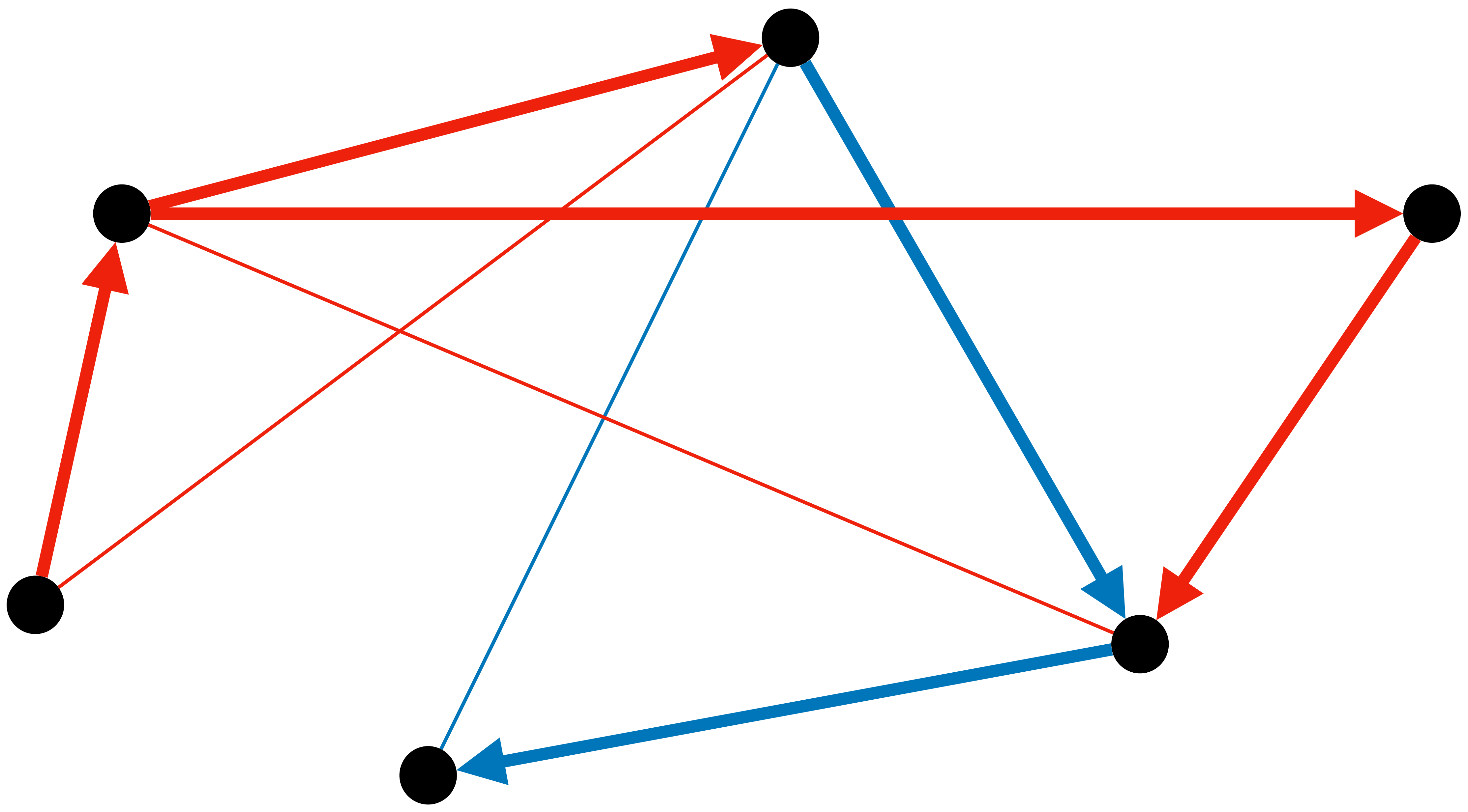
Streaming CSPs

Santhoshini Velusamy (Toyota Technological Institute at Chicago)

