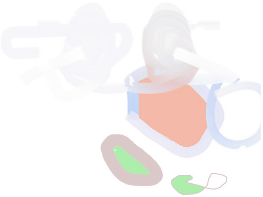
WL50 6 July 2018

Coherent Configurations and Graph Isomorphism:

The emergence of the Johnson graphs

László Babai
University of Chicago



Graph isomorphism

testing isomorphism of graphs with n vertices

moderately exponential $\exp(\sqrt{n \log n})$ Luks 1983

quasipolynomial $\exp((\log n)^c)$ B 2015+

Graph isomorphism

testing isomorphism of graphs with n vertices

moderately exponential $\exp(\sqrt{n \log n})$ Luks 1983

quasipolynomial $\exp((\log n)^c)$ B 2015+

divide-and-conquer algorithm

reduces instance of size n to $q(n)$ instances of
significantly smaller size ($\leq 0.9n$)

$q(n)$ – **multiplicative cost** (branching factor)

Graph isomorphism

testing isomorphism of graphs with n vertices

moderately exponential $\exp(\sqrt{n \log n})$ Luks 1983

quasipolynomial $\exp((\log n)^c)$ B 2015+

divide-and-conquer algorithm

reduces instance of size n to $q(n)$ instances of significantly smaller size ($\leq 0.9n$)

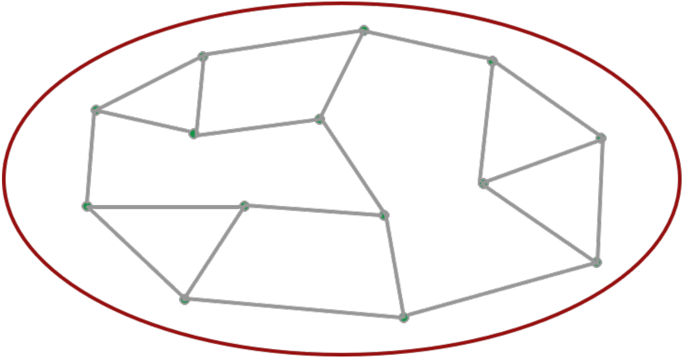
$q(n)$ – **multiplicative cost** (branching factor)

cost analysis: $t(n) \leq q(n)t(0.9n)$

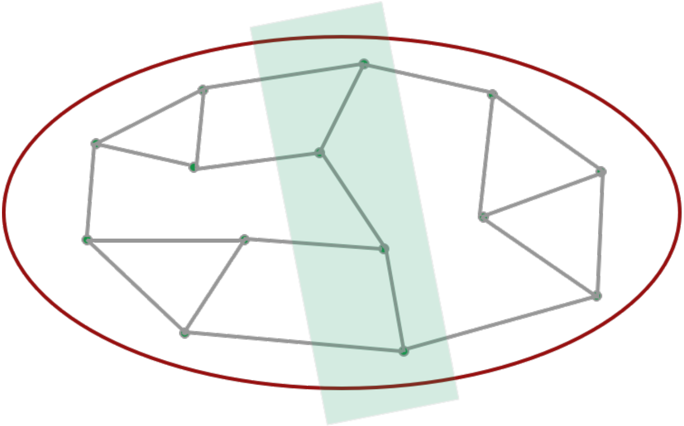
$$\to t(n) \leq q(n)^{O(\log n)}$$

suffices to keep $q(n)$ quasipolynomial

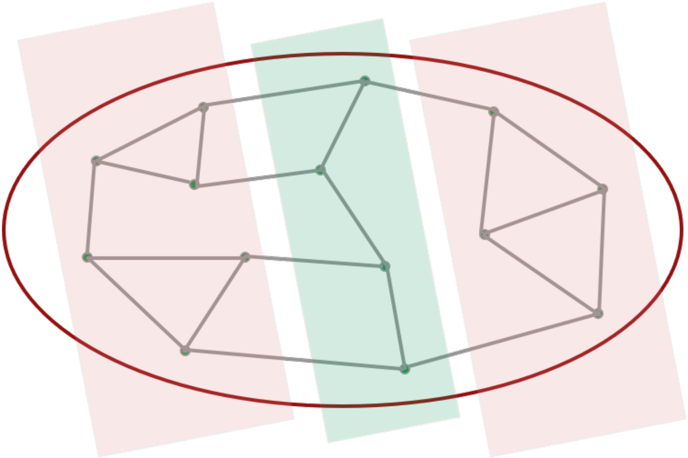
Exposing irregularity: canonical coloring



Exposing irregularity: canonical coloring



Exposing irregularity: canonical coloring



Exposing irregularity: canonical coloring

Good coloring: no dominant color (90% of vertices)

Good equipartition: equipartition of dominant color
nontrivial partition into equal parts

Canonicity: preserved under isomorphisms

Exposing irregularity: canonical coloring

Good coloring: no dominant color (90% of vertices)

Good equipartition: equipartition of dominant color
nontrivial partition into equal parts

Canonicity: preserved under isomorphisms

Why not consider
canonical partitions into unequal parts?

Exposing irregularity: canonical coloring

Good coloring: no dominant color (90% of vertices)

Good equipartition: equipartition of dominant color
nontrivial partition into equal parts

Canonicity: preserved under isomorphisms

Why not consider
canonical partitions into <u>unequal parts</u>?

$\rightarrow$ canonical coloring by size of block
each color equipartitioned

Exposing irregularity: canonical coloring

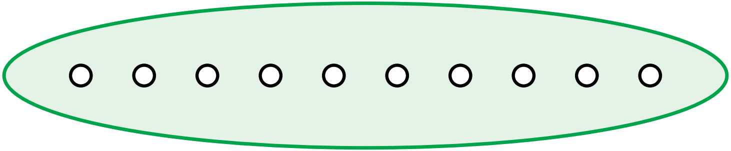
Good coloring: no dominant color (90% of vertices)

Good equipartition: equipartition of dominant color
nontrivial partition into equal parts

Canonicity: preserved under isomorphisms

Why not consider
canonical partitions into unequal parts?

→ canonical coloring by size of block
each color equipartitioned



Exposing irregularity: canonical coloring

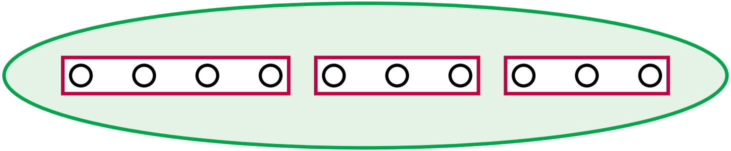
Good coloring: no dominant color (90% of vertices)

Good equipartition: equipartition of dominant color
nontrivial partition into equal parts

Canonicity: preserved under isomorphisms

Why not consider
canonical partitions into unequal parts?

→ canonical coloring by size of block
each color equipartitioned



Exposing irregularity: canonical coloring

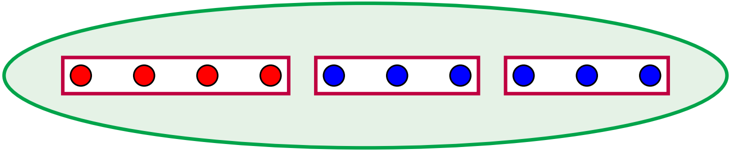
Good coloring: no dominant color (90% of vertices)

Good equipartition: equipartition of dominant color
nontrivial partition into equal parts

Canonicity: preserved under isomorphisms

Why not consider
canonical partitions into unequal parts?

→ canonical coloring by size of block
each color equipartitioned



Canonical assignment

Assignment $\mathbf{x} \mapsto F(\mathbf{x})$ structures

E.g. $\mathbf{x}$ – graph, $F(\mathbf{x})$ – coloring of vertices

F **canonical** if it also assigns
isomorphism $\mapsto$ isomorphism

$$\begin{tikzcd}
\mathbf{x} \arrow[r, "\sigma"] \arrow[d, Rightarrow] & \mathbf{y} \arrow[d, Rightarrow] \\
F(\mathbf{x}) & F(\mathbf{y})
\end{tikzcd}$$

Canonical assignment

Assignment $\quad \mathbf{x} \mapsto F(\mathbf{x})$ structures

E.g. $\quad \mathbf{x}$ – graph, $\quad F(\mathbf{x})$ – coloring of vertices

F **canonical** if it also assigns
isomorphism $\mapsto$ isomorphism

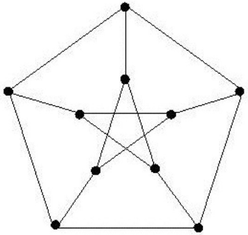
$$\begin{tikzcd} \mathbf{x} \arrow[r, "\sigma"] \arrow[d, Rightarrow] & \mathbf{y} \arrow[d, Rightarrow] \\ F(\mathbf{x}) \arrow[r, "F(\sigma)"'] & F(\mathbf{y}) \end{tikzcd}$$

FUNCTOR between categories of isomorphisms

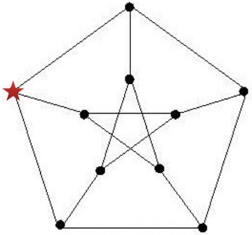
$$F(\sigma\tau) = F(\sigma)F(\tau)$$

e.g., $\quad F : \text{Graphs} \to \text{ColoredSets}$

Individualization/refinement: creating irregularity

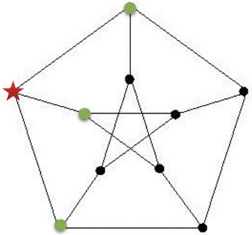


Individualization/refinement: creating irregularity



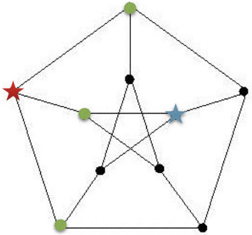
individualize vertex

Individualization/refinement: creating irregularity



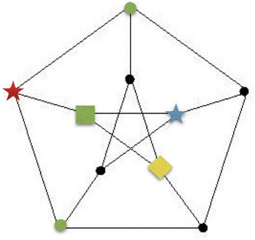
individualize vertex, refine

Individualization/refinement: creating irregularity



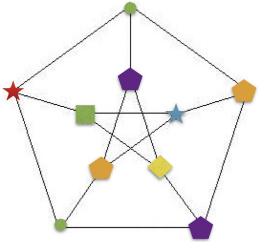
individualize vertex, refine
individualize second vertex

Individualization/refinement: creating irregularity



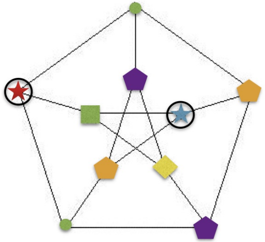
individualize vertex, refine
individualize second vertex, refine

Individualization/refinement: creating irregularity



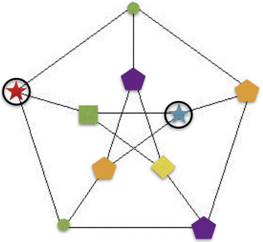
individualize vertex, refine
individualize second vertex, refine

Individualization/refinement: creating irregularity



individualize vertex, refine
individualize second vertex, refine

Individualization/refinement: creating irregularity



individualize vertex, refine
individualize second vertex, refine
multiplicative cost of individualization
of k vertices: $\mathbf{n^k}$

Cost of good partitioning

Given a nontrivial regular graph
can we find a canonical **good coloring** or **equipartition**
at **modest multiplicative cost?**

Cost of good partitioning

Given a nontrivial regular graph
can we find a canonical **good coloring** or **equipartition**
at **modest multiplicative cost?**

NO! **Johnson graphs**
resilient to good coloring/partition

DEF: $J(k,t)$ **Johnson graph** $t \geq 1 \quad k \geq 2t+1$

vertex set $V = \{v_T \mid T \subseteq \Delta, |T| = t\}$ where $|\Delta| = k$

$$|V| = \binom{k}{t}$$

adjacency: $v_T \sim v_S \iff |T \setminus S| = 1$

---

multiplicative cost of good coloring/partition $\exp(\Omega(k/t))$

important case: $t = 2$ so $n = \Theta(k^2)$ cost $\exp(\Omega(\sqrt{n}))$
Graph Isomorphism **bottleneck** for three decades

Split-or-Johnson

Johnson graphs are the *only* obstruction
to good partitioning

Split-or-Johnson

Johnson graphs are the *only* obstruction
to good partitioning

Theorem (Split-or-Johnson (abridged))

Given a nontrivial regular graph, at quasipolynomial multiplicative cost one can efficiently find either

