On the Parameterized Complexity of Some Optimization Problems Related to Multiple-Interval Graphs

Minghui Jiang

Utah State University

CPM June 22, 2010

Intersection Graphs

The *intersection graph* $\Omega(\mathcal{F})$ of a family of sets

$$\mathcal{F} = \{S_1, \ldots, S_n\}$$

is the graph with $\mathcal{F}$ as the vertex set and with two different vertices S_i and S_j adjacent if and only if $S_i \cap S_j \neq \emptyset$.

The family $\mathcal{F}$ is called a *representation* of the graph $\Omega(\mathcal{F})$.

Disk intersection graphs, interval graphs, circle graphs, circular-arc graphs…

Multiple-Interval Graphs

Let t be an integer at least two.

A *t-interval* is the union of t disjoint intervals in the real line. A *t-track interval* is the union of t disjoint intervals in t disjoint parallel lines called tracks, one interval on each track.

A *t-interval graph* is the intersection graph of a family of t-intervals. A *t-track interval graph* is the intersection graph of a family of t-track intervals.

If all intervals in the representation of a t-interval graph have unit lengths, then the graph is called a *unit t-interval graph*. Similarly for *unit t-track interval graphs*.

Graph Hierarchy

The t disjoint tracks for a t-track interval graph can be viewed as t disjoint "host" intervals in the real line for a t-interval graph.

t-track interval graphs $\subset$ t-interval graphs

t-interval graphs $\subset$ $(t+1)$-interval graphs
t-track interval graphs $\subset$ $(t+1)$-track interval graphs

unit t-interval graphs $\subset$ t-interval graphs
unit t-track interval graphs $\subset$ t-track interval graphs

The most basic subclass: unit 2-track interval graphs

Applications

As generalizations of the ubiquitous interval graphs, multiple-interval graphs such as t-interval graphs and t-track interval graphs have wide applications, traditionally to *scheduling and resource allocation* and more recently to *bioinformatics*.

In particular, 2-interval graphs and 2-track interval graphs are natural models for the *similar regions of DNA sequences* and for the *helices of RNA secondary structures*.

Parameterized Complexity

In general graphs, the following four optimization problems, parameterized by the optimal solution size k, are exemplary problems in parameterized complexity theory:

- k-Vertex Cover: in FPT
- k-Independent Set / k-Clique: W[1]-hard
- k-Dominating Set: W[2]-hard

In Multiple-Interval Graphs…

Since t-interval graphs are a special class of graphs, all FPT algorithms for k-VERTEX COVER in general graphs immediately carry over to t-interval graphs.

The parameterized complexities of k-INDEPENDENT SET, k-CLIQUE, and k-DOMINATING SET in t-interval graphs, however, are not at all obvious.

In general graphs, k-INDEPENDENT SET and k-CLIQUE are essentially the same problem, but in t-interval graphs, they manifest different parameterized complexities…

Previous Results

Fellows, Hermelin, Rosamond, and Vialette (2009) recently initiated the study of the parameterized complexity of multiple-interval graph problems.

They showed:

1. k-Independent Set in t-interval graphs is W[1]-hard for any constant $t \geq 2$.
2. k-Dominating Set in t-interval graphs is also W[1]-hard for any constant $t \geq 2$.
3. k-Clique in t-interval graphs admits an FPT algorithm parameterized by both k and t.

Three Open Questions

Fellows et al. then raised three open questions:

1. Are k-INDEPENDENT SET and k-DOMINATING SET in 2-track interval graphs W[1]-hard?
2. Is k-DOMINATING SET in t-interval graphs W[2]-hard?
3. Can the parametric time-bound of their FPT algorithm for k-CLIQUE in t-interval graphs be improved?

We answer the first open question in the affirmative, and make a little progress on the third open question...

k-Multicolored Clique

Given a graph G and a vertex-coloring

$$\kappa : V(G) \to \{1, 2, \ldots, k\},$$

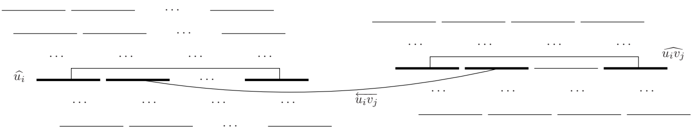
k-MULTICOLORED CLIQUE is the problem of deciding whether G has a clique of k vertices containing exactly one vertex of each color.

Fellows et al. proved that k-MULTICOLORED CLIQUE is W[1]-complete, then proved that both k-INDEPENDENT SET and k-DOMINATING SET in unit 2-interval graphs are W[1]-hard by FPT reductions from k-MULTICOLORED CLIQUE.

We prove:

Theorem 1. k-INDEPENDENT SET *and* k-DOMINATING SET *in unit 2-track interval graphs are* W[1]-*hard.*

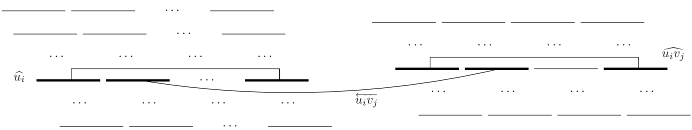
Previous Reduction for k-Independent Set



Let (G, κ, k) be an instance of k-MULTICOLORED CLIQUE. The construction consists of $k + \binom{k}{2}$ groups of unit intervals occupying disjoint regions of the real line:

- k groups are vertex gadgets, one for each color,
- $\binom{k}{2}$ groups are edge gadgets, one for each pair of distinct colors.

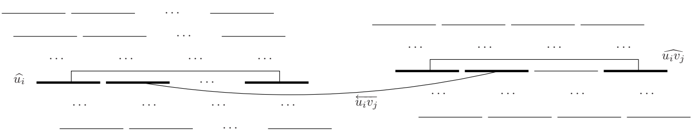
How It Works



The vertex gadgets and the edge gadgets are then linked together, according to the incidence relation between the vertices and the edges, by the validation gadget.

Each vertex gadget selects a vertex of a particular color.
Each edge gadget selects an edge of a particular pair of colors. The validation gadget ensures the consistency of the selections.

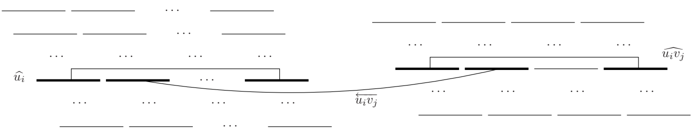
Vertex Selection



For each color i, $1 \leq i \leq k$, let V_i be the set of vertices with color i.

The vertex gadget for the color i consists of a group of intervals that can viewed as a table with $|V_i|$ rows and $k+1$ columns.