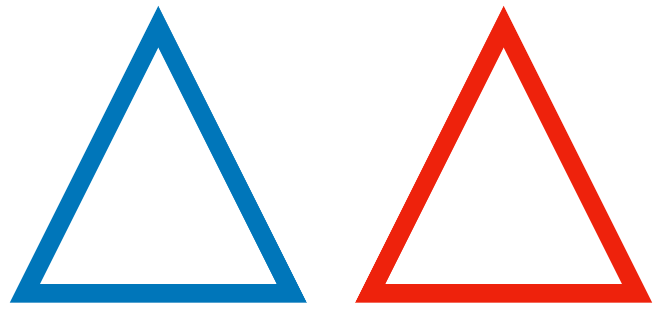
Sublinear Algorithms Program at the Simons Institute of Technology

Constraint Satisfaction Problems (CSPs)

n **variables**

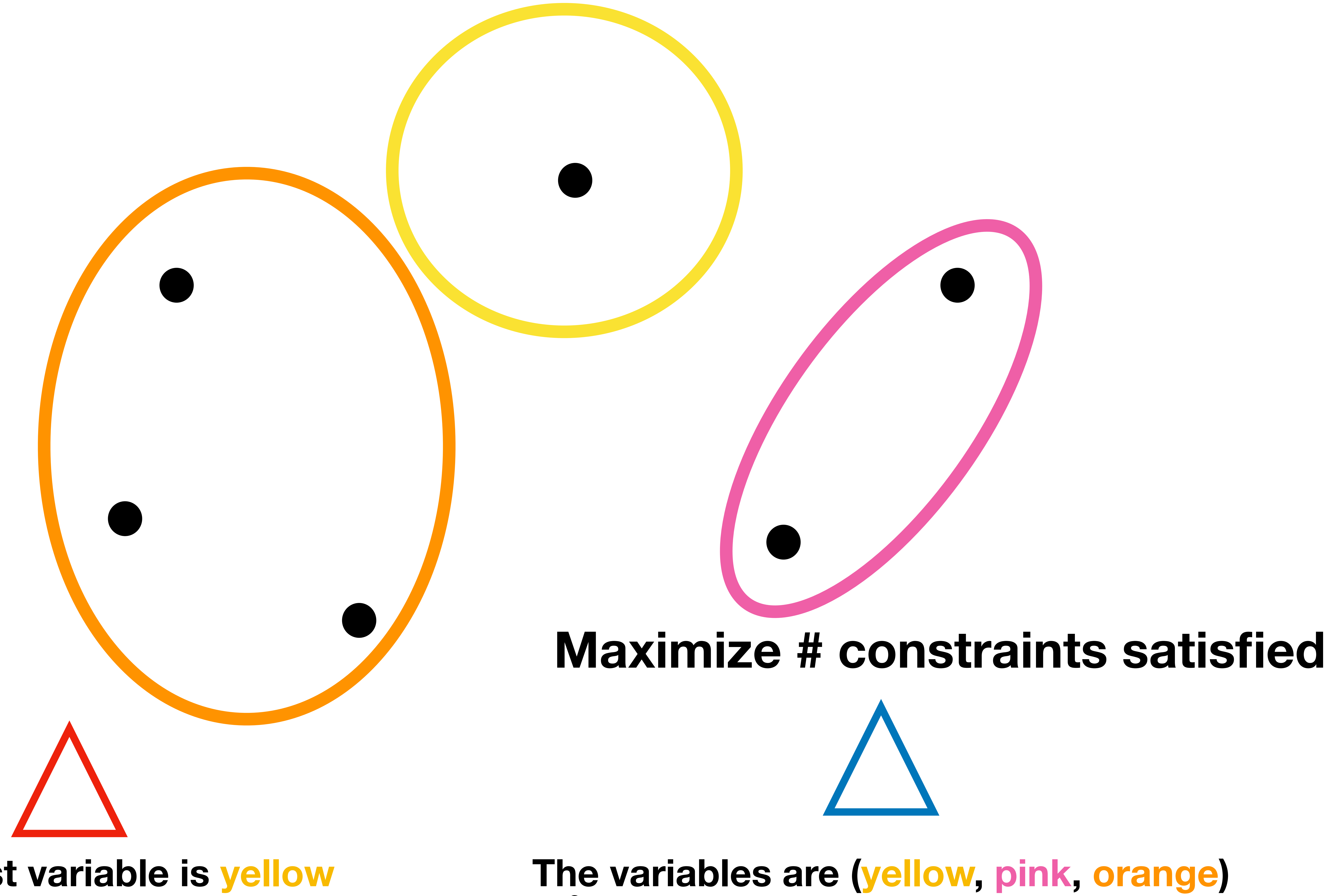


$k = 3$

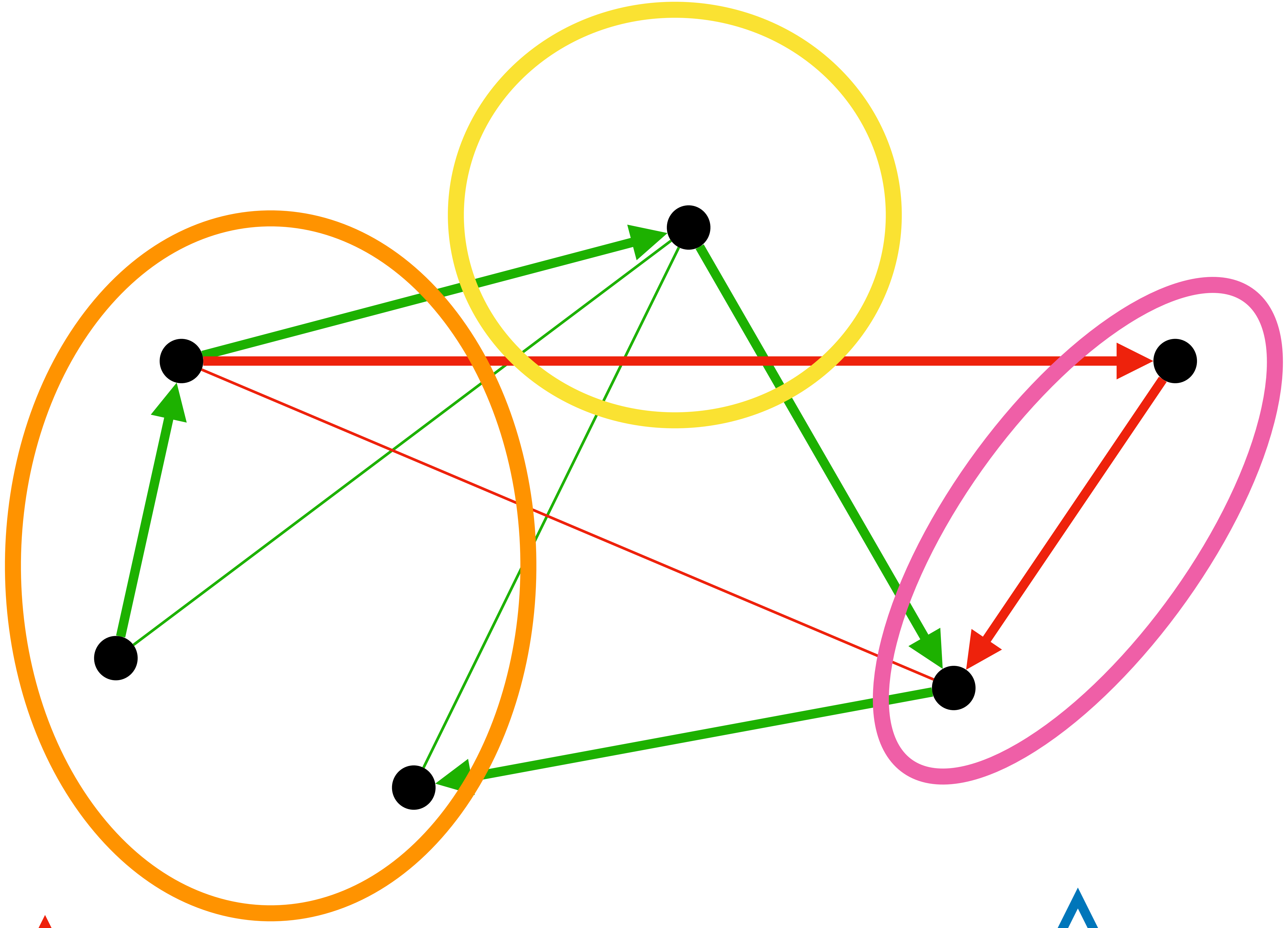


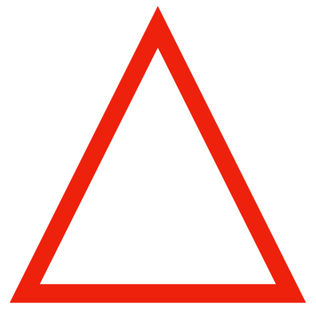
Constant number of constraint types of arity k