(a) *a good canonical vertex-coloring, or*

(b) *a good canonical equipartition, or*

(c) *a canonically embedded Johnson graph on dominant vertex-color class ($\geq 90\%$ of vertices)*

quasipolynomial: $\exp(O((\log n)^c))$

canonicity: functor: isomorphism of input graphs induces isomorphism of embedded structures

efficiently: in quasipolynomial time

Split-or-Johnson, inefficiently

What if we **omit efficiency requirement?**
$\to$ **assume** $\mathrm{Aut}(X)$ **known**

split into orbits; if no orbit dominant, done
let $G := \mathrm{Aut}(X)$ restricted to dominant orbit C, $|C| = m$
$G \neq A_m, S_m$ by symmetry defect
if G t-trans $\implies$ $t = O(\log^2 n / \log\log n)$ (Bochert 1892)
$t = O(\log n)$ (Wielandt 1934)
individualize $t-1$ points $\implies$ G trans, not doubly
if G imprimitive, pick max system of imprimitivity
(minimal domains of imprimitivity)
$\to$ not canonical, but there are $\leq n-1$ of these
b/c blocks containing $x \in \Omega$ are disjoint outside x
$\to$ individualize one of these systems of imprimitivity
remaining case: G **uniprimitive**

Split-or-Johnson, inefficiently

remaining case: G **uniprimitive**
$\rightarrow$ small or Cameron (CFSG)
if small, individualize polylog points $\rightarrow$ fix all
else, if Cameron then socle $(A_r)^s$
if $s \geq 2$ individualize s objects to get imprimitive
else, if $s = 1 \implies$ Johnson

Cameron's classification of large primitive groups
has been used for decades to identify "Luks bottleneck"
for graph isomorphism testing
$\rightarrow$ can be replaced by Split-or-Johnson
$\rightarrow$ eliminates one application of CFSG used to justify quasipolynomial ISO test
other appl of CFSG was eliminated by Pyber

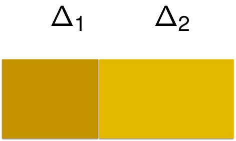
Twins, symmetry defect

$\mathfrak{X} = (\Gamma, \mathcal{R})$ — structure

DEF: $x \neq y \in \Gamma$ **twins** if transposition $(x, y) \in \mathrm{Aut}(\mathfrak{X})$
Fact: "twin or equal" — equivalence relation

DEF: $\Delta \subseteq \Gamma$ **set of twins**: subset of equivalence class
Fact: $\Delta \subseteq \Gamma$ **set of twins** $\iff$ $\mathrm{Sym}(\Delta) \leq \mathrm{Aut}(\mathfrak{X})$
DEF: Symmetricity of $\mathfrak{X}$:
relative size of largest twin equivalence class
DEF: Symmetry defect of $\mathfrak{X}$: $1-$ symmetricity$(\mathfrak{X})$

Example: if $\mathrm{Aut}(\mathfrak{X}) = \mathrm{Sym}(\Delta_1) \times \mathrm{Sym}(\Delta_2)$ where $\Gamma = \Delta_1 \cup \Delta_2$ then the defect of $\mathfrak{X}$ is $\min\{|\Delta_1|, |\Delta_2|\}$



Split-or-Johnson

Exercise. If X is a nontrivial regular graph (not empty, not complete) then defect$(X) \geq 1/2$.

Split-or-Johnson

Exercise. If X is a nontrivial regular graph (not empty, not complete) then $\mathrm{defect}(X) \geq 1/2$.

Theorem (Split-or-Johnson (unabridged))

Given a graph X with $\mathrm{defect}(X) \geq 0.1$, at quasipolynomial multiplicative cost one can efficiently find either

(a) *a good canonical vertex-coloring, or*

(b) *a good canonical equipartition, or*

(c) *a canonically embedded Johnson graph on dominant vertex-color class ($\geq 90\%$ of vertices)*

Why this is the real thing?
Symmetry defect condition obviously **necessary**

Reducing k-ary to binary: Design Lemma

k-ary relation on V: $R \subseteq V^k$
k-ary relational structure: $\mathfrak{X} = (V, \mathcal{R})$ where
$\mathcal{R} = (R_1, \dots, R_m)$ — k-ary relations

Theorem (Design Lemma)

Given a k-ary relational structure $\mathfrak{X}$ with $\mathrm{defect}(\mathfrak{X}) \geq 0.1$, *one can individualize $k-1$ vertices and find, in $n^{O(k)}$ time,*

(a) *a good canonical vertex-coloring, or*
(b) *a good canonical equipartition, or*
(c) *a canonically embedded regular graph on dominant vertex-color class ($\geq 90\%$ of vertices)*

Application to Graph Isomorphism: $k = O(\log n)$

Significance to Graph Isomorphism

Graph Isomorphism can be reduced * via group theory to

Encasement problem

Given a k-ary relational structure $\mathfrak{X} = (V, \mathcal{R})$ with $k = O(\log n)$ (where $n = |V|$) and defect$(\mathfrak{X}) \geq 0.1$, find $G \leq \mathrm{Sym}(V)$ (subgroup of the symmetric group) and $A \subseteq V$ such that

- $|A| = \mathrm{polylog}(n)$
- $\mathrm{Aut}(\mathfrak{X})_{(A)} \leq G$ (pointwise stabilizer of A)
- $|\mathrm{Sym}(V) : G| = \exp(\Omega(n))$

Design lemma & Split-or-Johnson solve this in quasipolynomial time.

Significance to Graph Isomorphism

Graph Isomorphism can be reduced * via group theory to

Encasement problem

Given a k-ary relational structure $\mathfrak{X} = (V, \mathcal{R})$ with $k = O(\log n)$ (where $n = |V|$) and $\text{defect}(\mathfrak{X}) \geq 0.1$, find $G \leq \text{Sym}(V)$ (subgroup of the symmetric group) and $A \subseteq V$ such that

- $|A| = \text{polylog}(n)$
- $\text{Aut}(\mathfrak{X})_{(A)} \leq G$ (pointwise stabilizer of A)
- $|\text{Sym}(V) : G| = \exp(\Omega(n))$

Design lemma & Split-or-Johnson solve this in quasipolynomial time.

* This reduction took 30 years to find; Split-or-Johnson, 3 weeks

Significance to Graph Isomorphism

Encasement problem

Given a k-ary relational structure $\mathfrak{X} = (V, \mathcal{R})$ with $k = O(\log n)$ (where $n = |V|$) and defect$(\mathfrak{X}) \geq 0.1$, find $G \leq \mathrm{Sym}(V)$ (subgroup of the symmetric group) and $A \subseteq V$ such that

- $|A| = \mathrm{polylog}(n)$
- $\mathrm{Aut}(X)_{(A)} \leq G$ (pointwise stabilizer of A)
- $|\mathrm{Sym}(V) : G| = \exp(\Omega(n))$

Design lemma & Split-ot-Johnson solve this in quasipolynomial time:

good coloring: $|S_n : S_k \times S_{n-k}| = \binom{n}{k} > (n/k)^k > 10^{n/10}$

good partition: even better; but what if **Johnson found?**

Significance to Graph Isomorphism

Encasement problem

Given a k-ary relational structure $\mathfrak{X} = (V, \mathcal{R})$ with $k = O(\log n)$ (where $n = |V|$) and defect$(\mathfrak{X}) \geq 0.1$, find $G \leq \mathrm{Sym}(V)$ (subgroup of the symmetric group) and $A \subseteq V$ such that

- $|A| = \mathrm{polylog}(n)$
- $\mathrm{Aut}(X)_{(A)} \leq G$ (pointwise stabilizer of A)
- $|\mathrm{Sym}(V) : G| = \exp(\Omega(n))$

Design lemma & Split-ot-Johnson solve this in quasipolynomial time:

good coloring: $|S_n : S_k \times S_{n-k}| = \binom{n}{k} > (n/k)^k > 10^{n/10}$
good partition: even better; but what if **Johnson found?**

$$|\mathrm{Aut}(J(k,t))| = k! \approx \exp(n^{1/t}) \leq \exp(\sqrt{n})$$

– dramatic reduction, can only be repeated $O(\log \log n)$ times

Naive refinement

Input vertex-colored graph

Output refined coloring

$c(x)$ — color of vertex x $\qquad$ I – set of colors

$d_i(x) := \#$ neighbors of x of color $i \in I$

refinement step

for $x \in V$ **do simultaneously**

$$c(x) \leftarrow (c(x); d_i(x) \mid i \in I)$$

Naive refinement

refinement step

for $x \in V$ **do simultaneously**

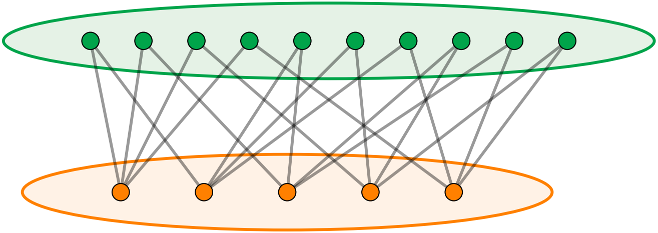
$$c(x) \leftarrow (c(x); d_i(x) \mid i \in I)$$

Repeat until coloring stable

Exercise stable iff **"equitable partition,"** i.e.,

(i) each color class induces regular subgraph

(ii) each pair of color classes induces semiregular bipartite subgraph

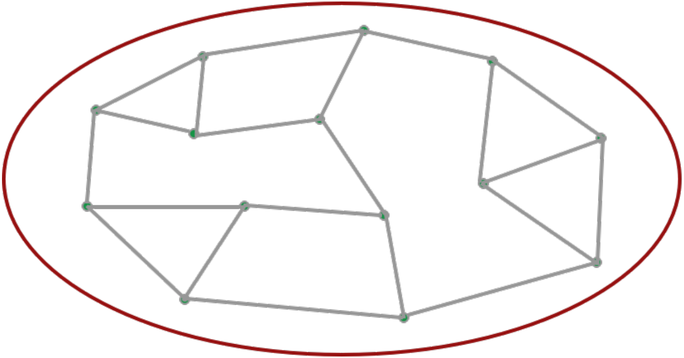


Naive refinement

If input: graph (just one color)

first round: **color** vertices **by degree**

regular graphs: stable, no refinement



Naive refinement

complete split: each vertex gets different color

$\rightarrow$ just 1 candidate isomorphism

Naive refinement

complete split: each vertex gets different color
$\rightarrow$ just 1 candidate isomorphism

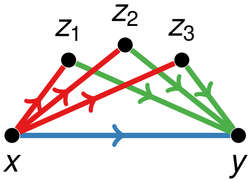
Theorem (B – Erdös – Selkow, 1979)

For almost all graphs, complete split in 2 rounds.
$\therefore$ *ISO test in linear time*

Exercise $\heartsuit$ (Abe Mowshovitz, $\approx$ 1970) If characteristic polynomial of adjacency matrix **irreducible** then
naive refinement $\rightarrow$ complete split.
(Hint: Only equitable partition: discrete)

Advanced refinement: Weisfeiler-Leman 1968

color all ordered pairs, refine by counting triples with shared base and same color composition



coherent configurations: stable under WL

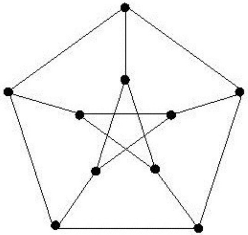
Advanced refinement: Weisfeiler-Leman 1968

Input: graph $\rightarrow$ rank 3 (3 colors):

diagonal $\{(x, x) \mid x \in V\}$,

edge, edge of complement

strongly regular graphs: stable under WL



Coherent configurations

Coherent configuration of rank r: partition of

$V \times V = R_1 \dot\cup \ldots \dot\cup R_r$

R_i — **constituent** (digraph) of color i

coloring: $c(x,y) = i$ if $(x,y) \in R_i$

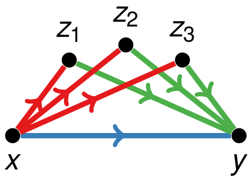
Axioms

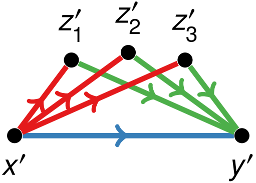
1. diag colors $\neq$ off-diag colors

   $\mathrm{diag}(V) = \{(x,x) \mid x \in V\}$

2. color $x \rightarrow y$ determines color $x \leftarrow y$
3. **intersection numbers:** for any $(x,y) \in R_k$

   p_{ij}^k : $\#\, z$ s.t. $(x,z) \in R_i$ and $(z,y) \in R_j$





Coherent configurations

Homogeneous CC: all vertices same color

Primitive CC: homogeneous and
every off-diagonal constituent (strongly) connected

Schurian case: $G \leq \mathrm{Sym}(\Omega)$
$\mathfrak{X}(G) = (\Omega; \text{orbitals})$ (orbits on $\Omega \times \Omega$)

