Block Realizations of Finite Metrics and the Tight-Span Construction I: The Embedding Theorem

Andreas W. M. Dress andreas@picb.ac.cn
CAS-MPG Partner Institute for Computational Biology,
320 Yue Yang Road, 200031 Shanghai, China.
Katharina T. Huber katharina.huber@cmp.uea.ac.uk
School of Computing Sciences, University of East Anglia,
Norwich, NR4 7TJ, UK.
Jacobus Koolen koolen@postech.ac.kr
Department of Mathematics, POSTECH,
Pohang, South Korea.
Vincent Moulton vincent.moulton@cmp.uea.ac.uk
School of Computing Sciences, University of East Anglia,
Norwich, NR4 7TJ, UK.

December 21, 2007

Abstract

Given a finite set X and a proper metric $D : X \times X \to \mathbb{R}_{\geq 0}$ defined on X, we show that every **block realization** of D can be "embedded" canonically into the tight span $T(D)$ of D and characterize the subsets of $T(X)$ that can be obtained in that way as the "canonical image" of the vertex set of a block realization.

Keywords and Phrases: metric, block realization, tight span, cut points, cut vertices.

1 Introduction

Given a finite set X and a proper metric $D : X \times X \to \mathbb{R}_{\geq 0} : (x, y) \mapsto xy$ defined on X (i.e., a metric for which $xy = 0$ holds for some $x, y \in X$ if and only if one has $x = y$), recall that a **block realization** $\mathfrak{B}$ of D as defined in [5] is a weighted block graph (V, E, ℓ), i.e., a triple consisting of two finite sets $V = V_{\mathfrak{B}}$ and $E = E_{\mathfrak{B}}$ and a **length-assignment map** $\ell = \ell_{\mathfrak{B}} : E_{\mathfrak{B}} \to \mathbb{R}_{>0}$ such that

(BR1) $E \subseteq \binom{V}{2}$ holds and the graph $G = (V, E)$ is a connected block graph,

(BR2) V contains X and $xy = d(x, y)$ holds, for all $x, y \in X$, for the (necessarily unique and proper) largest symmetric map $d = d_\ell : V \times V \to \mathbb{R} \cup \{+\infty\}$ defined on $V \times V$ for which $d(u, v) \leq \ell(u, v)$ holds for every edge $\{u, v\} \in E$, and

(BR3) every vertex v in $V - X$ has degree at least 3 and is a **cut vertex** of G.

We will show here that block realizations of D are closely related to the so-called **tight span**

$$T(D) = T(X, D) := \{f \in \mathbb{R}^X : \forall_{x \in X} f(x) = \sup_{y \in X} \big(xy - f(y)\big)\}$$

of D, the metric space associated with D consisting of the union of all compact faces of the (non-compact) convex polytope

$$P(D) = P(X, D) := \{f \in \mathbb{R}^X : \forall_{x,y \in X}\ xy \leq f(x) + f(y)\},$$

endowed with the ℓ_∞-metric $\|...\|_\infty$, cf. [13], see also [1]–[12].

Furthermore, defining a triple $\mathfrak{B} = (V, E, \ell)$ as above to be a **weak** block realization of D if (BR1) and (BR2) hold, and every vertex v in $V - X$ is a **cut vertex** of G, but does not necessarily have degree at least 3, we will show that this relationship with $T(D)$ does not only extend immediately to weak block realizations, but that these weak realizations present an even more natural conceptual framework for dealing with it.