Constraint Satisfaction Problems (CSPs)

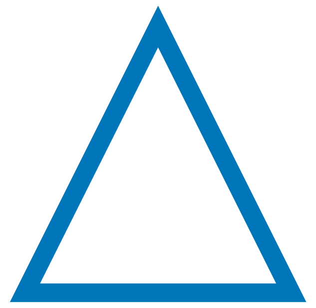


Constraint Satisfaction Problems (CSPs)





Last variable is yellow

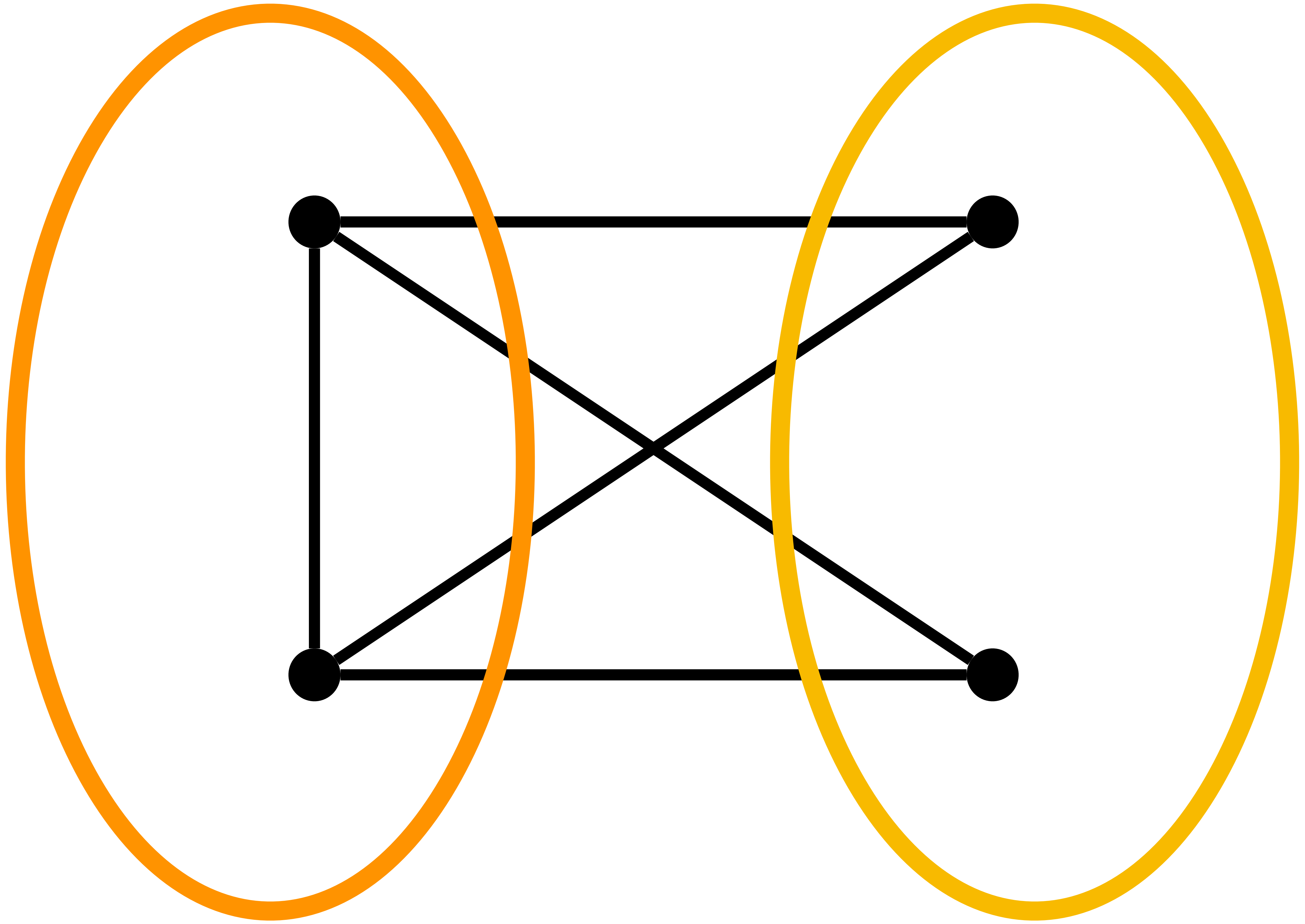


The variables are (yellow, pink, orange)