$\mathfrak{X}(G)$ homogeneous iff G transitive

$\mathfrak{X}(G)$ primitive iff G primitive

$\mathfrak{X}(G)$ uniprimitive iff G primitive and not doubly transitive

Primitice CCs: Johnson schemes

DEF: $\mathfrak{J}(k,t)$ **Johnson scheme** $t \geq 1 \quad k \geq 2t+1$

vertex set $V = \{v_T \mid T \subseteq \Delta, |T| = t\}$ where $|\Delta| = k$

$$|V| = \binom{k}{t}$$

colors: $c(v_T, v_S) = |T \setminus S|$

Example: $\mathfrak{J}(5,2) \quad \sim$ Petersen's graph

Much symmetry: $\mathrm{Aut}(\mathfrak{J}(k,t)) = \mathrm{Aut}(J(k,t)) \cong S_k$

$k! \approx \exp(n^{1/t})$ automorphisms

Coherent configurations

Homogeneous CC: all vertices same color

Primitive CC: homogeneous and
every off-diagonal constituent (strongly) connected

Uniprimitive CC (UPCC):
primitive of rank ≥ 3 (not a clique)

Coherent configurations: a crash course

vertex-color: $c(x) := c(x,x)$

$V = C_1 \cup \cdots \cup C_s$ vertex-color classes

Fact. Edge-color $c(x,y)$ "aware of"
vertex colors $c(x)$ and $c(y)$ $x \to y$

Corollary. $(\forall i)(\exists j,k)(R_i \subseteq C_j \times C_k)$
constituent **homogeneous** $(j = k)$ or **bipartite** $(j \neq k)$

Fact. **Homogeneous** constituent: <u>biregular digraph</u>

Fact. **Bipartite** constituent: <u>semiregular bipartite graph</u>

Coherent configurations: a crash course

Notation. For $R \subseteq V \times V$ and $x \in V$

$R(x) = \{y \in V \mid x \xrightarrow{R} y\}$ — set of **out-neighbors**

Def. **Strong component** of digraph:
equiv class of mutually accessible vertices

Weak component: component of its symmetrization
(ignore orientation)

Fact. The weak components of a *homogeneous constituent* are its strong component.

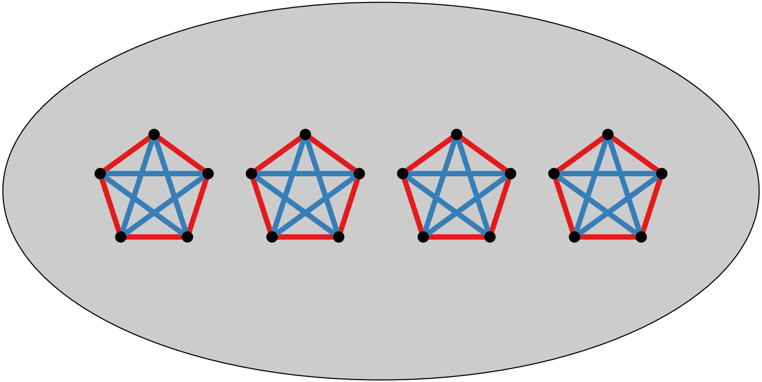
Proof.

DEF: Eulerian digraph: $(\forall x \in V)(\deg^+(x) = \deg^-(x))$

Lemma. Eulerian digraph weakly $\Longrightarrow$ strongly conn

Coherent configurations: a crash course

Fact. The weak components of a constituent have equal order (number of vertices).



Coherent configurations: induced subconfigurations

$X = (V, E)$ digraph — $E \subseteq V \times V$
$\mathfrak{X} = (V, c)$ coherent configuration

Induced subgraph of digraph: for $A \subseteq V$

$$X[A] = (A, E \cap (A \times A))$$

Induced subconfiguration: for $A \subseteq V$

$$\mathfrak{X}[A] = (A, c_{|A \times A})$$

Bipartite graph: $(V_1, V_2; E)$ where $E \subseteq V_1 \times V_2$

Induced bipartite subgraph of digraph:

for $A, B \subseteq V$, $A \cap B = \emptyset$

$$X[A, B] = (A, B; E \cap (A \times B))$$

Induced bipartite subconfiguration:

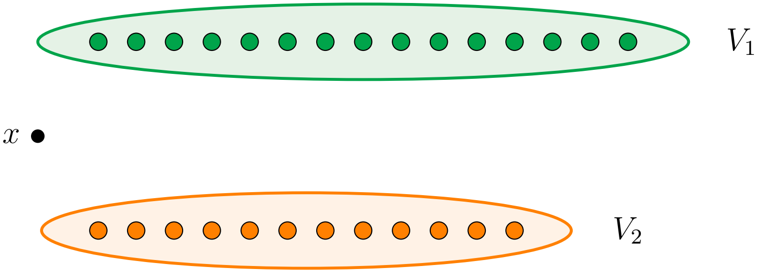
for $A, B \subseteq V$, $A \cap B = \emptyset$

$$\mathfrak{X}[A, B] = (A, B; c_{|A \times B})$$

Fact. $A \subseteq V$ union of color-classes $\implies$ $\mathfrak{X}[A]$ — CC

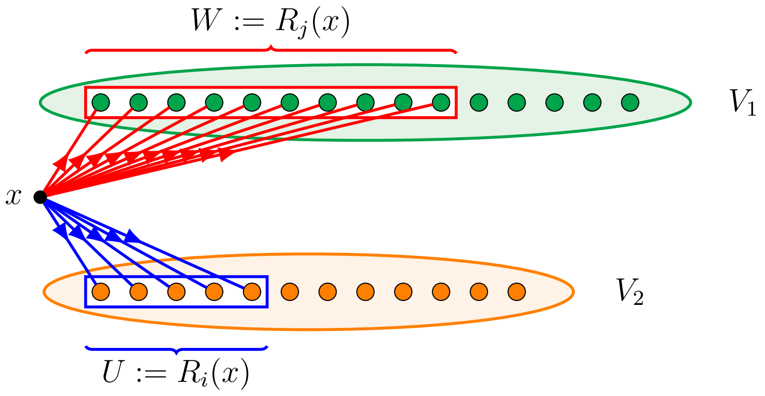
Coherent configurations: the fix

Exercise. $\mathfrak{X}$ CC, V_1, V_2 vertex-color classes, $x \in V$.
Claim. $R_k[U, W]$ semiregular ($U = R_i(x)$, $W = R_j(x)$)



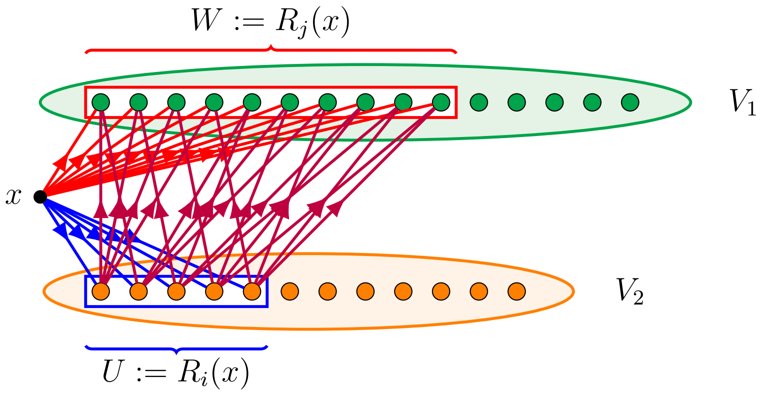
Coherent configurations: the fix

Exercise. $\mathfrak{X}$ CC, V_1, V_2 vertex-color classes, $x \in V$.
Claim. $R_k[U, W]$ semiregular ($U = R_i(x)$, $W = R_j(x)$)



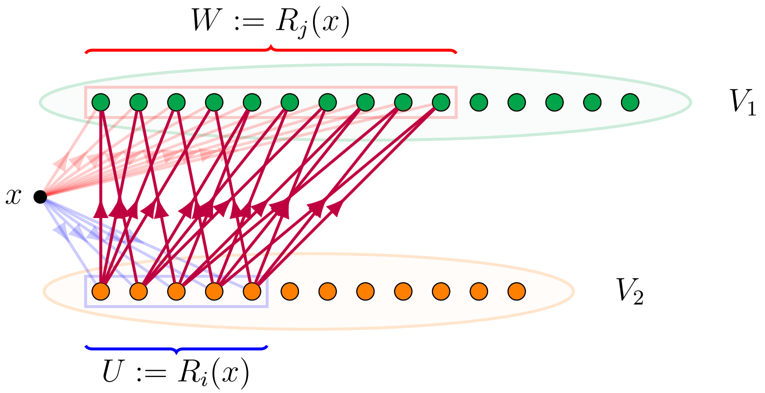
Coherent configurations: the fix

Exercise. $\mathfrak{X}$ CC, V_1, V_2 vertex-color classes, $x \in V$.
Claim. $R_k[U, W]$ semiregular ($U = R_i(x)$, $W = R_j(x)$)



Coherent configurations: the fix

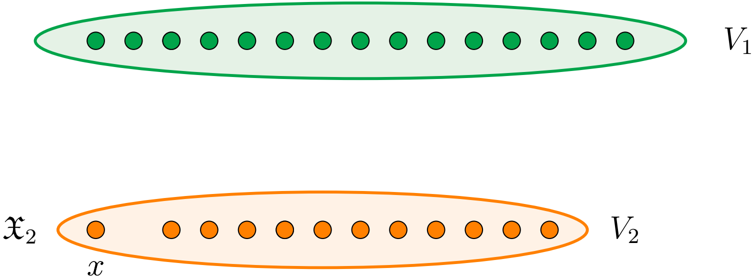
Exercise. $\mathfrak{X}$ CC, V_1, V_2 vertex-color classes, $x \in V$.
Claim. $R_k[U, W]$ semiregular ($U = R_i(x)$, $W = R_j(x)$)



Coherent configurations: the fix

Theorem

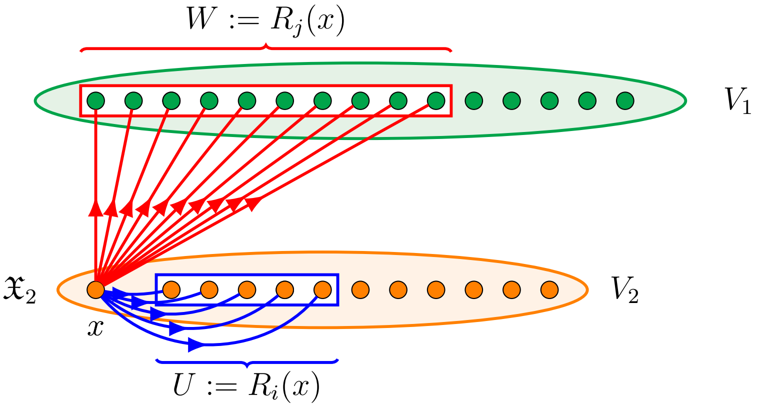
Assume $\mathfrak{X}_2 := \mathfrak{X}[V_2]$ UPCC, $|V_1| > |W| > |V_1|/2$, where $x \in V_2$, $U := R_i(x) \subseteq V_2$, $W = R_j(x) \subseteq V_1$. Then $Y := R_j[U, W]$ nontrivial.



Coherent configurations: the fix

Theorem

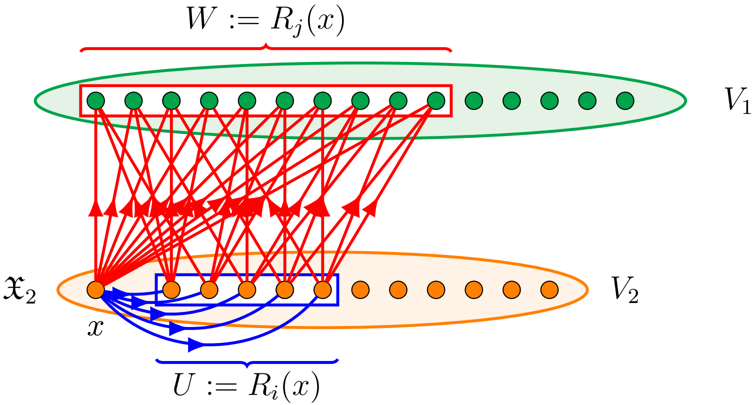
Assume $\mathfrak{X}_2 := \mathfrak{X}[V_2]$ UPCC, $|V_1| > |W| > |V_1|/2$, where $x \in V_2$, $U := R_i(x) \subseteq V_2$, $W = R_j(x) \subseteq V_1$.
Then $Y := R_j[U, W]$ nontrivial.