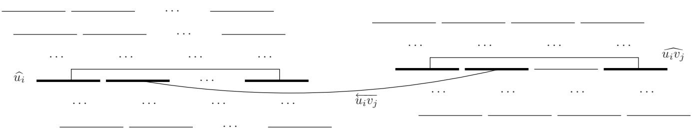
Vertex Selection



Each row of the table corresponds to a distinct vertex $u \in V_i$:

- the first interval and the last interval together form a *vertex 2-interval* $\widehat{u_i}$;
- the other intervals, each associated with a distinct color $j \in \{1, \ldots, k\} \setminus \{c\}$ and denoted by $\overline{u_i * j}$, and are used for validation.

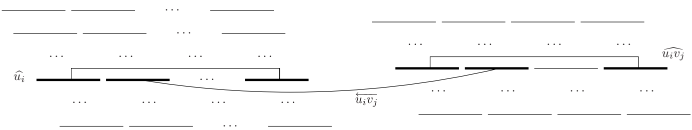
Vertex Selection



The intervals in the table are arranged in a parallelogram formation with slanted columns:

- the intervals in each row are disjoint;
- the intervals in each column intersect at a common point;
- the intervals in lower rows have larger horizontal offsets such that each interval also intersects all intervals in higher rows in the next column.

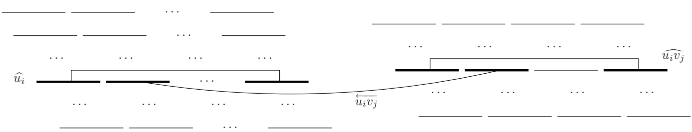
Edge Selection



For each pair of distinct colors i and j, $1 \leq i < j \leq k$, let E_{ij} be the set of edges uv such that u has color i and v has color j.

The edge gadget for the pair of colors ij consists of a group of intervals that can viewed as a table with $|E_{ij}|$ rows and 4 columns. Again the intervals in the table are arranged in a parallelogram formation.

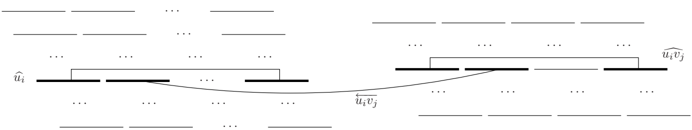
Edge Selection



Each row of the table corresponds to a distinct edge $uv \in E_{ij}$:

- the first interval and the fourth interval together form an *edge 2-interval* $\widehat{u_iv_j}$;
- the first interval and the fourth interval together form an the second and the third intervals, denoted by $\overline{u_iv_j}$ and $\overline{v_ju_i}$, respectively, are used for validation.

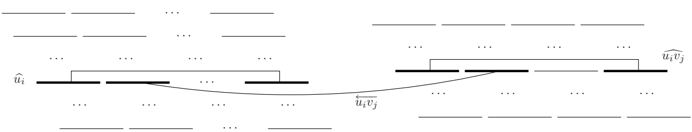
Validation



For each edge uv such that u has color i and v has color j, the validation gadget includes two *validation 2-intervals* $\overleftarrow{u_iv_j}$ and $\overrightarrow{u_iv_j}$:

- the 2-interval $\overleftarrow{u_iv_j}$ consists of the interval $\overline{u_iv_j}$ and the interval $\overline{u_i{*}_j}$;
- the 2-interval $\overrightarrow{u_iv_j}$ consists of the interval $\overline{v_ju_i}$ and the interval $\overline{v_j{*}_i}$.

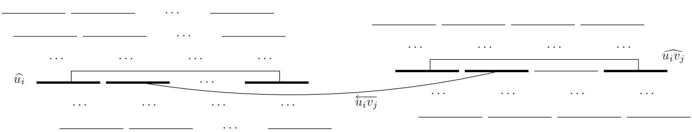
Validation



The vertex 2-interval $\widehat{u_i}$ selects the vertex u for the color i. The edge 2-interval $\widehat{u_iv_j}$ selects the edge uv for the pair of colors ij. The validation 2-interval $\overleftarrow{u_iv_j}$ validates the selections.

$$\mathcal{F} = \left\{\widehat{u_i} \mid u \in V_i,\, 1 \leq i \leq k\right\}$$
$$\cup \left\{\widehat{u_iv_j}, \overleftarrow{u_iv_j}, \overrightarrow{u_iv_j} \mid uv \in E_{ij},\, 1 \leq i < j \leq k\right\}.$$

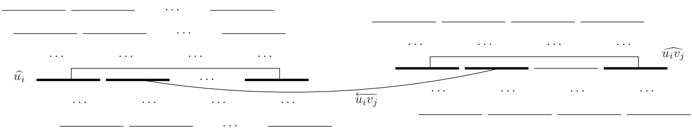
The Claim



Set the parameter $k' = k + 3\binom{k}{2}$.

> G has a k-multicolored clique if and only if $\mathcal{F}$ has a k'-independent set.

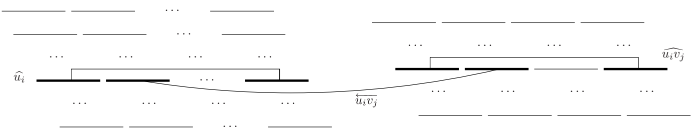
Direct Implication



If $K \subseteq V(G)$ is a k-multicolored clique, then the following subset of 2-intervals is a k'-independent set in $\mathcal{F}$:

$$\left\{ \widehat{u_i} \mid u \in K,\ i = \kappa(u) \right\}$$
$$\cup \left\{ \widehat{u_i v_j}, \overleftarrow{u_i v_j}, \overrightarrow{u_i v_j} \mid u, v \in K,\ i = \kappa(u),\ j = \kappa(v) \right\}.$$

Reverse Implication

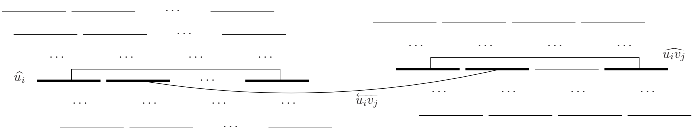


Suppose that $\mathcal{I}$ is a k'-independent set in $\mathcal{F}$.

By construction, $\mathcal{I}$ can include at most one vertex 2-interval for each color, and at most one edge 2-interval plus at most two validation 2-intervals for each pair of distinct colors.

Since $k' = k + 3\binom{k}{2}$, $\mathcal{I}$ must include exactly one vertex 2-interval for each color, and exactly one edge 2-interval plus two validation 2-intervals for each pair of distinct colors.

Reverse Implication



It follows that the $2\binom{k}{2} = (k-1)k$ validation 2-intervals in $\mathcal{I}$ have exactly two intervals in each edge gadget, and exactly $k-1$ intervals in each vertex gadget.

Moreover, in each vertex gadget, the intervals of the vertex 2-interval and the $k-1$ validation 2-intervals in $\mathcal{I}$ must be in the same row. Similarly, in each edge gadget, the intervals of the edge 2-interval and the two validation 2-intervals in $\mathcal{I}$ must be in the same row.