We'll employ the following notational conventions, definitions, and facts:

(N1) Given any simple graph $G = (V, E)$ with vertex set V and edge set $E \subseteq \binom{V}{2}$, and any vertex $v \in V$, we denote

   (i) by $G(v)$ the connected component of G containing v,

   (ii) by $\pi_0(G) := \{G(u) : u \in V\}$ the partition of V into the set of connected components of G,

   (iii) by $G - v$ the graph induced by G on the set $V - v := V - \{v\}$ so that $(G - v)(u)$ denotes, for any vertex $u \in V - v$, the connected component of $G - v$ containing u, and v is a cut vertex of G if and only if $|\pi_0(G - v)| > 1$ holds,

   (iv) and by $[E]$ the edge set of the smallest block graph with vertex set V containing the edge set E, i.e., the union of E and the collection of all 2-subsets e of V for which some "circuit" $e_1, \dots, e_m \in E$ exists[1] for which $e \subseteq \bigcup_{i=1}^m e_i$ holds.

[1] i.e., some family $e_1, \dots, e_m$ of pairwise distinct edges $e_1, \dots, e_m \in E$ with $e_i \cap e_j = \emptyset \iff 1 < |i - j| < m - 1$ for any two distinct indices $i, j \in \{1, \dots, m\}$.

(N2) Given any proper finite metric space (V, d) with point set V and metric d, and a subset E of $\binom{V}{2}$, note that the metric d coincides with the map $d_{\ell_{E,d}} : V \times V \to \mathbb{R} \cup \{+\infty\}$ induced on $V \times V$ by the length-assignment map $\ell_{E,d} : E \to \mathbb{R} : \{u, v\} \mapsto d(u, v)$ obtained by "restricting" d to E if and only if E contains the set

$$E(d) := \{\{u, v\} \in \binom{V}{2} : \forall_{w \in V - \{u,v\}}\ d(u, v) < d(u, w) + d(w, v)\}.$$

Thus, denoting by $\mathfrak{B}(V) = \mathfrak{B}(V, d) := (V, [E(d)], \ell_{[E(d)],d})$ the weighted block graph with vertex set V, edge set $[E(d)]$, and weighting map $\ell_{[E(d)],d}$, the map $d_{\ell_{[E(d)],d}}$ induced on $V \times V$ by $\mathfrak{B}(V)$ always coincides with the input metric d, and any cut point v of (V, d) must therefore also be a cut vertex of the associated block graph $(V, [E(d)])$.

(N3) Further, associating the so-called **Kuratowski map**

$$k_x = k_x^D : X \to \mathbb{R} : y \mapsto D(x, y)$$

(cf. [14]) to a given point $x \in X$, defines an isometry $K = K_D : x \mapsto k_x$ from the metric space (X, D) into $T(D)$ that maps X bijectively onto the subset $K(X) = K(X, D)$ consisting of all $f \in T(D)$ for which the support $\text{supp}(f) := \{x \in X : f(x) \neq 0\}$ is distinct from X: Indeed, given any map $f \in T(D)$ and any $x \in X$, one has $k_x \in T(D)$ and $f(x) = \|f, k_x\|_\infty$ and, therefore, $\|k_y, k_x\|_\infty = k_y(x) = yx$ for all $x, y \in X$.

We will henceforth identify each point $x \in X$ with the corresponding Kuratowski map $k_x \in T(D)$ and, thus, think of X as coinciding with the image $K(X)$ of the map $K = K_D$ in $T(D)$, and of the map $K = K_D$ as being the identity on X.

(N4) And finally, recall that a map $f \in P(D)$ is either a Kuratowski map or a **cut point** of $T(D)$ — that is, f is contained in $T(D)$ and the complement $T(D) - f$ of the one-point subset $\{f\}$ of $T(D)$ is disconnected — if and only if the graph (X, E_f) with edge set

$$E_f := \big\{\{x, y\} \in \binom{X}{2} : f(x) + f(y) > D(x, y)\big\}$$

is disconnected – and it is a cut point of $T(D)$ if and only if the induced subgraph $\Gamma_f := (\text{supp}(f), E_f)$ is disconnected[2]. Henceforth, we'll denote the set of all cut points of $T(D)$ by $cut(D)$, and the union of $cut(D)$ and $K(X)$ by $Cut(D)$.