Coherent configurations: the fix

Theorem

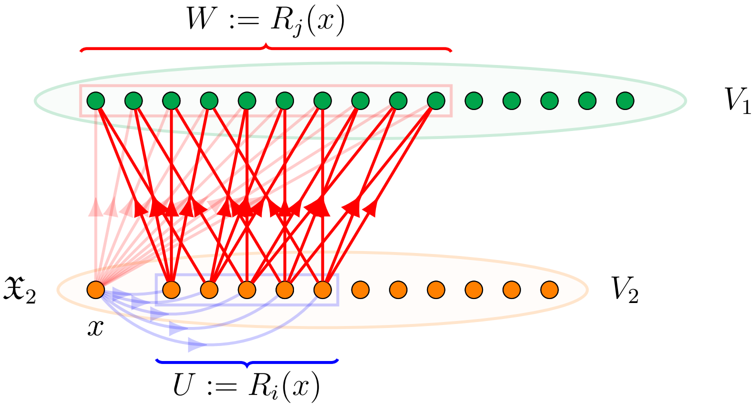
Assume $\mathfrak{X}_2 := \mathfrak{X}[V_2]$ UPCC, $|V_1| > |W| > |V_1|/2$, where $x \in V_2$, $U := R_i(x) \subseteq V_2$, $W = R_j(x) \subseteq V_1$. Then $Y := R_j[U, W]$ nontrivial.



Coherent configurations: the fix

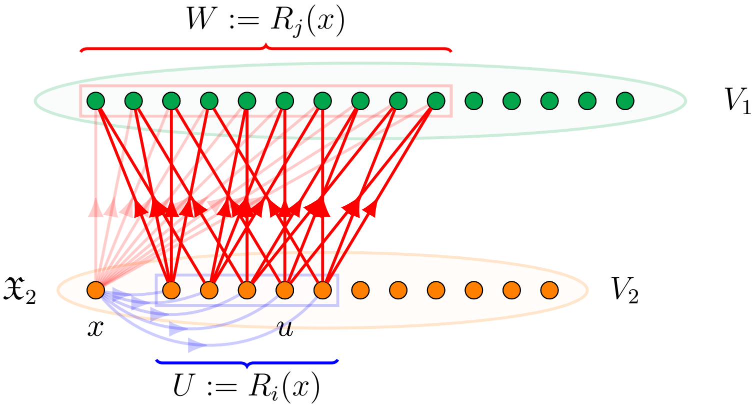
Theorem

Assume $\mathfrak{X}_2 := \mathfrak{X}[V_2]$ UPCC, $|V_1| > |W| > |V_1|/2$, where $x \in V_2$, $U := R_i(x) \subseteq V_2$, $W = R_j(x) \subseteq V_1$. Then $Y := R_j[U, W]$ nontrivial.



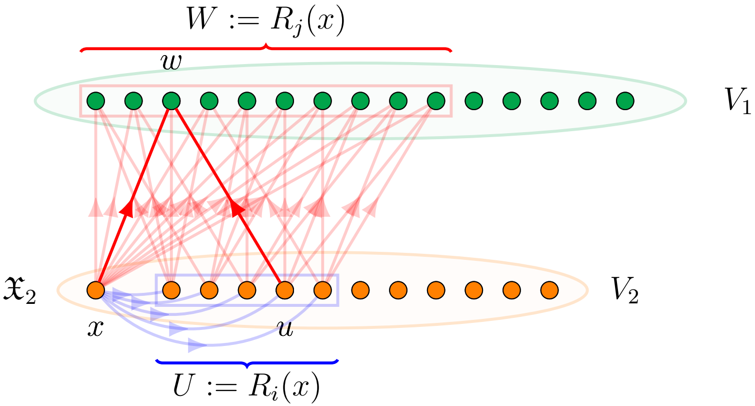
Coherent configurations: the fix

Not empty: $u \in V_2 \implies |R_j(u)| = |R_j(x)| > |V_1|/2$
so $(\exists w \in W)(c(u,w) = j) \quad \therefore (u,v) \in E(Y)$



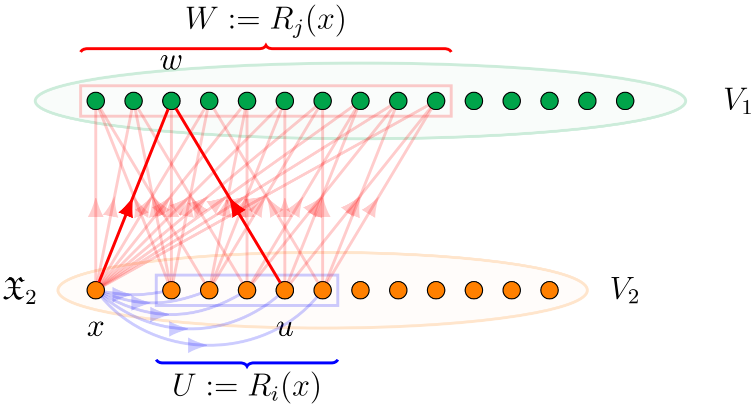
Coherent configurations: the fix

Not empty: $u \in V_2 \implies |R_j(u)| = |R_j(x)| > |V_1|/2$
so $(\exists w \in W)(c(u,w) = j) \quad \therefore (u,v) \in E(Y)$



Coherent configurations: the fix

Not empty: $u \in V_2 \implies |R_j(u)| = |R_j(x)| > |V_1|/2$
so $(\exists w \in W)(c(u, w) = j) \quad \therefore (u, v) \in E(Y)$

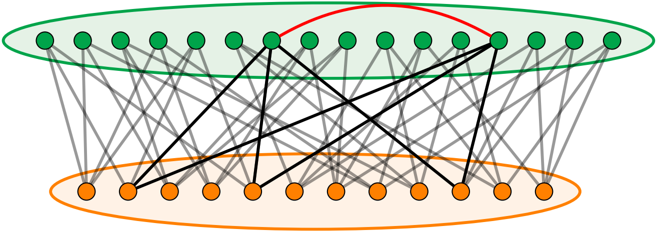


Remains to show: <u>not complete</u>

Coherent configurations: the fix

Lemma (Twin awareness)

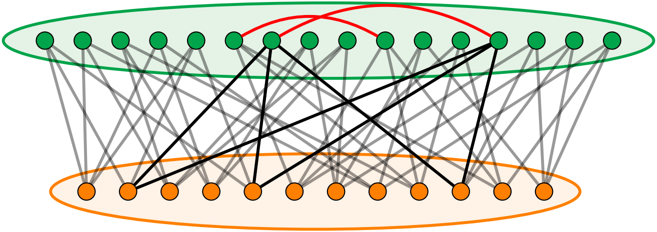
A, B vertex-color classes in CC $\mathfrak{X} = (V; R_1, \ldots, R_r)$ and $R_i \subseteq A \times B$. Then for $x \neq y \in A$ the color $c(x, y)$ determines whether or not x, y are twins for R_i



Coherent configurations: the fix

Lemma (Twin awareness)

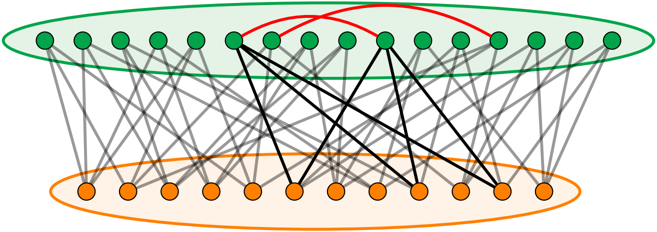
A, B vertex-color classes in CC $\mathfrak{X} = (V; R_1, \dots, R_r)$ and $R_i \subseteq A \times B$. Then for $x \neq y \in A$ the color $c(x, y)$ determines whether or not x, y are twins for R_i



Coherent configurations: the fix

Lemma (Twin awareness)

A, B vertex-color classes in CC $\mathfrak{X} = (V; R_1, \ldots, R_r)$ and $R_i \subseteq A \times B$. Then for $x \neq y \in A$ the color $c(x, y)$ determines whether or not x, y are twins for R_i

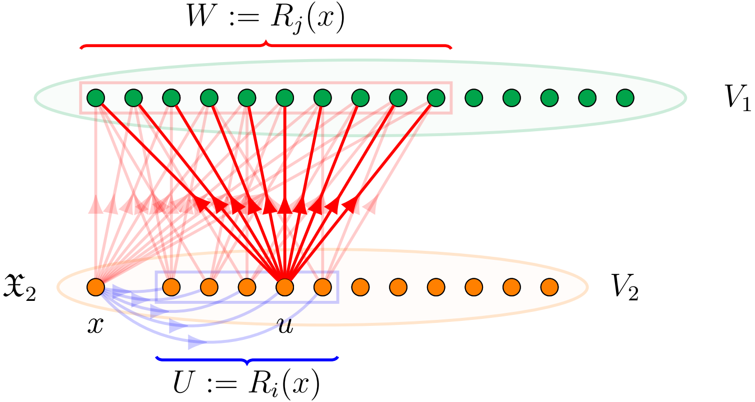


Theorem

Assume $\mathfrak{X}_2$ *UPCC,* $x \in V_2$, $U := R_i(x)$, $W = R_j(x)$, $|W| < |V_1|$. *Then* $Y := R_j[U, W]$ *is not complete.*

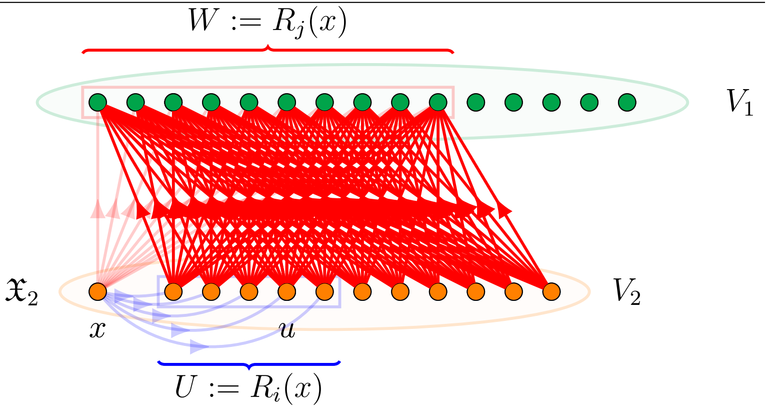
Coherent configurations: the fix

If complete $\implies (\forall u \in U)(R_j(u) \supseteq W)$
But $|R_j(u)| = |R_j(x)| = |W| \implies R_j(u) = W$
$\implies x, u$ twins $\implies$ all pairs of color i are twins
but $(V_2; R_i)$ **connected** $\implies$ all vertices in V_2 twins
$\implies (\forall w \in V_1 \setminus W)(deg_i^-(w) = 0) \quad \rightarrow\leftarrow \quad$ QED(Thm)



Coherent configurations: the fix

If complete $\implies (\forall u \in U)(R_j(u) \supseteq W)$
But $|R_j(u)| = |R_j(x)| = |W| \implies R_j(u) = W$
$\implies x, u$ twins $\implies$ all pairs of color i are twins
but $(V_2; R_i)$ **connected** $\implies$ all vertices in V_2 twins
$\implies (\forall w \in V_1 \setminus W)(deg_i^-(w) = 0) \quad \rightarrow\leftarrow \quad$ QED(Thm)



Split-or-Johnson

Dominant vertex-color: more than 90% of vertices
Good coloring: no dominant color
Good equipartition: dominant color nontriv equipartitioned

Theorem (Split-or-Johnson)

Given a graph with $\text{defect}(X) \geq 0.1$*, at quasipolynomial multiplicative cost one can efficiently find either*

(a) *a good canonical vertex-coloring, or*
(b) *a good canonical equipartition, or*
(c) *a canonically embedded Johnson graph on dominant vertex-color class*

canonicity: functor: isomorphism of input graphs induces isomorphism of embedded structures

Reduces to "SoJ for UPCCs" via WL

Split-or-Johnson for UPCCs

Theorem (Split-or-Johnson for UPCCs)

Given a UPCC on vertex-set V, at quasipolynomial multiplicative cost one can efficiently find either

(a) *a good canonical coloring of V, or*

(b) *a good canonical equipartition of V, or*

(c) *a canonically embedded Johnson graph on dominant vertex-color class.*

Reduces to "SoJ for bipartite graphs"

SoJ for semiregular bipartite graphs

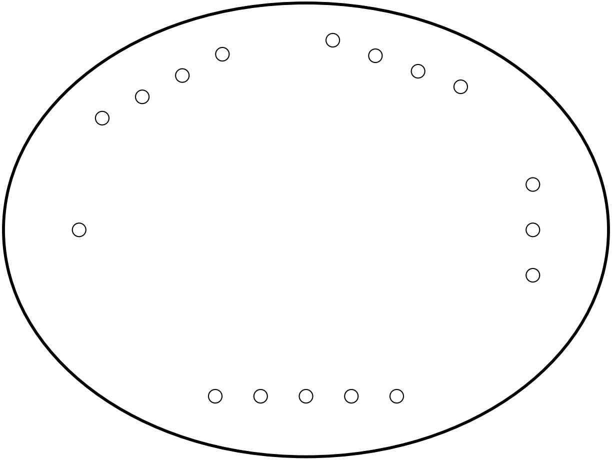
Bipartite graph $(V_1, V_2; E)$ where $E \subseteq V_2 \times V_1$
Notation: $n_i = |V_i|$
Semiregular: all vertices in V_i have same degree $(i = 1, 2)$
Trivial: complete or empty