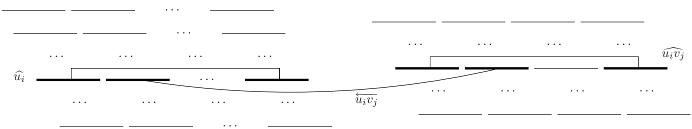
Some Intuition

The central idea behind the construction is essentially a geometric packing argument:

- Consider each vertex 2-interval as a container of capacity $k-1$, each edge 2-interval as a container of capacity 2, and the validation 2-intervals as items to be packed.

- In order to pack each container to its full capacity, the items in each container must be arranged in a regular pattern, that is, all intervals in each vertex or edge gadget must be in the same row.

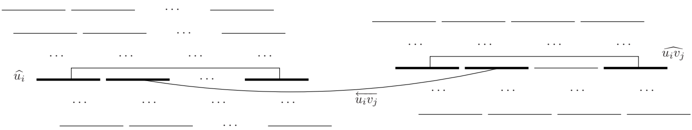
Reverse Implication



Since all intervals in the same row of a vertex gadget are associated with the same vertex, and all intervals in the same row of an edge gadget are associated with the same edge, the vertex selection and the edge selection must be consistent.

The k vertex 2-intervals in $\mathcal{I}$ corresponds to a k-multicolored clique in G.

New Reduction

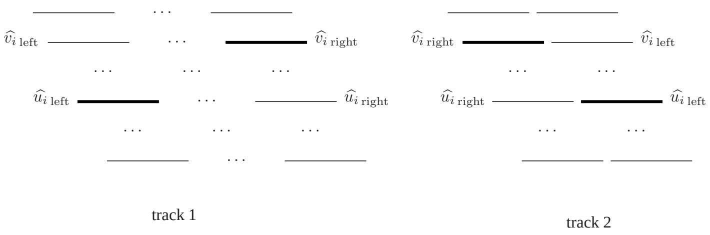


We now modify the previous construction to transform each 2-interval into a 2-track interval.

Move all vertex gadgets to track 1, and move all edge gadgets to track 2. Then all validation 2-intervals are immediately transformed into 2-track intervals.

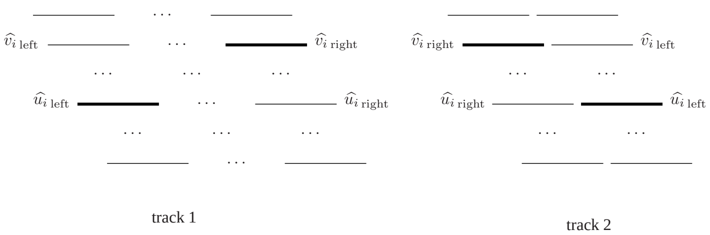
It remains to fix the vertex 2-intervals on track 1 and the edge 2-intervals on track 2.

Vertex 2-Intervals



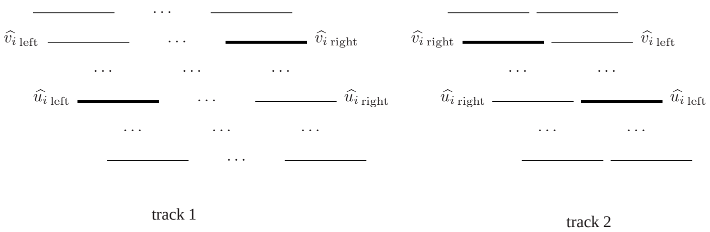
To fix the vertex 2-intervals in the vertex gadget for the vertices V_i with color i, we replace each 2-interval $\widehat{u_i}$ by two 2-track intervals $\widehat{u_i}_{\text{left}}$ and $\widehat{u_i}_{\text{right}}$.

Vertex 2-Intervals



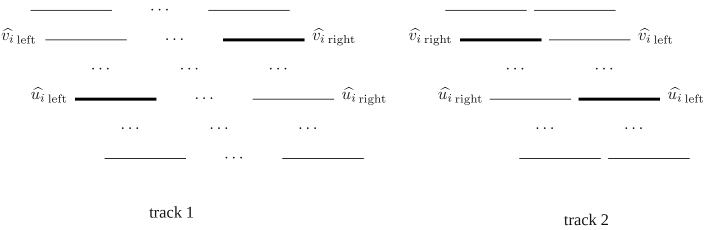
On track 1, let the intervals of $\widehat{u_i}_{\text{ left}}$ and $\widehat{u_i}_{\text{ right}}$ be the left and the right intervals, respectively, of $\widehat{u_i}$.

Vertex 2-Intervals



On track 2, put the intervals of $\widehat{u_i}_{\text{ left}}$ and $\widehat{u_i}_{\text{ right}}$ for all $u \in V_i$ in a separate region, and arrange them in a parallelogram formation with $|V_i|$ rows and 2 columns: $\widehat{u_i}_{\text{ left}}$ in the right column, $\widehat{u_i}_{\text{ right}}$ in the left column.

Vertex 2-Intervals



As usual, the intervals are disjoint in each row and are pairwise intersecting in each column, moreover the columns are slanted such that each interval in the left column intersects all intervals in higher rows in the right column.

Edge 2-Intervals

In a similar way (with the roles of track 1 and track 2 reversed), we replace each edge 2-interval $\widehat{u_i v_j}$ by two 2-track intervals $\widehat{u_i v_j}_{\text{left}}$ and $\widehat{u_i v_j}_{\text{right}}$.

$$\begin{aligned}\mathcal{F} = &\left\{\widehat{u_i} \mid u \in V_i,\, 1 \le i \le k\right\}\\ &\quad \cup \left\{\widehat{u_i v_j}, \overleftarrow{u_i v_j}, \overrightarrow{u_i v_j} \mid uv \in E_{ij},\, 1 \le i < j \le k\right\}.\end{aligned}$$

$$\Downarrow$$

$$\begin{aligned}\mathcal{F} = &\left\{\widehat{u}_{i\,\text{left}}, \widehat{u}_{i\,\text{right}} \mid u \in V_i,\, 1 \le i \le k\right\}\\ &\quad \cup \left\{\widehat{u_i v_j}_{\text{left}}, \widehat{u_i v_j}_{\text{right}}, \overleftarrow{u_i v_j}, \overrightarrow{u_i v_j} \mid uv \in E_{ij},\, 1 \le i < j \le k\right\}.\end{aligned}$$

The Claim

Set the parameter $k' = k + 3\binom{k}{2}$.

$$\Downarrow$$

Set the parameter $k' = 2k + 4\binom{k}{2}$.

G has a k-multicolored clique if and only if $\mathcal{F}$ has a k'-independent set.