Here are our main results

[2] More specifically (cf. [4]), associating to each connected component $A \in \pi_0(\Gamma_f)$ of the graph Γ_f the open subset $O_f(A) := \{g \in T(D) : f(x) < g(x) \text{ for all } x \in \text{supp}(f) - A\}$ of $T(D) - f$ defines a canonical bijection $\mathcal{O}_f : \pi_0(\Gamma_f) \to \pi_0\big(T(D) - f\big) : A \mapsto O_f(A)$ from the set $\pi_0(\Gamma_f)$ onto the set $\pi_0\big(T(D) - f\big)$ of connected components of the space $T(D) - f$.

Theorem 1 (i) *Given a weak block realization* $\mathfrak{B} = (G_{\mathfrak{B}} = (V_{\mathfrak{B}}, E_{\mathfrak{B}}), \ell_{\mathfrak{B}})$ *of* D*, the associated map*

$$\Phi_{\mathfrak{B}} : V_{\mathfrak{B}} \to \mathbb{R}^X : v \mapsto \big(f_v : X \to \mathbb{R}\colon x \mapsto d_{\mathfrak{B}}(x, v)\big)$$

maps $V_{\mathfrak{B}}$ *isometrically onto a finite subset of* $Cut(D)$ *of* $T(D)$*, and any point* $x \in X \subseteq V_{\mathfrak{B}}$ *onto itself, considered as an element of* $T(D)$*.*

Furthermore, given any $v \in V_{\mathfrak{B}}$*, the map* $\Phi_{\mathfrak{B}}$ *induces a well-defined surjective "inverse" mapping* $\pi_0^{(v)}$ *from* $\pi_0(\Gamma_{f_v})$ *onto* $\pi_0(G - v)$ *that maps any connected component* $\Gamma_{f_v}(x)$ $\big(x \in \mathrm{supp}(f_v)\big)$ *of* Γ_{f_v} *onto the connected component* $(G - v)(x)$ *of* $G - v$ *containing* x*.*

(ii) *Conversely, given a finite subset* V *of* $Cut(D)$ *that contains* X*, the weighted block graph* $\mathfrak{B}(V) = \mathfrak{B}(V, \|...\|_\infty)$ *associated with* V *considered as a metric space relative to the* ℓ_∞*-metric* $\|...\|_\infty$ *restricted to* V*, is a weak block realization of* D *for which the associated map* $\Phi_{\mathfrak{B}(V)}$ *coincides with the identity map* $\mathbf{Id}_V$ *on* V*.*

2 Two Lemmata on Block Realizations

In this section, we present two simple observations concerning weak block realizations.

Lemma 2.1 *Given a weak block realization* $\mathfrak{B} = (G = (V, E), \ell)$ *of a proper metric* D *defined on a finite set* X*, then* $X \cap C \neq \emptyset$ *holds for every vertex* $v \in V$ *and every connected component* $C \in \pi_0(G - v)$*.*

Proof: If $X \cap C$ were empty, every vertex in C would be a cut vertex of G. Choose any vertex $w \in C$ for which the distance $d(w, v)$ of w to v is maximal. Since $w \in C \subseteq V - X$ must be a cut vertex of G, there exists some vertex $w' \in V$ in $(G - w)(v)$. But then, $d(v, w') = d(v, w) + d(w, w') > d(v, w)$ and, therefore, also $w' \in C$ — contradicting together our choice of w. ∎

Lemma 2.2 *With* X, D*, and* $\mathfrak{B}$ *as in Lemma* 2.1*, there exist, for any two vertices* u, v *in* V*, some* $x, y \in X$ *such that* $xy = d(x, y) = d(x, u) + d(u, v) + d(v, y)$ *holds.*

Proof: It suffices to show that there exists some $x \in X$ such that $d(x, v) = d(x, u) + d(u, v)$ holds. Clearly, we may assume that $u \notin X$ holds, implying that u must be a cut vertex of G. Thus, there must exist a connected component C of $G - u$ with $v \notin C$. So, choosing any $x \in C \cap X$, we have $d(x, v) = d(x, u) + d(u, v)$ as claimed. ∎