Theorem (Split-or-Johnson for semireg bipartite graphs)

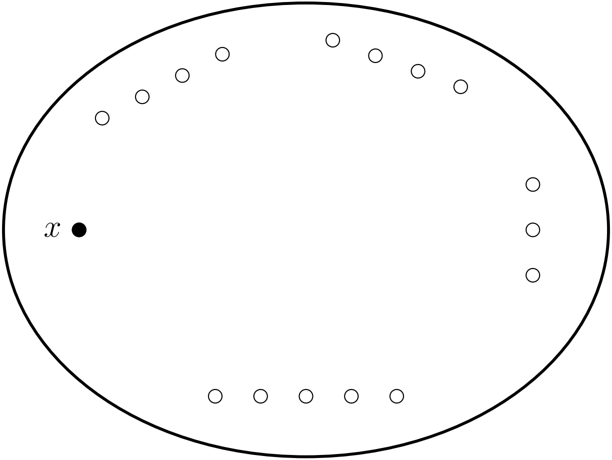
Given nontrivial semiregular bipartite graph $X = (V_1, V_2; E)$ such that $n_2 \leq 0.9 n_1$, at quasipolynomial multiplicative cost one can efficiently find either

(a) *a good canonical coloring of V_1, or*
(b) *a good canonical equipartition of V_1, or*
(c) *a canonically embedded Johnson graph on dominant vertex-color class of V_1*

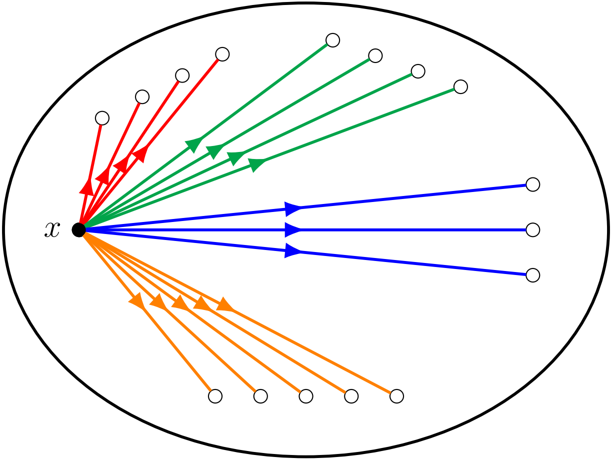
Reduction: UPCC-SoJ to semiregular bipartite SoJ



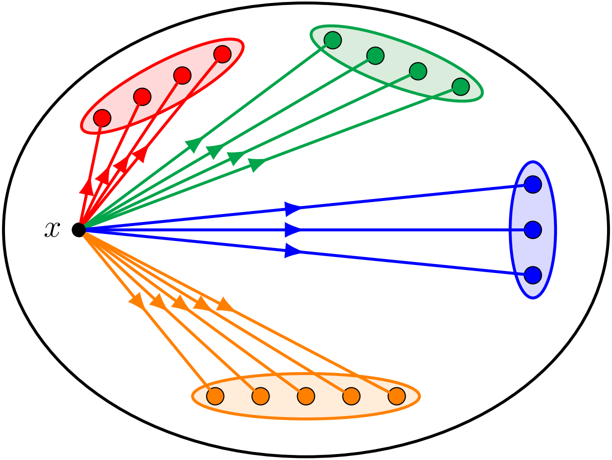
Reduction: UPCC-SoJ to semiregular bipartite SoJ



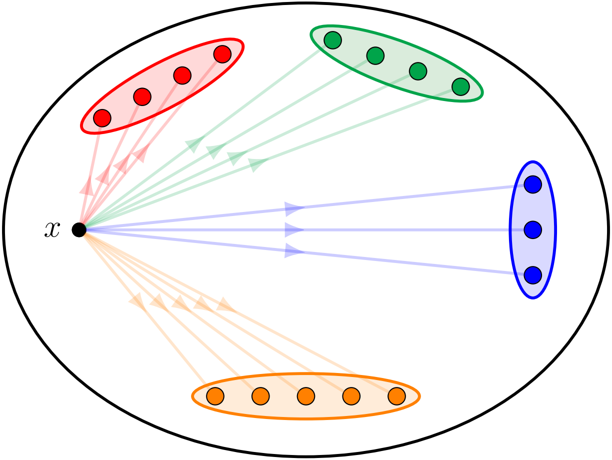
Reduction: UPCC-SoJ to semiregular bipartite SoJ



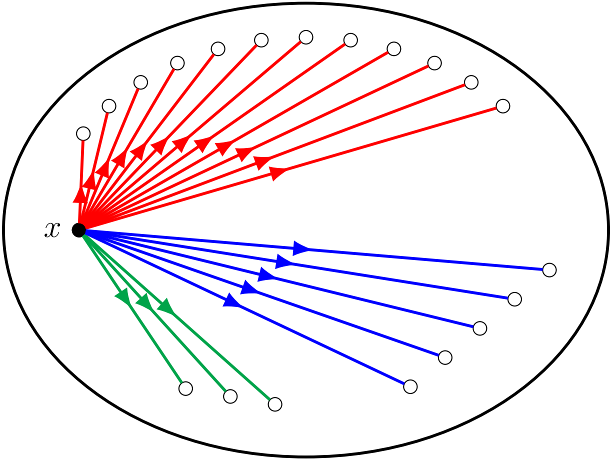
Reduction: UPCC-SoJ to semiregular bipartite SoJ



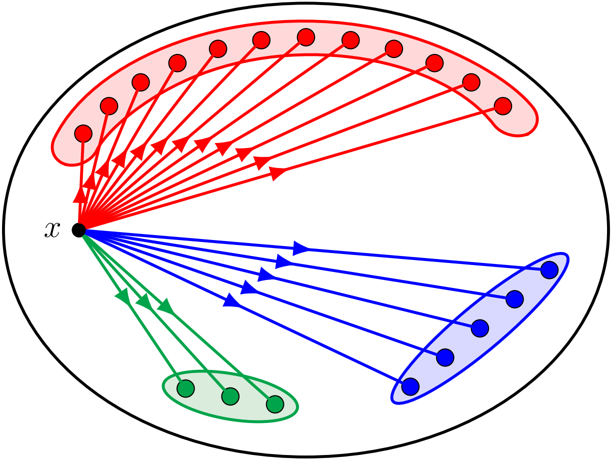
Reduction: UPCC-SoJ to semiregular bipartite SoJ



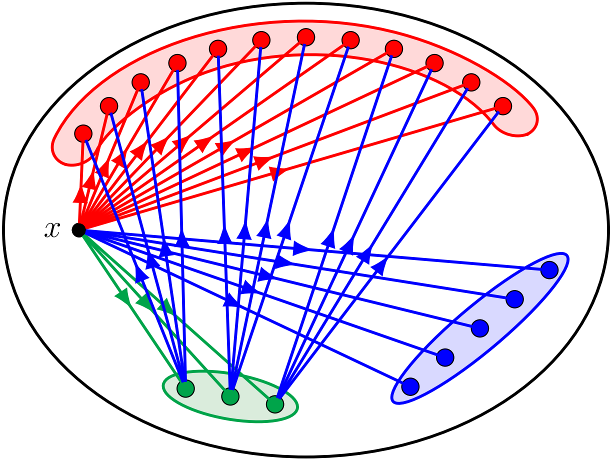
Reduction: UPCC-SoJ to semiregular bipartite SoJ



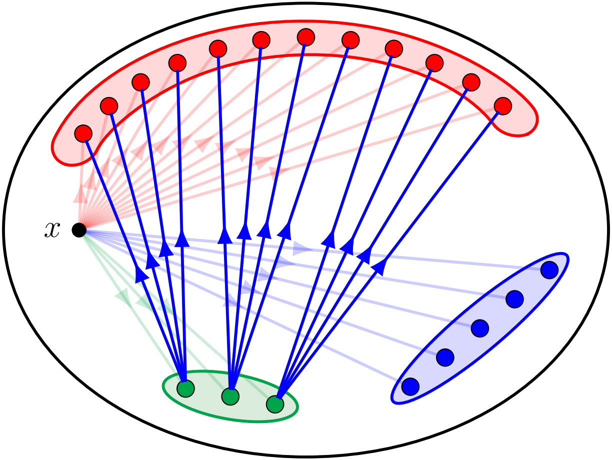
Reduction: UPCC-SoJ to semiregular bipartite SoJ



Reduction: UPCC-SoJ to semiregular bipartite SoJ



Reduction: UPCC-SoJ to semiregular bipartite SoJ



Twins, symmetry defect

Structure $\quad \mathfrak{X} = (V, \mathcal{R})$

Def: $x \neq y \in V$ are **twins** if transposition
$\tau = (x, y) \in \mathrm{Aut}(\mathfrak{X})$

Fact: "twin or equal" — equivalence relation

Def: **Symmetry defect** of $\mathfrak{X}$:

$$\mathrm{defect}(\mathfrak{X}) = 1 - \frac{\max |T|}{|V|}$$ where T — twin equivalence class

Fact: T twin equivalence class $\iff$ $T \subseteq V$ maximal s.t.
$\mathrm{Sym}(T) \leq \mathrm{Aut}(\mathfrak{X})$

Example. If $\mathrm{Aut}(\mathfrak{X}) = S_k \times S_{n-k}$ where $n = |V|$ and $k \leq n/2$ then $\quad \mathrm{defect}(\mathfrak{X}) = k/n$

Twins, symmetry defect for bipartite graphs

Bipartite graph $X = (V_1, V_2; E)$ where $E \subseteq V_1 \times V_2$

$X(v)$ – set of neighbors of v

Fact: $x \neq y \in V_i$ twins $\iff X(x) = X(y)$

Def: Symmetry defect of X in part V_1:

$$\mathrm{defect}_1(X) = 1 - \frac{\max |T|}{|V_1|} \text{ where } T \subseteq V_1 \text{ — twin equiv class}$$

Exercise: If X nontrivial (not empty or complete) semiregular bipartite graph then $\mathrm{defect}_1(X) \geq 1/2$

SoJ for bipartite graphs: unabridged

Theorem (Split-or-Johnson for bipartite graphs)

Given a bipartite graph $X = (V_1, V_2; E)$ with symmetry defect ≥ 0.1 such that $n_2 \leq 0.9n_1$, then at quasipolynomial multiplicative cost one can efficiently find either

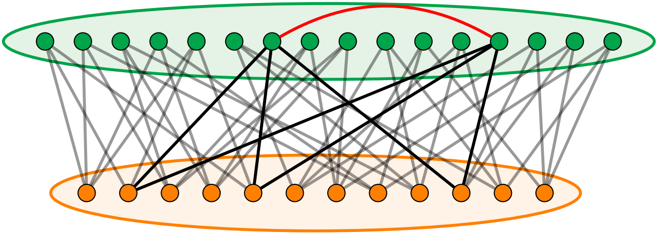
(a) *a good canonical coloring of V_1, or*

(b) *a good canonical equipartition of V_1, or*

(c) *a canonically embedded Johnson graph on dominant vertex-color class of V_1*

Why this is the real thing?

Symmetry defect condition obviously **necessary**



Proof of SoJ for bipartite graphs

Input: bipartite graph $X = (V_1, V_2; E)$ where $n_2 \le 0.9 n_1$,
defect$_X(V_1) \ge 0.1$

Goal: good partition of V_1 or
Johnson graph on $> 90\%$ of V_1

Inductive goal: either achieve "goal"
or halve the size of V_2

May assume defect$_1(X) \ge 0.9$ – otherwise
twin equiv classes — good partition

Start: apply Weisfeiler–Leman to X, obtain CC $\mathfrak{X}$.

$\mathfrak{X}_1 := \mathfrak{X}[V_1], \quad \mathfrak{X}_2 = \mathfrak{X}[V_2]$: CCs

Bipartite part of $\mathfrak{X}$: $\mathfrak{X}_{21} := \mathfrak{X}[V_2, V_1] = \{R_i \mid R_i \subseteq V_2 \times V_1\}$.