Direct Implication

If $K \subseteq V(G)$ is a k-multicolored clique, then the following subset of 2-track intervals is a k'-independent set in $\mathcal{F}$:

$$\left\{\widehat{u_i} \mid u \in K,\, i = \kappa(u)\right\}$$
$$\cup \left\{\widehat{u_i v_j}, \overleftarrow{u_i v_j}, \overrightarrow{u_i v_j} \mid u, v \in K,\, i = \kappa(u),\, j = \kappa(v)\right\}.$$

$$\Downarrow$$

$$\left\{\widehat{u_i}_{\,\mathrm{left}}, \widehat{u_i}_{\,\mathrm{right}} \mid u \in K,\, i = \kappa(u)\right\}$$
$$\cup \left\{\widehat{u_i v_j}_{\,\mathrm{left}}, \widehat{u_i v_j}_{\,\mathrm{right}}, \overleftarrow{u_i v_j}, \overrightarrow{u_i v_j} \mid u, v \in K,\, i = \kappa(u),\, j = \kappa(v)\right\}.$$

Reverse Implication

Suppose $\mathcal{I}$ is a k'-independent set in $\mathcal{F}$.

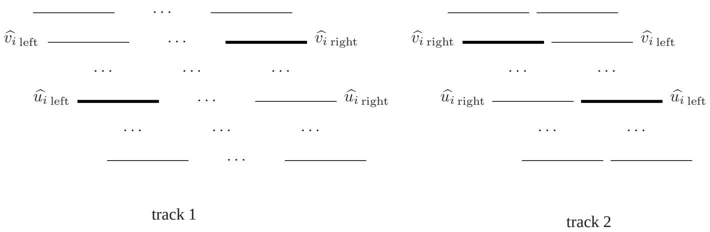
1. The same argument as before shows that $\mathcal{I}$ must include exactly *two* vertex 2-track intervals for each color, and exactly *two* edge 2-track intervals plus two validation 2-track intervals for each pair of distinct colors.

2. We can assume that the two vertex 2-track intervals for each color i form a pair $\widehat{u_i}_{\text{left}}\ \widehat{u_i}_{\text{right}}$ for the same vertex u.

3. Similarly, we can assume that the two edge 2-track intervals for each pair of colors ij form a pair $\widehat{u_i v_j}_{\text{left}}\ \widehat{u_i v_j}_{\text{right}}$ for the same edge uv.

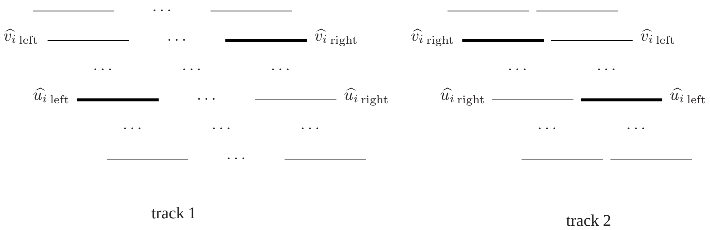
4. Then the same argument as before completes the proof.

Why Can We Assume That?



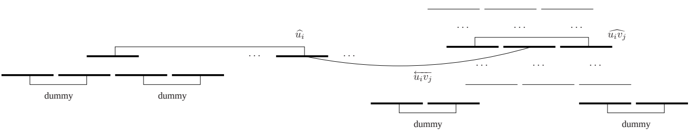
Let $\widehat{u}_{i\,\mathrm{left}}$ and $\widehat{v}_{i\,\mathrm{right}}$ be the two vertex 2-track intervals in $\mathcal{I}$ for some color i. The intersection pattern of the vertex 2-track intervals for V_i on track 2 ensures that the row of u must not be higher than the row of v.

Why Can We Assume That?



Without loss of generality, we can assume that they are in the same row, i.e., $u = v$, so that the set of validation intervals in the middle columns on track 1 that are dominated by $\widehat{u_i}_{\text{ left}}\, \widehat{v_i}_{\text{ right}}$ is minimal (or, in terms of geometric packing, this gives the container $\widehat{u_i}_{\text{ left}}\, \widehat{v_i}_{\text{ right}}$ the largest capacity on track 1).

Previous Reduction for k-Dominating Set

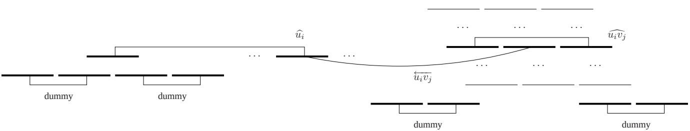


Let (G, κ, k) be an instance of k-MULTICOLORED CLIQUE.

The reduction again constructs k vertex gadgets, one for each color, and $\binom{k}{2}$ edge gadgets, one for each pair of distinct colors.

The vertex gadgets and the edge gadgets are then linked together by the validation gadget.

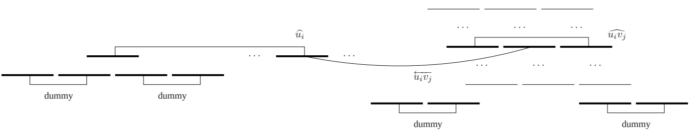
Vertex Selection



For each color i, $1 \leq i \leq k$, let V_i be the set of vertices with color i.

The vertex gadget for the color i includes one interval $\overline{*_i}$ for the color i and one interval $\overline{u_i}$ for each vertex $u \in V_i$.

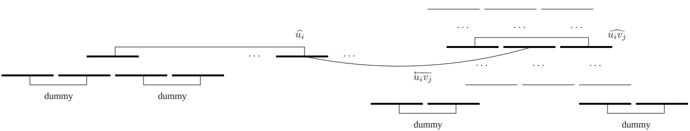
Vertex Selection



The interval $\overline{*_i}$ is combined with each interval $\overline{u_i}$ to form a *vertex 2-interval* $\widehat{u_i}$.

The vertex gadget for V_i also includes two disjoint dummy 2-intervals that contain the left and the right endpoints, respectively, of the interval $\overline{*_i}$.

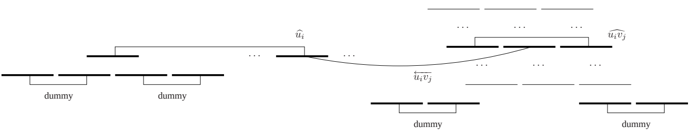
Edge Selection



For each pair of distinct colors i and j, $1 \leq i < j \leq k$, let E_{ij} be the set of edges uv such that u has color i and v has color j.

The edge gadget for the pair of colors ij includes a group of intervals that can viewed as a table with $|E_{ij}|$ rows and 3 columns. Again the intervals in the table are arranged in a parallelogram formation.

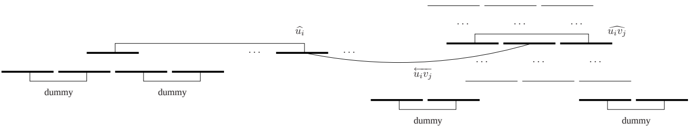
Edge Selection



Each row of the table corresponds to a distinct edge $uv \in E_{ij}$: the left interval and the right interval together form an *edge 2-interval* $\widehat{u_i v_j}$; the middle interval, denoted by $\overline{u_i v_j}$, is used for validation.

