence structure at every resolution induces a partial order in the blocks of the corresponding quasi-partition. As done in previous examples, we can interpret this partial order as a relative importance ordering. E.g., we can say that at resolution  $\delta_1^{**} = 0.787$ , MAN is more important than MIN which in turn is more important than UTI which is less important than PRO. However, PRO and MAN are not comparable at this resolution. At resolution  $\delta_4^{**} = 0.929$ , after the red and green clusters have merged together at resolution  $\delta = 0.869$ , we depict the combined cluster as red. This representation is not arbitrary, the red color of the combined cluster is inherited from the most important of the two component cluster. The fact that the red cluster is more important than the green one is represented by the edge from the former to the latter in the quasi-partition at resolution  $\delta_3^{**}$ . In this sense, the edge component  $E_C^*$  of the quasi-dendrogram formalizes a hierarchical structure between clusters at a fixed resolution apart from the hierarchical structure across resolutions given by the dendrogram component  $D_C^*$  of the quasi-dendrogram. E.g., if we focus only on the dendrogram  $D_C^*$  in Fig. 6-(a), the nodes MIN and MAN seem to play the same role. However, when looking at the quasi-partitions at resolutions  $\delta_1^{**}$  and  $\delta_2^{**}$ , it is clear that MAN has influence over a larger set of nodes than MIN and hence plays a more important role in the clustering for increasing resolutions. Indeed, if we delete the three nodes with the strongest influence structure, namely PRO, FIR, and MAN, and apply the quasi-clustering method  $\tilde{\mathcal{H}}^*$  on the remaining 11 nodes, the first merging occurs between the mining MIN and utilities UTI sectors at  $\delta = 0.960$ . At this same resolution, in the original dendrogram component in Fig. 6-(a), a main cluster composed of 12 nodes only excluding ‘Other services, except government’ (OSE) and EHS is formed. This indicates that by removing influential sectors of the economy, the tendency to cluster of the remaining sectors is decreased.