Proof of SoJ for bipartite graphs

$\mathfrak{X}_1 := \mathfrak{X}[V_1], \quad \mathfrak{X}_2 = \mathfrak{X}[V_2]$: CCs

Bipartite part of $\mathfrak{X}$: $\mathfrak{X}_{21} := \mathfrak{X}[V_2, V_1] = \{R_i \mid R_i \subseteq V_2 \times V_1\}$.

If $\mathfrak{X}_1$ has no dominant color – goal achieved, **exit**

Otherwise, one can reduce to $\mathfrak{X} \leftarrow \mathfrak{X}[U, W]$ where
$W \subseteq V_1$ dominant color class, $U \subseteq V_2$ color class,
$\mathfrak{X}_{21}$ not monochromatic

$\therefore \mathfrak{X}_1, \mathfrak{X}_2$ homogeneous

May assume $\mathfrak{X}_1$ primitive
otherwise equipartition $\rightarrow$ goal achieved, **exit**

Update $X \leftarrow$ one of the constituents in $\mathfrak{X}_{21}$

So X is **semiregular** $\therefore \text{defect}_i(X) \geq 1/2$

Proof of SoJ for bipartite graphs

Input: CC $\mathfrak{X}$ on $V_1 \cup V_2$
where $\mathfrak{X}_1$ primitive, $\mathfrak{X}_2$ homogeneous,
$\mathfrak{X}_{21}$ not monochromatic, X one of its constituents

Claim: no X-twins in V_1

Proof. By twin-awareness and primitivity of $\mathfrak{X}_1$
the alternative: V_1 is a twin equivalence class
$\therefore$ X trivial, $\mathfrak{X}_{21}$ monochromatic $\rightarrow\leftarrow$ QED

Cases based on $\mathfrak{X}_2$:

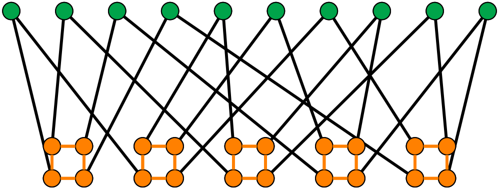
(1) $\mathfrak{X}_2$ imprimitive

(2) $\mathfrak{X}_2$ clique (rank-2 CC)

(3) $\mathfrak{X}_2$ UPCC (uniprimitive coherent configuration)

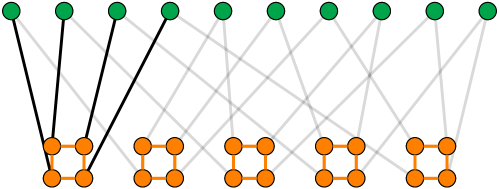
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$



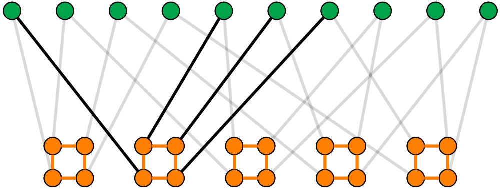
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$
if $\text{defect}_1(X[B_i, V_1]) \geq 1/2$
  return $X \leftarrow X[B_i, V_1]$
  progress: $|B_i| \leq n_2/2$ — V_2 halved



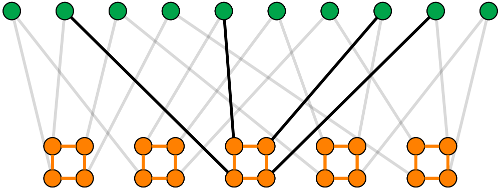
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$
if $\mathrm{defect}_1(X[B_i, V_1]) \geq 1/2$
return $X \leftarrow X[B_i, V_1]$
progress: $|B_i| \leq n_2/2$ — V_2 halved



SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$
if $\mathrm{defect}_1(X[B_i, V_1]) \geq 1/2$
  return $X \leftarrow X[B_i, V_1]$
  progress: $|B_i| \leq n_2/2$ — V_2 halved



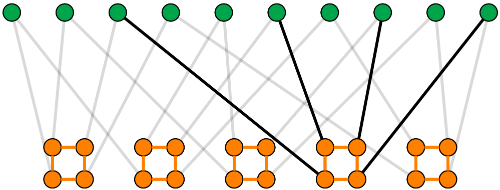
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$

if $\text{defect}_1(X[B_i, V_1]) \geq 1/2$

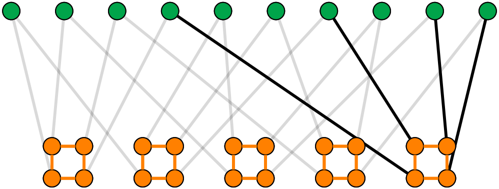
  return $X \leftarrow X[B_i, V_1]$

  progress: $|B_i| \leq n_2/2$ — V_2 halved



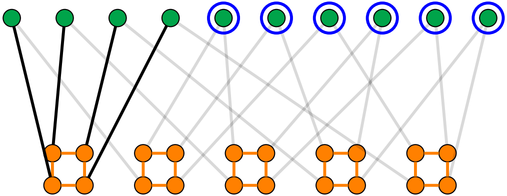
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$
if defect$_1(X[B_i, V_1]) \geq 1/2$
  return $X \leftarrow X[B_i, V_1]$
  progress: $|B_i| \leq n_2/2$ — V_2 halved



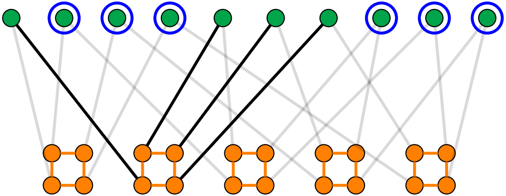
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$
else (:$(\forall i)(\text{defect}_1(X[B_i, V_1]) < 1/2)$:)



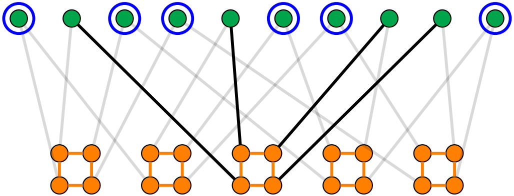
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$
else (:$(\forall i)(\text{defect}_1(X[B_i, V_1]) < 1/2)$:)



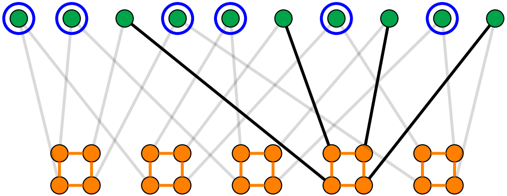
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$
else (:$(\forall i)(\text{defect}_1(X[B_i, V_1]) < 1/2)$:)



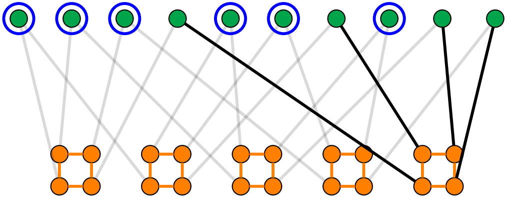
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$
else (:$(\forall i)(\text{defect}_1(X[B_i, V_1]) < 1/2)$:)



SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$
else (:$(\forall i)(\text{defect}_1(X[B_i, V_1]) < 1/2)$:)

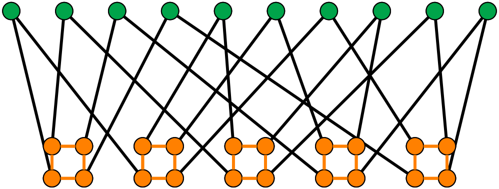


SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$

$Y \leftarrow$ contract each B_i to a vertex
return $X \leftarrow Y$

Lemma: Y semiregular, nontrivial $\therefore$ $\text{defect}_1(Y) \geq 1/2$



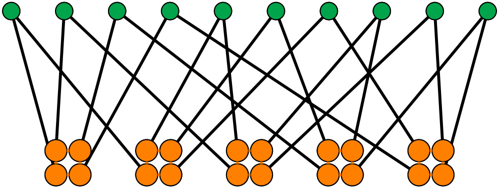
SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$

$Y \leftarrow$ contract each B_i to a vertex

return $X \leftarrow Y$

Lemma: Y semiregular, nontrivial $\therefore$ $\text{defect}_1(Y) \geq 1/2$

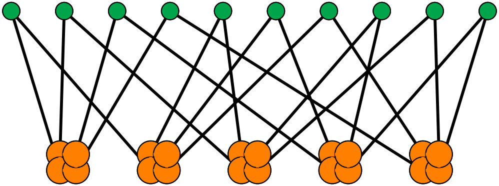


SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot{\cup} \dots \dot{\cup} B_k$

$Y \leftarrow$ contract each B_i to a vertex
return $X \leftarrow Y$

Lemma: Y semiregular, nontrivial $\therefore$ $\text{defect}_1(Y) \geq 1/2$

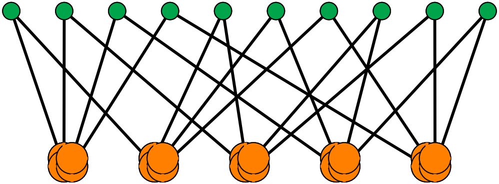


SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$

$Y \leftarrow$ contract each B_i to a vertex
return $X \leftarrow Y$

Lemma: Y semiregular, nontrivial $\therefore$ $\text{defect}_1(Y) \geq 1/2$

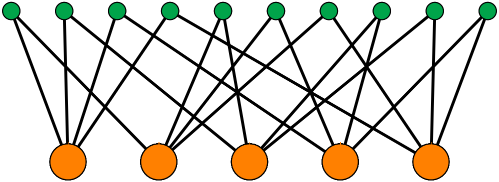


SoJ for bipartite graphs: imprimitive case

Case (1): $\mathfrak{X}_2$ imprimitive: blocks $V_2 = B_1 \dot\cup \ldots \dot\cup B_k$

$Y \leftarrow$ contract each B_i to a vertex
return $X \leftarrow Y$

Lemma: Y semiregular, nontrivial $\therefore$ $\text{defect}_1(Y) \geq 1/2$



Proof of SoJ for bipartite graphs

Input: CC $\mathfrak{X}$ on $V_1 \cup V_2$
where $\mathfrak{X}_1$ primitive, $\mathfrak{X}_2$ homogeneous,
$\mathfrak{X}_{21}$ not monochromatic, X one of its constituents

Cases based on $\mathfrak{X}_2$:

(1) $\mathfrak{X}_2$ imprimitive ✓

(2) $\mathfrak{X}_2$ clique (rank-2 CC)

(3) $\mathfrak{X}_2$ UPCC (uniprimitive coherent configuration)

SoJ for bipartite graphs: UPCC case

Case (3): $\mathfrak{X}_1$ homogeneous, $\mathfrak{X}_2$ UPCC
$\mathfrak{X}_{21}$ not monochromatic

Recursive goal:
Either find good coloring/equipartition of V_1
or halve V_2

SoJ for bipartite graphs: UPCC case

Case (3): $\mathfrak{X}_1$ homogeneous,$\mathfrak{X}_2$ UPCC
$\mathfrak{X}_{21}$ not monochromatic

Recursive goal:
Either find good coloring/equipartition of V_1
or halve V_2

Algorithm. Individualize $x \in V_2$, refine:
new color of $v \in V_1$; $\quad c'(v) = c(x, v)$

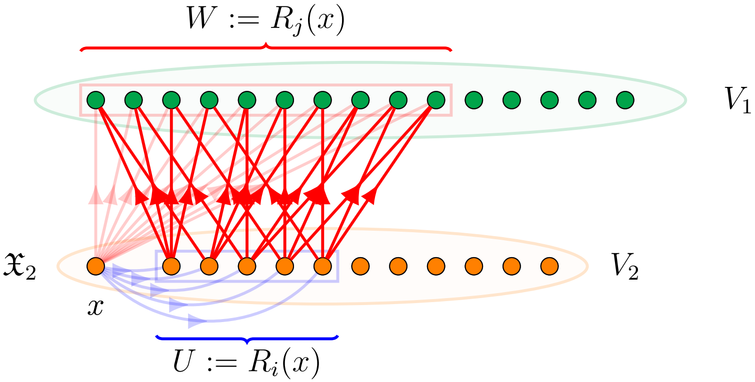
if there is no dominant color in $\mathfrak{X}_{21}$,
this is a good coloring of V_1, goal achieved, **exit**

else: R_j dominant color in $\mathfrak{X}_{21}$

SoJ for bipartite graphs: UPCC case

Situation: R_j dominant color in $\mathfrak{X}_{21}$.
Let $W = R_j(x)$, so $|W| > 0.9n_1$.

Pick a color R_i in $\mathfrak{X}_2$ such that $|R_i(x)| < n_2/2$.
Such color exists b/c $\mathfrak{X}_2$ is not a clique. Let $U = R_i(x)$.