The edge gadget for E_{ij} also includes two disjoint dummy 2-intervals that intersect the left intervals and the right intervals, respectively, of all edge 2-intervals $\widehat{u_i v_j}$.

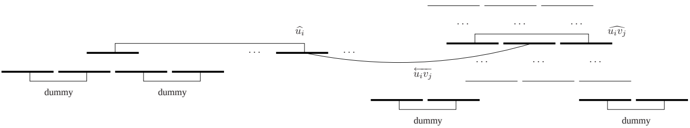
Validation



For each edge uv such that u has color i and v has color j, the validation gadget includes two *validation 2-intervals* $\overleftarrow{u_iv_j}$ and $\overrightarrow{u_iv_j}$:

- the 2-interval $\overleftarrow{u_iv_j}$ consists of the interval $\overline{u_iv_j}$ and the interval $\overline{u_i}$;
- the 2-interval $\overrightarrow{u_iv_j}$ consists of the interval $\overline{u_iv_j}$ and the interval $\overline{v_j}$.

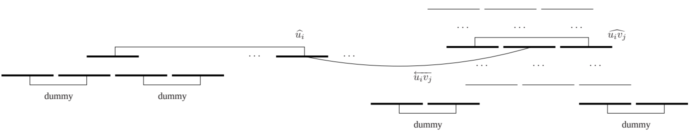
Put Together



$$\begin{aligned}\mathcal{F} = &\Big\{\widehat{u_i} \mid u \in V_i,\, 1 \le i \le k\Big\} \\ &\cup \Big\{\widehat{u_iv_j}, \overleftarrow{u_iv_j}, \overrightarrow{u_iv_j} \mid uv \in E_{ij},\, 1 \le i < j \le k\Big\} \\ &\cup \textit{DUMMIES},\end{aligned}$$

where *DUMMIES* is the set of $2k + 2\binom{k}{2}$ dummy 2-intervals, two in each vertex or edge gadget.

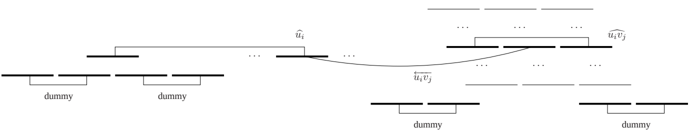
The Claim



Set the parameter $k' = k + \binom{k}{2}$.

G has a k-multicolored clique if and only if $\mathcal{F}$ has a k'-dominating set.

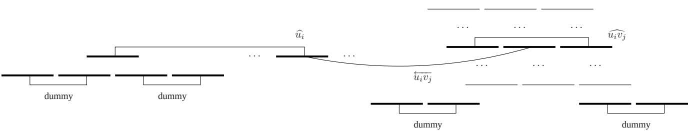
Direct Implication



If $K \subseteq V(G)$ is a k-multicolored clique, then the following subset of 2-intervals is a k'-dominating set in $\mathcal{F}$:

$$\left\{\widehat{u_i} \mid u \in K,\ i = \kappa(u)\right\}$$
$$\cup \left\{\widehat{u_iv_j} \mid u, v \in K,\ i = \kappa(u),\ j = \kappa(v)\right\}.$$

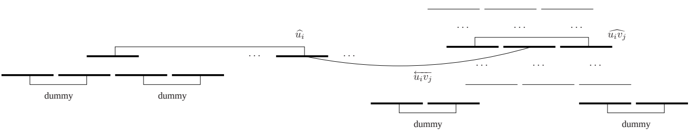
Reverse Implication



Suppose that $\mathcal{I}$ is a k'-dominating set in $\mathcal{F}$.

Because every dummy 2-interval can be replaced by an adjacent vertex or edge 2-interval in a dominating set, we can assume without loss of generality that $\mathcal{I}$ does not include any dummy 2-intervals.

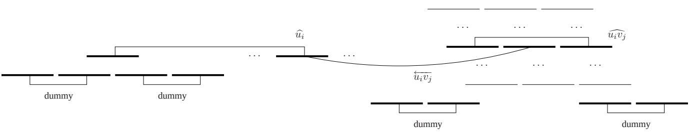
Reverse Implication



Then, to dominate the dummy 2-intervals, $\mathcal{I}$ must include at least one vertex 2-interval for each color, and at least one edge 2-interval for each pair of distinct colors.

Since $k' = k + \binom{k}{2}$, $\mathcal{I}$ must include exactly one vertex 2-interval for each color, and exactly one edge 2-interval for each pair of distinct colors.

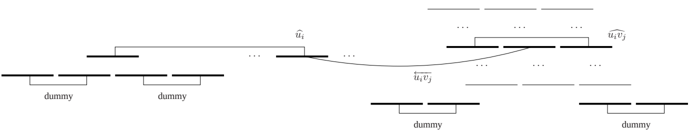
Reverse Implication



It follows that for each pair of distinct colors ij, the two validation 2-intervals $\overleftarrow{u_iv_j}$ and $\overrightarrow{u_iv_j}$ must be dominated by the two vertex 2-intervals $\widehat{u_i}$ and $\widehat{v_j}$, respectively.

Therefore the vertex selection and the edge selection are consistent: the k vertex 2-intervals in $\mathcal{I}$ corresponds to a k-multicolored clique in G.

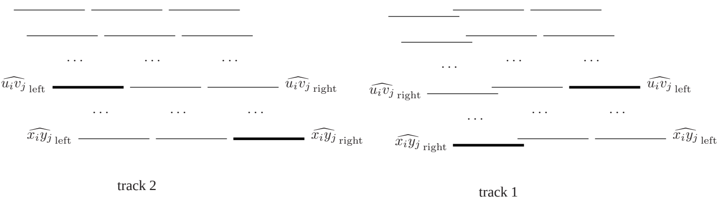
New Reduction



We now modify the previous construction to transform each 2-interval into a 2-track interval. To transform the vertex 2-intervals into 2-track intervals, move the intervals $\overline{u_i}$ to track 1, and move the intervals $\overline{*_i}$ to track 2. Then, to transform the validation 2-intervals into 2-track intervals, move all edge gadgets to track 2. The dummy 2-intervals can be fixed accordingly.

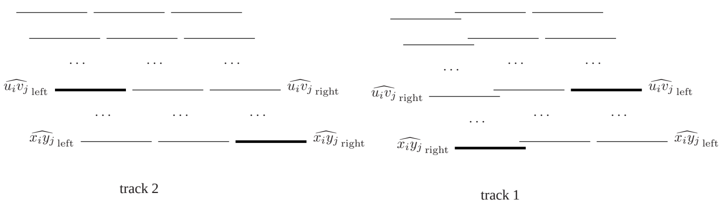
It remains to fix the edge 2-intervals now on track 2.