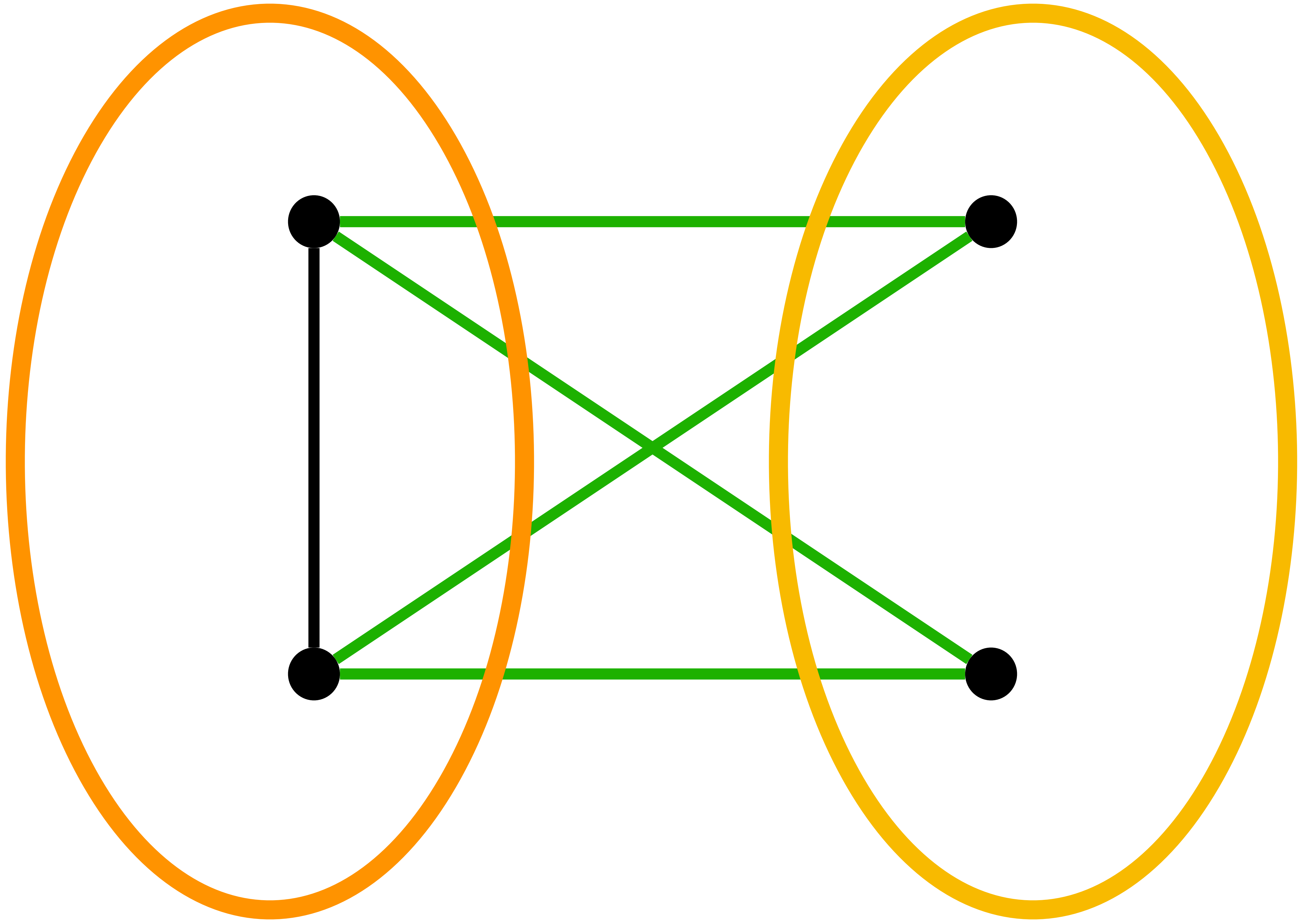
Why study CSPs?