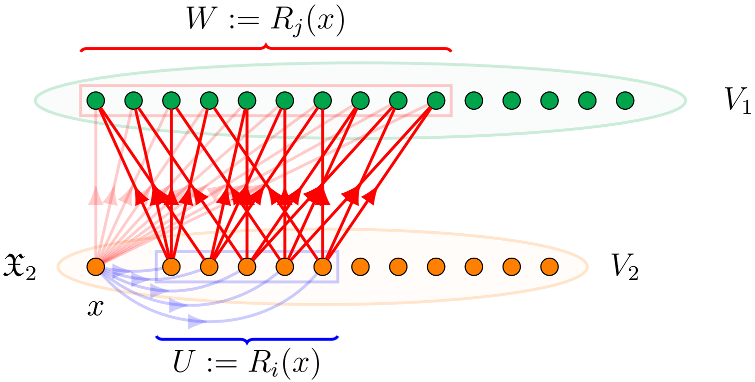


SoJ for bipartite graphs: UPCC case

Let $Y = R_j[U, W]$

return $Y \leftarrow X$. Progress: $|U| < n_2/2$.

Y is semiregular and **nontrivial** by **Theorem**



SoJ for bipartite graphs: UPCC case

Cost analysis

$f(n, m)$ – cost of SoJ for bipartite graph $n = |V_1|$, $m = |V_2|$

reduced $m \leftarrow m/2$ at a cost of 1 individualization

$f(n, m) \leq m \cdot f(n, m/2)\}$

$\therefore \quad f(n, m) = n^{O(\log m)}$ – quasipolynomial QED

Proof of SoJ for bipartite graphs

Input: CC $\mathfrak{X}$ on $V_1 \cup V_2$
where $\mathfrak{X}_1$ primitive, $\mathfrak{X}_2$ homogeneous,
$\mathfrak{X}_{21}$ not monochromatic, X one of its constituents

Cases based on $\mathfrak{X}_2$:

(1) $\mathfrak{X}_2$ imprimitive ✓

(2) $\mathfrak{X}_2$ clique (rank-2 CC)

(3) $\mathfrak{X}_2$ UPCC (uniprimitive CC) ✓

Neighborhood hypergraph in bipartite graph

Case (2): $\mathfrak{X}_2$ clique

Desired progress: move to Case (1) or Case (3)
without increasing the n_i

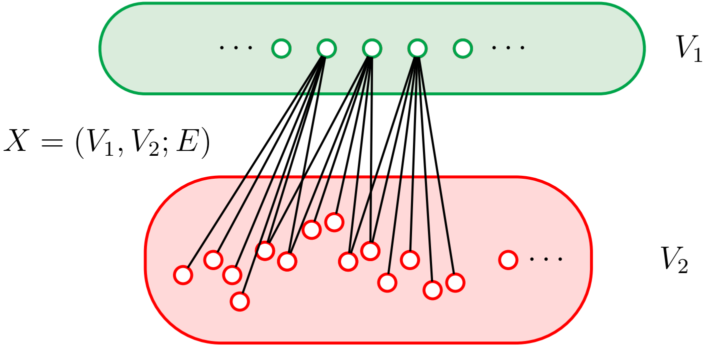
Notation: X graph v vertex $X(v)$: set of neighbors of v

May assume X **semiregular:** $(\forall v \in V_i)(\deg(v) = d_i)$,
nontrivial

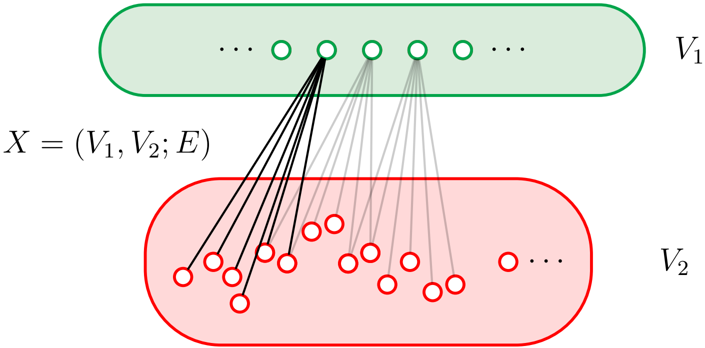
Neighborhood hypergraph: $\mathcal{H} = \{X(v) \mid v \in V_1\}$

d_1-uniform d_2-regular hypergraph on V_2

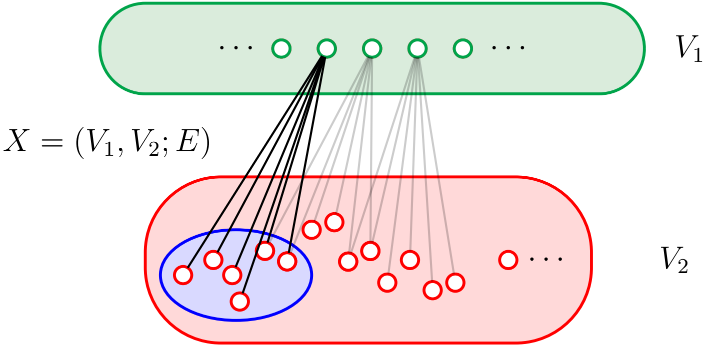
Neighborhood hypergraph in bipartite graph



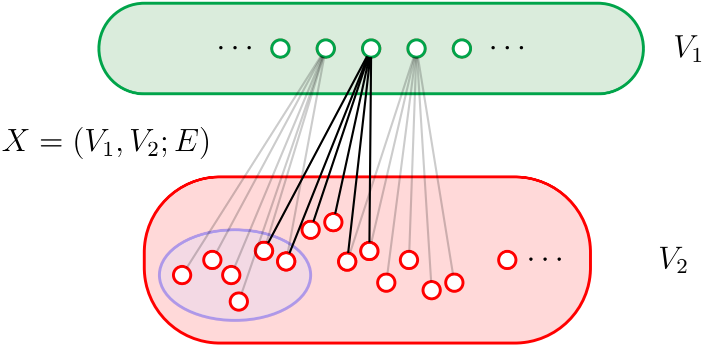
Neighborhood hypergraph in bipartite graph



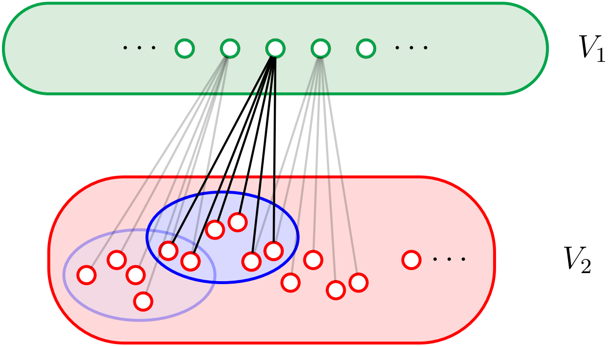
Neighborhood hypergraph in bipartite graph



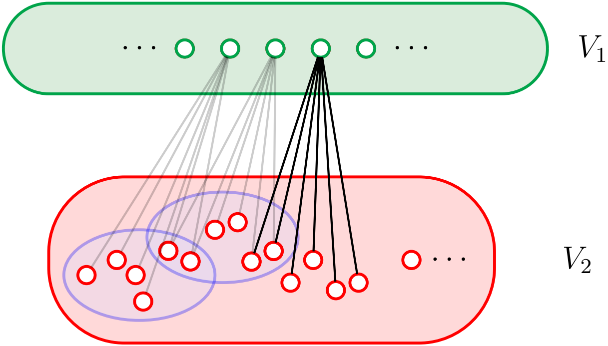
Neighborhood hypergraph in bipartite graph



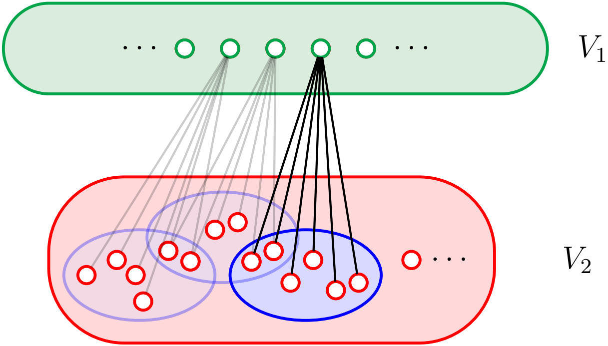
Neighborhood hypergraph in bipartite graph



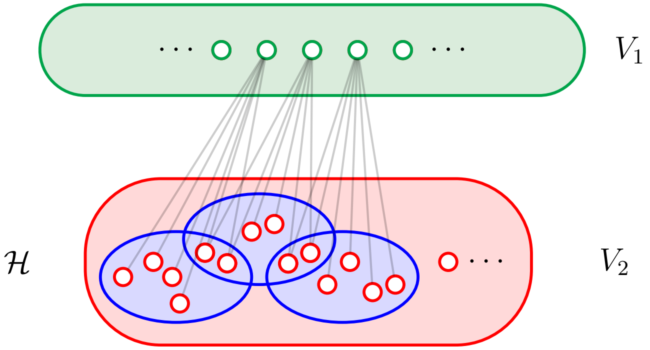
Neighborhood hypergraph in bipartite graph



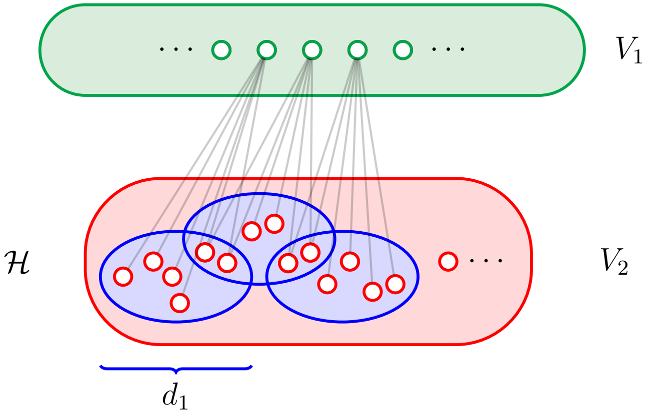
Neighborhood hypergraph in bipartite graph



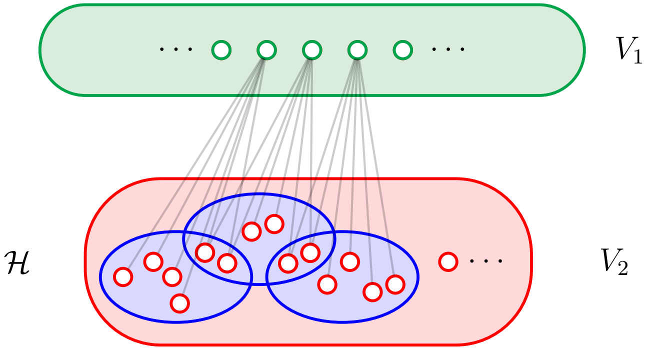
Neighborhood hypergraph in bipartite graph



Neighborhood hypergraph in bipartite graph



Neighborhood hypergraph in bipartite graph



Q: **What if $\mathcal{H}$ complete d_1-uniform hypergraph**?

Neighborhood hypergraph in bipartite graph

Q: **What if $\mathcal{H}$ complete d_1-uniform hypergraph**?

Neighborhood hypergraph in bipartite graph

Q: **What if $\mathcal{H}$ complete d_1-uniform hypergraph**?

$\rightarrow$ **Johnson scheme** $\mathfrak{J}(n_2, d_1)$ canonically on V_1

Canonical: for $x, y \in V_1$ isomorphisms preserve
$|X(x) \cap X(y)|$

This is exactly the case when V_2 is a twin equivalence class for $\mathcal{H}$: $\mathrm{Aut}(\mathcal{H}) = \mathrm{Sym}(V_2)$

Same works for defect$(\mathcal{H}) \leq 1/2$. **Assume** defect$(\mathcal{H}) > 1/2$.

If $d_1 \leq (7/3)\log_2 n_1$: apply **Design Lemma** to $\mathcal{H}$

Neighborhood hypergraph in bipartite graph

If $d_1 > (7/3)\log_2 n_1$: let $t = (7/4)\log_2 n_1$

Def. t-skeleton $\mathcal{H}^{(t)}$ of hypergraph $\mathcal{H}$: all t-subsets of the edges of $\mathcal{H}$.