Edge 2-Intervals



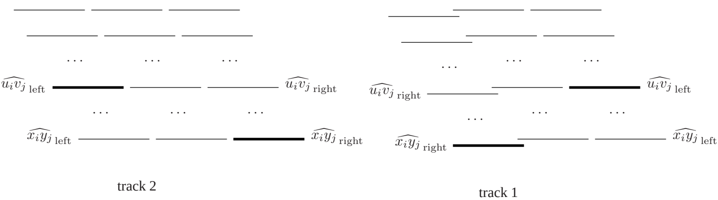
To fix the edge 2-intervals in the edge gadget for the edges E_{ij} with colors ij, we replace each 2-interval $\widehat{u_iv_j}$ by two 2-track intervals $\widehat{u_iv_j}_{\text{ left}}$ and $\widehat{u_iv_j}_{\text{ right}}$.

Edge 2-Intervals



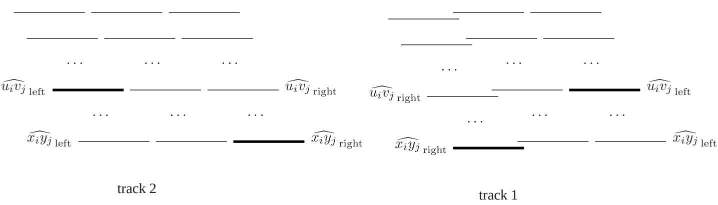
On track 2, let the intervals of $\widehat{u_iv_j}_{\text{ left}}$ and $\widehat{u_iv_j}_{\text{ right}}$ be the left and the right intervals, respectively, of $\widehat{u_iv_j}$.

Edge 2-Intervals



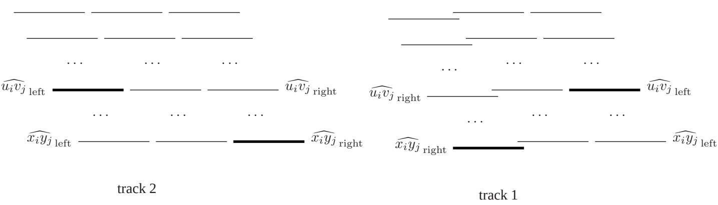
On track 1, put the intervals of $\widehat{u_iv_j}_{\text{ left}}$ and $\widehat{u_iv_j}_{\text{ right}}$ for all $uv \in E_{ij}$ in a separate region, then arrange them, together with $|E_{ij}|$ additional dummy intervals, in a parallelogram formation with $|E_{ij}|$ rows and 3 columns: $\widehat{u_i}_{\text{ left}}$ in the right column, $\widehat{u_i}_{\text{ right}}$ in the left column, and dummies in the middle column.

Edge 2-Intervals



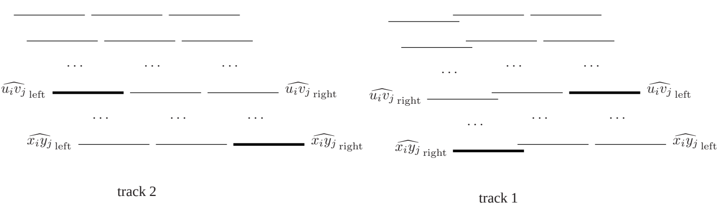
As usual, the intervals are pairwise intersecting in each column, and the columns are slanted. But in each row the three intervals are not all disjoint: the left interval and the middle interval slightly overlap, and are both disjoint from the right interval.

Edge 2-Intervals



Now each interval in the right column intersects all intervals in lower rows in the middle column, and each interval in the left column intersects all intervals in the same or higher rows in the middle column.

Edge 2-Intervals



Finally, each of the $|E_{ij}|$ dummy intervals in the middle column is combined with an isolated dummy interval on track 2 to form a dummy 2-track interval.

Put Together

$$\begin{aligned}\mathcal{F} =& \left\{ \widehat{u_i} \mid u \in V_i,\, 1 \le i \le k \right\} \\ & \cup \left\{ \widehat{u_i v_j}, \overleftarrow{u_i v_j}, \overrightarrow{u_i v_j} \mid uv \in E_{ij},\, 1 \le i < j \le k \right\} \\ & \cup \mathit{DUMMIES},\end{aligned}$$

where *DUMMIES* is the set of $2k + 2\binom{k}{2}$ dummy 2-intervals, two in each vertex or edge gadget.

$$\Downarrow$$

$$
\begin{aligned}
\mathcal{F} = & \left\{ \widehat{u_i} \mid u \in V_i,\ 1 \le i \le k \right\} \\
& \cup \left\{ \widehat{u_i v_j}_{\text{ left}}, \widehat{u_i v_j}_{\text{ right}}, \overleftarrow{u_i v_j}, \overrightarrow{u_i v_j} \mid uv \in E_{ij},\ 1 \le i < j \le k \right\} \\
& \cup \textit{DUMMIES},
\end{aligned}
$$

where *DUMMIES* is the set of $2k + 2\binom{k}{2} + |E(G)|$ dummy 2-track intervals, two in each vertex or edge gadget as before, and one more for each edge (recall the middle column of each edge gadget on track 1).

The Claim

Set the parameter $k' = k + \binom{k}{2}$.

$$\Downarrow$$

Set the parameter $k' = k + 2\binom{k}{2}$.

G has a k-multicolored clique if and only if $\mathcal{F}$ has a k'-dominating set.

Direct Implication

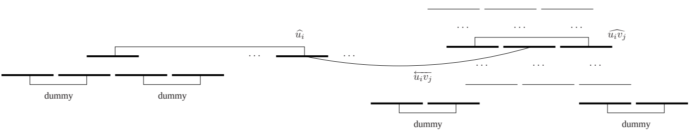
If $K \subseteq V(G)$ is a k-multicolored clique, then the following subset of 2-track intervals is a k'-dominating set in $\mathcal{F}$:

$$\left\{\widehat{u_i} \mid u \in K,\, i = \kappa(u)\right\}$$
$$\cup \left\{\widehat{u_i v_j} \mid u, v \in K,\, i = \kappa(u),\, j = \kappa(v)\right\}.$$

$$\Downarrow$$

$$\left\{\widehat{u_i} \mid u \in K,\, i = \kappa(u)\right\}$$
$$\cup \left\{\widehat{u_i v_j}_{\,\mathrm{left}}, \widehat{u_i v_j}_{\,\mathrm{right}} \mid u, v \in K,\, i = \kappa(u),\, j = \kappa(v)\right\}.$$

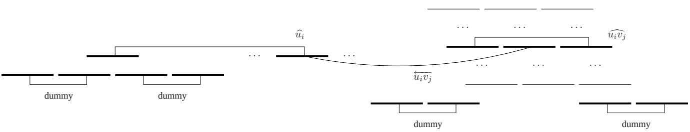
Reverse Implication



Suppose that $\mathcal{I}$ is a k'-dominating set in $\mathcal{F}$.