3 Proof of the Main Result

(i) Assume that $\mathfrak{B} = (V, E, \ell)$ is weak block realization of D, and consider the associated metric $d = d_{\mathfrak{B}}$ induced by ℓ and E on V and map

$$\Phi = \Phi_{\mathfrak{B}} : V \to \mathbb{R}^X : v \mapsto \big(f_v : X \to \mathbb{R} \colon x \mapsto d(x, v)\big).$$

Clearly, Φ maps every vertex $x \in X$ onto the corresponding Kuratowski map, i.e., we have $f_x = x \in Cut(D)$ for all $x \in X \subseteq V$. To show that $f := f_v \in Cut(D)$ also holds in case $v \in V - X$, note first that

$$f(x) + f(y) = d(x, v) + d(y, v) \geq d(x, y) = xy$$

holds for all $x, y \in X$, i.e., we surely have $f \in P(D)$. Furthermore, we have

$$f(x) + f(y) = d(x, v) + d(y, v) = d(x, y) = xy$$

for all $x, y \in X$ for which the connected components $(G - v)(x), (G - v)(y)$ in $\pi_0(G - v)$ containing x and y, respectively, are distinct. So, the edge set E_f of the graph Γ_f cannot contain any edge connecting two vertices x, y with $(G-v)(x) \neq (G-v)(y)$ implying that Φ induces indeed a well-defined "inverse" mapping $\pi_0^{(v)}$ from $\pi_0(\Gamma_f)$ onto $\pi_0(G - v)$ that maps any connected component of Γ_f of the form $\Gamma_f(x)$ for some $x \in \mathrm{supp}(f)$ onto the connected component $(G - v)(x)$ of $G - v$ containing x.

Moreover, as every connected component of Γ_f is of this form for some x in $\mathrm{supp}(f)$ in view of Lemma 2.1, this map must be surjective implying in particular that $\#\pi_0(T(D) - f) = \#\pi_0(\Gamma_f) \geq \#\pi_0(G - v)$ must hold for every $v \in V$ for $f = f_v$ and that, therefore, f must be a cut point of $T(D)$ for every $v \in V - X$.

Next, note that Φ is **non-expansive**, i.e.,

$$\|f_u, f_v\|_\infty \leq d(u, v)$$

holds for all $u, v \in V$. Indeed, we have

$$|f_u(x) - f_v(x)| = |d(u, x) - d(v, x)| \leq d(u, v)$$

for all $x \in X$ in view of the triangle inequality applied to d.

Thus, Φ must actually be an isometry as, given any two vertices $u, v \in V$, we may choose points $x, y \in X$ according to Lemma 2.2 so that

$$xy = d(x, y) = d(x, u) + d(u, v) + d(v, y)$$

holds. So, the fact that Φ is non-expansive combined with the triangle inequality implies

$$\begin{aligned} xy = d(x, y) &= d(x, u) + d(u, v) + d(v, y) \\ &\geq \|k_x, f_u\|_\infty + \|f_u, f_v\|_\infty + \|f_v, k_y\|_\infty \\ &\geq \|k_x, k_y\|_\infty = xy \end{aligned}$$

which in turn implies that $d(u,v) = \|f_u, f_v\|_\infty$ must hold for any two vertices $u, v \in V$.

Together, this implies the first part of Theorem 1.

(ii) To establish the second part, it suffices to note that, given a finite subset V of $Cut(D)$ that contains X,

- the triple $\mathfrak{B} := \mathfrak{B}(V, \|...\|_\infty)$ associated with the finite metric space $(V, \|...\|_\infty)$ according to (N2) is, by construction, a connected weighted block graph whose vertex set contains X,
- the induced metric $d_{\mathfrak{B}}$ coincides with the input metric $\|...\|_\infty$ implying that also $xy = d_{\mathfrak{B}}(x,y)$ must hold for all $x, y \in X$,
- the map $\Phi_{\mathfrak{B}}$ therefore coincides with the identity map $\mathbf{Id}_V$ on V,
- and any point $v \in V - X$ must, therefore, be a cut vertex of the graph $G_V := (V, [E(d)])$.