Lemma (Skeleton defect lemma)

Let $\mathcal{H}$ be a nontrivial, regular, d-uniform hypergraph with n vertices and m edges where $d \le n/2$. Let $(7/4)\log_2 m \le t \le (3/4)d$. Then

$$\text{defect}(\mathcal{H}^{(t)}) \ge 1/4$$

Now apply the Design Lemma to $\mathcal{H}^{(t)}$.

k-ary coherent configurations: definition

DEF: k-ary **partition structure:**

$$\mathfrak{X} = (V, c) \text{ where } \qquad c : V^k \to \{\text{colors}\}$$

DEF: k-ary **configuration:**

- color $c(x_1, \dots, x_k)$ determines equiv relation on $[k]$
$$i \sim j \iff x_i = x_j$$
- $(\forall f : [k] \to [k])$
  color $c(x_1, \dots, x_k)$ determines $c(x_{f(1)}, \dots, x_{f(k)})$

DEF: k-ary **coherent configuration:**

$(\forall i_1, \dots, i_k \le k)(c(x_1, \dots, x_k)$ determines
$(\#z \in V)(\forall j \le k)(c(x_1, \dots, z_{j\text{-th position}}, \dots, x_k) = i_j\,)$
$(r^{k+1}$ **intersection numbers** where $r = |\{\text{colors}\}|)$

k-ary coherent configurations: restriction

DEF: k-ary **partition structure:**

$\mathfrak{X} = (V, c)$ where $\quad c : V^k \to \{\text{colors}\}$

DEF: $(k-\ell)$-ary **restriction** of

k-ary partition structure $\mathfrak{X} = (V, c)$:

let $\vec{x} = (x_1, \ldots, x_\ell) \in V^\ell$, $\; V' = V \setminus \{x_1, \ldots, x_\ell\}$

$\mathfrak{X}_{\vec{x}} := (V', c')$ where for $\vec{y} \in (V')^{k-\ell}$ we set

$$c'(\vec{y}) = c(\vec{x}\vec{y})$$

Fact: Restriction of CC_k is $CC_{k-\ell}$

k-ary coherent configurations: skeleton

DEF: k-ary **partition structure:**

$\mathfrak{X} = (V, c)$ where $\quad c : V^k \to \{\text{colors}\}$

DEF: ℓ-ary **skeleton** of

k-ary partition structure $\mathfrak{X} = (V, c)$:

ℓ-ary partition structure $\mathfrak{X}^{(\ell)} = (V, c')$ where

$$c'(x_1, \ldots, x_\ell) = c(x_1, \ldots, x_\ell, \underbrace{x_\ell, \ldots, x_\ell}_{(k-\ell)\text{ times}})$$

Fact: ℓ-skeleton on CC_k is CC_ℓ

Design Lemma

Theorem (Design Lemma)

Given a k-ary relational structure $\mathfrak{X}$ with $\text{defect}(\mathfrak{X}) \geq 0.1$, *one can individualize $k-1$ vertices and find, in $n^{O(k)}$ time,*

(a) *a good canonical vertex-coloring, or*
(b) *a good canonical equipartition, or*
(c) *a canonically embedded regular graph on dominant vertex-color class ($\geq 90\%$ of vertices)*

An application of Fisher's inequality

Lemma (Large clique lemma)

$\mathfrak{X} = (V, c)$ (classical) coherent configuration with vertex-color classes $W_1, \ldots, W_s$. Assume

- *W_1 dominant:* $(\forall i \geq 2)(|W_1| > |W_i|)$
- *W_1 induces a clique in $\mathfrak{X}$*

Then W_1 is a twin equivalence class in $\mathfrak{X}$.

Proof uses Fisher's inequality for block design:

in a BIBD, $\quad \#$ of blocks $\geq \#$ of vertices

Design Lemma: idea of proof

Let $\mathfrak{X}$ be k-ary coherent, with defect ≥ 0.1.

If 2-skeleton $\mathfrak{X}^{(2)}$ has no dominant color class
DONE (good coloring)

Assume **dominant vertex-color** $C \subseteq V$ $(|C| > 0.9n)$

If $\mathfrak{X}^{(2)}[C]$ **imprimitive**
DONE (good equipartition)

If $\mathfrak{X}^{(2)}[C]$ **uniprimitive** (primitive, not clique)
DONE (reduced to binary case $\rightarrow$ SoJ)

If $\mathfrak{X}^{(2)}[C]$ **clique**
need to destroy it by individualizing $(k-1)$ points

Design Lemma: idea of proof

Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\leq (k-1)$ points to **individualize?**

Design Lemma: idea of proof

Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\leq (k-1)$ points to **individualize?**

C too large to be a twin equivalence class for $\mathfrak{X}$
$\therefore \quad (\exists u, v \in C)(\tau := (u, v) \notin \mathrm{Aut}(\mathfrak{X}))$

Design Lemma: idea of proof

Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\le (k-1)$ points to **individualize?**

C too large to be a twin equivalence class for $\mathfrak{X}$

$\therefore \quad (\exists u, v \in C)(\tau := (u, v) \notin \mathrm{Aut}(\mathfrak{X}))$

$\therefore \quad (\exists \vec{x} \in V^k)(c(\vec{x}^{\,\tau}) \neq c(\vec{x}))$

Design Lemma: idea of proof

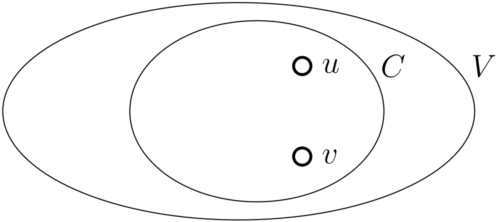
Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\le (k-1)$ points to **individualize?**

C too large to be a twin equivalence class for $\mathfrak{X}$

$\therefore \quad (\exists u, v \in C)(\tau := (u, v) \notin \mathrm{Aut}(\mathfrak{X}))$

$\therefore \quad (\exists \vec{x} \in V^k)(c(\vec{x}^{\,\tau}) \neq c(\vec{x}))$



Design Lemma: idea of proof

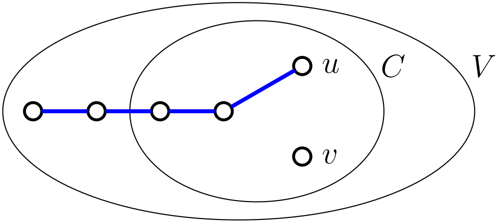
Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\leq (k-1)$ points to **individualize?**

C too large to be a twin equivalence class for $\mathfrak{X}$

$\therefore \quad (\exists u, v \in C)(\tau := (u, v) \notin \mathrm{Aut}(\mathfrak{X}))$

$\therefore \quad (\exists \vec{x} \in V^k)(c(\vec{x}^{\,\tau}) \neq c(\vec{x}\,))$



Design Lemma: idea of proof

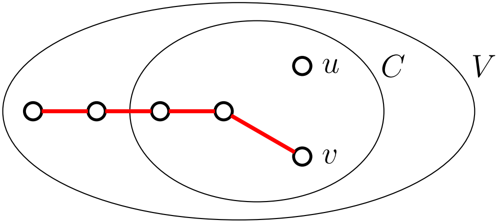
Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\leq (k-1)$ points to **individualize?**

C too large to be a twin equivalence class for $\mathfrak{X}$

$\therefore \quad (\exists u, v \in C)(\tau := (u, v) \notin \mathrm{Aut}(\mathfrak{X}))$

$\therefore \quad (\exists \vec{x} \in V^k)(c(\vec{x}^{\,\tau}) \neq c(\vec{x}))$



Design Lemma: idea of proof

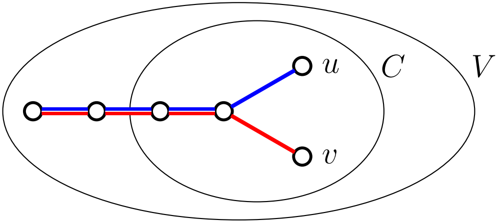
Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\le (k-1)$ points to **individualize?**

C too large to be a twin equivalence class for $\mathfrak{X}$

$\therefore \quad (\exists u, v \in C)(\tau := (u, v) \notin \mathrm{Aut}(\mathfrak{X}))$

$\therefore \quad (\exists \vec{x} \in V^k)(c(\vec{x}^{\,\tau}) \neq c(\vec{x}\,))$



Design Lemma: idea of proof

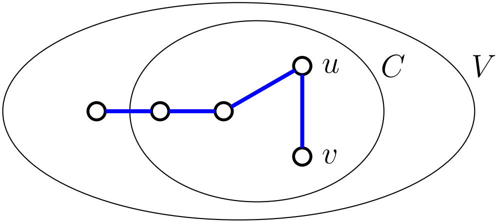
Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\leq (k-1)$ points to **individualize?**

C too large to be a twin equivalence class for $\mathfrak{X}$

$\therefore \quad (\exists u, v \in C)(\tau := (u, v) \notin \mathrm{Aut}(\mathfrak{X}))$

$\therefore \quad (\exists \vec{x} \in V^k)(c(\vec{x}^{\,\tau}) \neq c(\vec{x}))$



Design Lemma: idea of proof

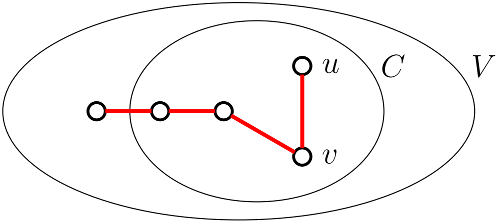
Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\leq (k-1)$ points to **individualize?**

C too large to be a twin equivalence class for $\mathfrak{X}$

$\therefore \quad (\exists u, v \in C)(\tau := (u, v) \notin \mathrm{Aut}(\mathfrak{X}))$

$\therefore \quad (\exists \vec{x} \in V^k)(c(\vec{x}^{\,\tau}) \neq c(\vec{x}))$



Design Lemma: idea of proof

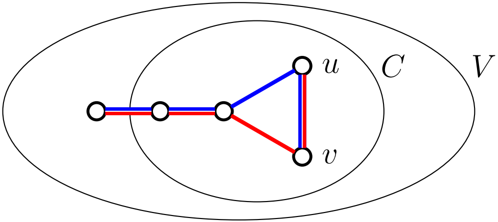
Case: $\mathfrak{X}^{(2)}[C]$ is a **clique**

How do we **guess** the $\leq (k-1)$ points to **individualize?**

C too large to be a twin equivalence class for $\mathfrak{X}$

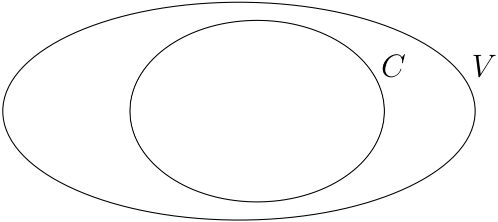
$\therefore \quad (\exists u, v \in C)(\tau := (u,v) \notin \mathrm{Aut}(\mathfrak{X}))$

$\therefore \quad (\exists \vec{x} \in V^k)(c(\vec{x}^{\,\tau}) \neq c(\vec{x}\,))$



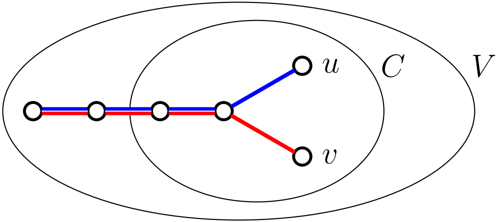
Design Lemma: idea of proof

$\vec{y}$: individualized prefix of $\vec{x}$
Need to destroy clique on $C \setminus \mathrm{supp}(\vec{y})$



Design Lemma: idea of proof

$\vec{y}$: individualized prefix of $\vec{x}$
Need to destroy clique on $C \setminus \text{supp}(\vec{y})$



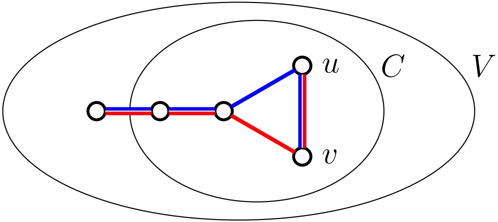
$c(u) \neq c(v)$ after individualizing tail

$\vec{y} = (y_1, \ldots, y_{k-1})$

$\therefore C \setminus \{y_1, \ldots, y_{k-1}\}$ not homogeneous in $\mathfrak{X}_{\vec{y}}$

Design Lemma: idea of proof

$\vec{y}$: individualized prefix of $\vec{x}$
Need to destroy clique on $C \setminus \mathrm{supp}(\vec{y})$



$c(u, v) \neq c(v, u)$ after individualizing tail

$\vec{y} = (y_1, \ldots, y_{k-2})$

$\therefore C \setminus \{y_1, \ldots, y_{k-2}\}$ not a clique in $\mathfrak{X}_{\vec{y}}^{(2)}$

The moral: Hidden irregularity or . . . ?

Split-or-Johnson dichotomy rephrased:

opposite of **hidden irregularity** is

not “hidden regularity” but

The moral: Hidden irregularity or . . . ?

Split-or-Johnson dichotomy rephrased:

opposite of **hidden irregularity** is

not “hidden regularity” but

hidden **robust symmetry**

symmetry that can only be destroyed at high cost