Note that any one of the (original) two dummy 2-track intervals in each vertex or edge gadget can be replaced by an adjacent vertex or edge 2-interval in a dominating set. Thus we can assume without loss of generality that $\mathcal{I}$ includes none of these $2k + 2\binom{k}{2}$ dummies.

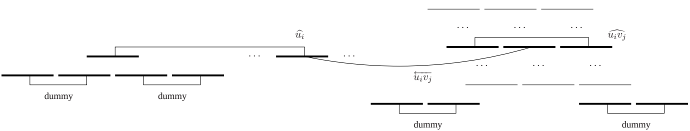
Reverse Implication



Then, to dominate these dummies, $\mathcal{I}$ must include at least one vertex 2-track interval for each color, and at least two edge 2-track intervals for each pair of distinct colors.

Since $k' = k+2\binom{k}{2}$, $\mathcal{I}$ must include exactly one vertex 2-track interval for each color, and exactly two edge 2-track intervals for each pair of distinct colors.

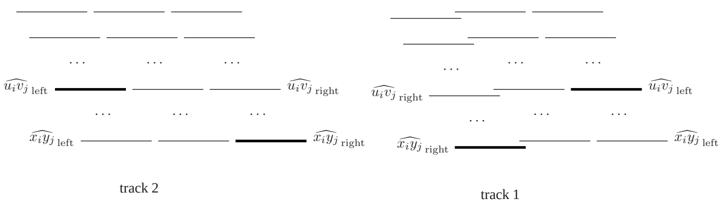
Reverse Implication



We can assume that the two edge 2-track intervals for each pair of colors ij form a pair $\widehat{u_i v_j}_{\,\mathrm{left}}\, \widehat{u_i v_j}_{\,\mathrm{right}}$ for the same edge uv.

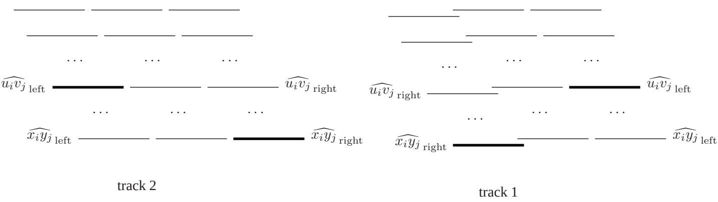
Then the same argument as before shows that the k vertex 2-track intervals in $\mathcal{I}$ corresponds to a k-multicolored clique in G.

Why Can We Assume That?



Let $\widehat{u_iv_j}_{\text{left}}$ and $\widehat{x_iy_j}_{\text{right}}$ be the two edge 2-track intervals in $\mathcal{I}$ for some pair of colors ij. The intersection pattern of the edge 2-track intervals for E_{ij} on track 1 ensures that, in order to dominate all the (new) dummies in the middle column, the row of xy must not be higher than the row of uv.

Why Can We Assume That?



Without loss of generality, we can assume that they are in the same row, i.e., $uv = xy$, so that the set of validation intervals in the middle column on track 2 that are dominated by $\widehat{u_iv_j}_{\text{ left}}\widehat{x_iy_j}_{\text{ right}}$ is maximal.

k-Clique

Fellows et al. presented an FPT algorithm for k-CLIQUE in t-interval graphs parameterized by both k and t. They estimated that the running time of their algorithm is $t^{O(k \log k)} \cdot \mathrm{poly}(n)$, where n is the number of vertices in the graph, and asked whether the parametric time-bound can be improved.

We show:

Theorem 2. *For any constant $c \geq 3$, there is an algorithm for k-CLIQUE in t-interval graphs with running time $O(t^{ck}) \cdot O(n^c)$ if $k \leq \frac{1}{4} \cdot n^{1-1/c}$, where n is the number of vertices in the graph. In particular, there is an FPT algorithm for k-CLIQUE in t-interval graphs with running time $\max\{t^{O(k)}, 2^{O(k \log k)}\} \cdot \mathrm{poly}(n)$.*

Previous Algorithm

Fellows et al. presented the following algorithm $\textsc{Clique}(G, k)$ that decides whether a given t-interval graph G has a k-clique:

$\textsc{Clique}(G, k)$:

1. If $|V(G)| < k$, then return NO.
2. Let v be a vertex of minimum degree in G.
3. If $\deg(v) \geq 2tk$, then return YES.
4. If v is in a k-clique of G, then return YES.
5. Return $\textsc{Clique}(G - v, k)$.

The crucial step of this algorithm, step 3, is justified by a structural lemma that “if G is a t-interval graph with no k-cliques then G has a vertex of degree less than $2tk$.”

Previous Analysis

Step 4 can be implemented in $O(k^2 \cdot \binom{2tk}{k})$ time by brute force; all other steps have running time polynomial in n. The total number of recursive calls, in step 5, is at most n. The overall time complexity of the algorithm is

$$O(k^2 \cdot \binom{2tk}{k}) \cdot \mathrm{poly}(n).$$

Fellows et al. estimated that

$$O(k^2 \cdot \binom{2tk}{k}) = t^{O(k \log k)}. \tag{1}$$

Improved Algorithm

Our FPT algorithm has two components:

1. The first component is an algorithm CLIQUE*(G,k) slightly modified from CLIQUE(G,k).

2. The second component is the obvious brute-force algorithm that enumerates and checks all k-subsets of vertices for k-cliques.

CLIQUE*

CLIQUE*(G, k):

1. If $|V(G)| < k$, then return NO.
2. Let v be a vertex of minimum degree in G.
3. If $\deg(v) \geq 2tk$, then return YES.
4. If CLIQUE*$(\text{neighbors}(v), k-1)$ returns YES, then return YES.
5. Return CLIQUE*$(G - v, k)$.

CLIQUE*(G, k) is identical to CLIQUE(G, k) except step 4. The following recurrence on the time bound $f(k) \cdot g(n)$ captures the recursive behavior of CLIQUE*(G, k):

$$f(k) \cdot g(n) \leq f(k-1) \cdot g(2tk) + f(k) \cdot g(n-1) + O(n^2).$$

Two Components Put Together

Lemma 1. *For any constant $c \geq 3$, if $k \leq \frac{1}{4} \cdot n^{1-1/c}$, then the running time of* $\textsc{Clique}^*(G, k)$ *is* $O(t^{ck}) \cdot O(n^c)$.

Lemma 2. *For any constant $c \geq 3$, if $k > \frac{1}{4} \cdot n^{1-1/c}$, then the running time of the brute-force algorithm is* $2^{O(k \log k)}$.