So, $\mathfrak{B} = \mathfrak{B}(V, \|...\|_\infty)$ must be a weak block realization of D as claimed. ∎

Remark 3.1 *In* [10]*, we will characterize all weak block realizations* $\mathfrak{B} = (V, E, \ell)$ *of* D *for which* $V_{\mathfrak{B}}$ *is, as above, a finite subset* V *of* $Cut(D)$ *that contains* X *and* $\Phi_{\mathfrak{B}}$ *coincides with the identity map* $\mathbf{Id}_V$ *on* V*. And we will also characterize the "proper" block realizations* $\mathfrak{B} = (V, E, \ell)$ *among them.*

Acknowledgment The authors would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK, for hosting them in the context of its Phylogenetics Programme where part of the reserach presented in this paper was carried out. V. Moulton was supported in part by Engineering and Physical Sciences Research Council (EPSRC), grant EP/D068800/1, K. Huber and J. Koolen thank the Royal Society for supporting them through the International Joint Projects scheme, and A. Dress thanks the BMBF, Germany, and the Chinese Academy of Sciences for their support.

References

[1] H.-J. Bandelt, A. Dress, A canonical decomposition theory for metrics on a finite set, *Adv. in Math.*, **92**, 47-105, 1992.

[2] A. Dress, Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: A note on combinatorial properties of metric spaces, *Adv. in Math.*, **53**, 321-402, 1984.

[3] A. Dress, B. Holland, K. T. Huber, V. Moulton, J. Koolen, J. Weyer-Menkhoff, Δ additive and Δ ultra-additive maps Gromov's trees, and the Farris transform, *Disc. App. Math.*, **146**, 51-73, 2005.

[4] A. Dress, K. T. Huber, J. Koolen, V. Moulton, Cut points in metric spaces, *Appl. Math. Letters*, in press.

[5] A. Dress, K. T. Huber, J. Koolen, V. Moulton, Block graphs and connectivity maps, in preparation.

[6] A. Dress, K. T. Huber, J. Koolen, V. Moulton, A graph-theoretical invariant of toplogical spaces, in preparation.

[7] A. Dress, K. T. Huber, J. Koolen, V. Moulton, Compatible decompositions and block realizations of finite metric spaces, *Europ. J. Comb.* **28** (2007), Available online at "http://www.sciencedirect.com".

[8] A. Dress, K. T. Huber, J. Koolen, V. Moulton, An algorithm for computing virtual cut points in finite metric spaces, Proceedings of COCOA 2007, Lecture Notes in Computing Science, **4616** (2007), 4-10.

[9] A. Dress, K. T. Huber, J. Koolen, V. Moulton, A. Spillner, A note on metric cut points and bridges, submitted.

[10] A. Dress, K. Huber, J. Koolen, and V. Moulton, Block realizations of finite metrics and the tight-span construction II: On standard block realizations of finite metrics, in preparation.

[11] A. Dress, K. T. Huber, and V. Moulton, A comparison between two distinct continuous models in projective cluster theory: The median and the tight-span construction, *Ann. of Comb.*, **2**, 299-311, 1998.

[12] A. Dress, V. Moulton, W. Terhalle, T-theory: An overview, *Europ. J. Comb.*, **17**, 161-175, 1996.

[13] J. Isbell, Six theorems about metric spaces, *Comment. Math. Helv.*, **39**, 65-74, 1964.

[14] C. Kuratowski, Quelques problèmes concernant les espaces métriques nonséperables, *Fund. Math.*, **25**, 534-545, 1935.