Finally, for any constant $c \geq 3$, by choosing the algorithm $\textsc{Clique}^*(G, k)$ when $k \leq \frac{1}{4} \cdot n^{1-1/c}$, and choosing the brute-force algorithm when $k > \frac{1}{4} \cdot n^{1-1/c}$, we obtain an FPT algorithm with a parametric time-bound of

$$\max\{t^{O(k)}, 2^{O(k \log k)}\}. \tag{2}$$

Really?

Compare our bound (2) with the previous bound (1). It appears that we have obtained an improvement[1], but asymptotically this improvement is negligible. Check that the estimate in (1) is not tight:

$$O(k^2 \cdot \binom{2tk}{k}) = O(k^2(2tk)^k) = t^{O(k)}2^{O(k \log k)}$$
$$= \max\{(t^{O(k)})^2, (2^{O(k \log k)})^2\} = \max\{t^{O(k)}, 2^{O(k \log k)}\}.$$

[1]Under the condition that $k \leq \frac{1}{4} \cdot n^{1-1/c}$ for some constant $c \geq 3$, $\textsc{Clique}^*(G,k)$ clearly improves $\textsc{Clique}(G,k)$: in particular, for $t = \Theta(\log k)$, the parametric bound of $\textsc{Clique}^*(G,k)$ is $2^{O(k \log\log k)}$, and the parametric bound of $\textsc{Clique}(G,k)$ is $2^{O(k \log k)}$.

Open Question Reformulated

In light of this delicate distinction, perhaps the open question on k-CLIQUE in t-interval graphs could be stated more precisely as follows:

Question 1. *Is there an FPT algorithm for k-*CLIQUE *in t-interval graphs with a parametric time-bound of $t^{O(k)}$?*

Note that a parametric time-bound of $2^{O(k \log k)}$ alone is beyond reach. This is because every graph of n vertices is a t-interval graph for $t \geq n/4$. If the parameter t does not appear in the bound, then we would have an FPT algorithm for the W[1]-hard problem of k-CLIQUE in general graphs.

Comparative Genomics

A *genomic map* is a sequence of gene markers.

A *gene marker* appears in a genomic map in either positive or negative orientation.

In comparative genomics, the first step of sequence analysis is usually to decompose two or more genomes into *syntenic blocks* that are segments of homologous chromosomes.

For the reliable recovery of syntenic blocks, *noise and ambiguities* in the genomic maps need to be removed first.

Maximal Strip Recovery

Given d genomic maps as signed sequences of gene markers, MAXIMAL STRIP RECOVERY (MSR-d) is the problem of finding d subsequences, one subsequence of each genomic map, such that the total length ℓ of the maximal strips in these subsequences is maximized.

A *strip* is a string of at least two markers such that either the string itself or its signed reversal appears contiguously as a substring in each of the d subsequences in the solution.

An Example

The two genomic maps (the markers in negative orientation are underlined)

$$\begin{array}{cccccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\
\underline{8} & \underline{5} & \underline{7} & \underline{6} & 4 & 1 & 3 & 2 & \underline{12} & \underline{11} & \underline{10} & 9
\end{array}$$

have two subsequences

$$\begin{array}{ccccccccccccc}
1 & 3 & & \quad & 6 & 7 & 8 & \quad & 10 & 11 & 12 \\
\underline{8} & \underline{7} & \underline{6} & \quad & & 1 & 3 & \quad & \underline{12} & \underline{11} & \underline{10}
\end{array}$$

of the maximum total strip length 8.

Our Result

MSR-d admits a polynomial-time $2d$-approximation and is NP-hard to approximate within $\Omega(d/\log d)$ (Jiang 2010).

Our following theorem gives the first parameterized intractability result for MSR-d:

Theorem 3. MSR-d *for any constant* $d \geq 4$ *is* W[1]-*hard when the parameter is either the total length of the strips, or the total number of adjacencies in the strips, or the number of strips in the optimal solution. This holds even if all gene markers are distinct and appear in positive orientation in each genomic map.*

FPT- Reduction

Let ℓ-MSR-d be the problem MSR-d parameterized by the total length ℓ of the strips in the solution. We prove that ℓ-MSR-4 is W[1]-hard by an FPT-reduction from k-INDEPENDENT SET in 2-track interval graphs.

Let $(\mathcal{F}, k)$ be an instance of k-INDEPENDENT SET in 2-track interval graphs, where $\mathcal{F} = \{I_1, \ldots, I_n\}$ is a set of n 2-track intervals.

We construct four genomic maps $G_{\rightarrow}, G_{\leftarrow}, G_1, G_2$, where each map is a permutation of $2n$ distinct markers all in positive orientation:

$$\overset{i}{\subset} \quad \text{and} \quad \overset{i}{\supset}, \quad 1 \leq i \leq n.$$

$G_{\rightarrow}$ and $G_{\leftarrow}$

$G_{\rightarrow}$ and $G_{\leftarrow}$ are concatenations of the n pairs of markers with ascending and descending indices, respectively, and ensure that each strip must be a pair of markers:

$$
\begin{array}{lcccccc}
G_{\rightarrow}: & \overset{1}{\subset} & \overset{1}{\supset} & \cdots & \overset{n}{\subset} & \overset{n}{\supset} \\
G_{\leftarrow}: & \overset{n}{\subset} & \overset{n}{\supset} & \cdots & \overset{1}{\subset} & \overset{1}{\supset}
\end{array}
$$

G_1 and G_2

G_1 and G_2 encode the intersection pattern of the 2-track intervals by pairs of markers:

1. Modify the representation of the 2-track interval graph for $\mathcal{F}$ until the $2n$ endpoints of the n intervals on each track are all distinct.

2. On each track, mark the left and the right endpoints of the interval for I_i by the left and the right markers $\overset{i}{\subset}$ and $\overset{i}{\supset}$, respectively. Thus we obtain two sequences of markers.

Set the parameter $\ell = 2k$. Then $\mathcal{F}$ has a k-independent set if and only $G_{\rightarrow}, G_{\leftarrow}, G_1, G_2$ have four subsequences of total strip length ℓ.

Summary

Theorem 1. k-INDEPENDENT SET *and* k-DOMINATING SET *in unit 2-track interval graphs are* W[1]*-hard.*

Theorem 2. *For any constant $c \geq 3$, there is an algorithm for k-*CLIQUE *in t-interval graphs with running time $O(t^{ck}) \cdot O(n^c)$ if $k \leq \frac{1}{4} \cdot n^{1-1/c}$, where n is the number of vertices in the graph. In particular, there is an FPT algorithm for k-*CLIQUE *in t-interval graphs with running time $\max\{t^{O(k)}, 2^{O(k \log k)}\} \cdot \mathrm{poly}(n)$.*

Theorem 3. MSR-d *for any constant $d \geq 4$ is* W[1]*-hard when the parameter is either the total length of the strips, or the total number of adjacencies in the strips, or the number of strips in the optimal solution. This holds even if all gene markers are distinct and appear in positive orientation in each genomic map.*