- Structure to study infinitely many problems simultaneously
- Contains several problems of interest: Max-CUT, Max-DICUT, Max-k-SAT,…
- Allows for finite classifications through dichotomy theorems[Schaefer '78, Raghavendra '08, Khot-Tulsiani-Worah '14, Bulatov '17, Zhuk '20, Chou-Golovnev-Sudan-V ’21, Ghoshal-Lee '22, Kol-Parmanov-Saxena-Yu ’23]
  - Insights about what lies at the heart of approximating CSPs
- Discovery of new techniques that are broadly applicable. For example, SDP rounding, Sum of squares, Unique Games,..

Max-CUT



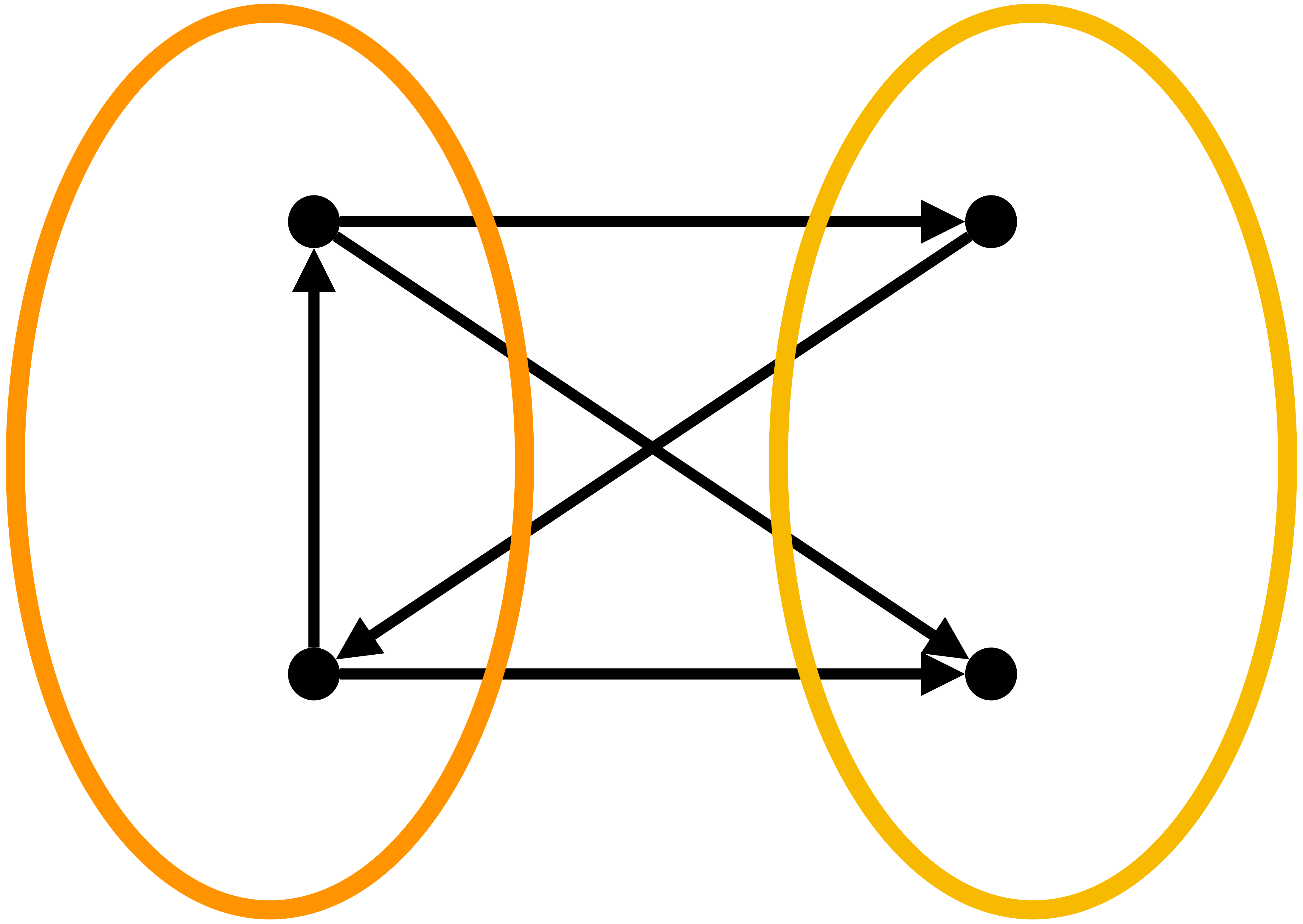
Max-CUT



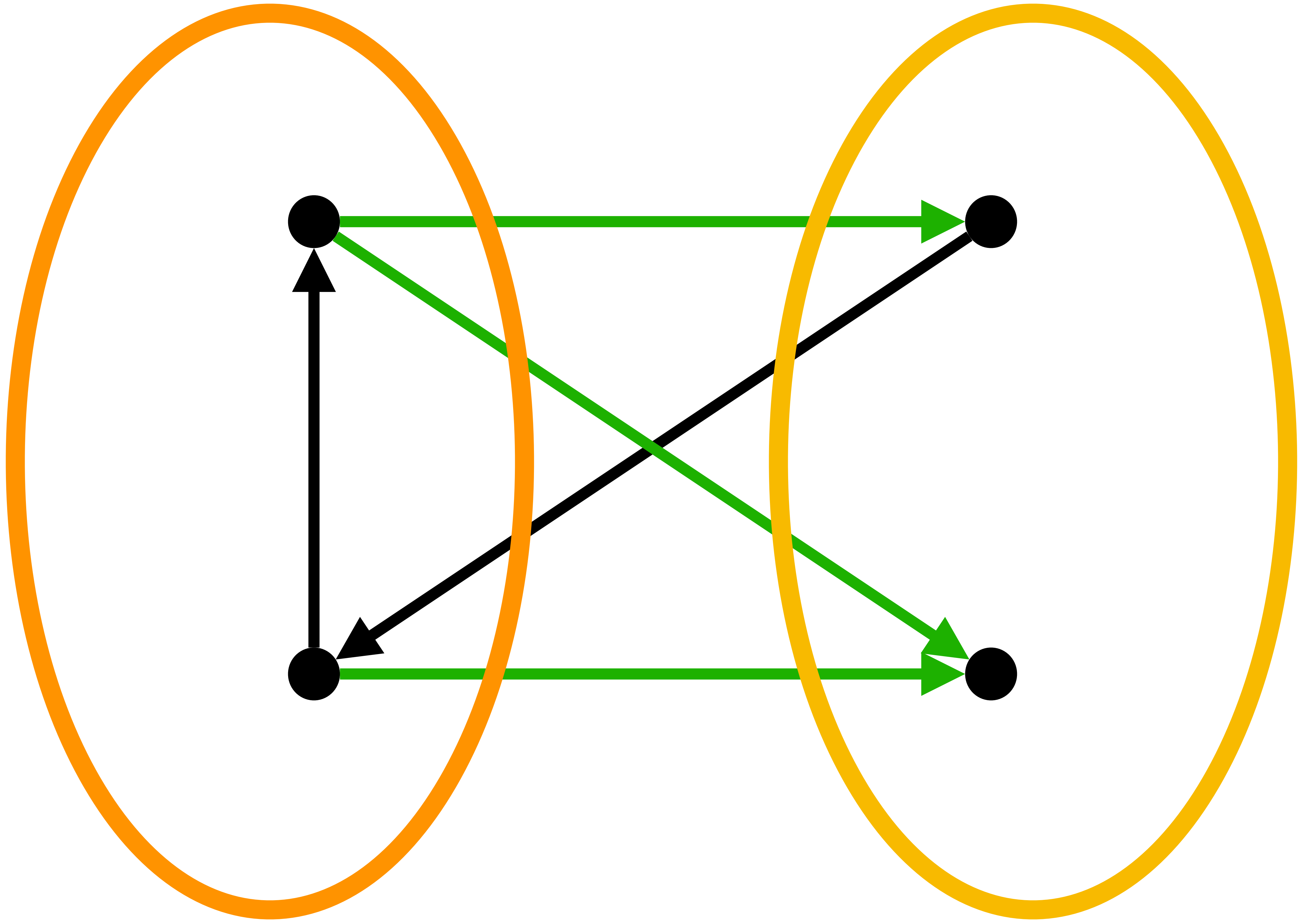
CUT size = 4

Maximize CUT size

Max-DICUT



Max-DICUT



DICUT size = 3

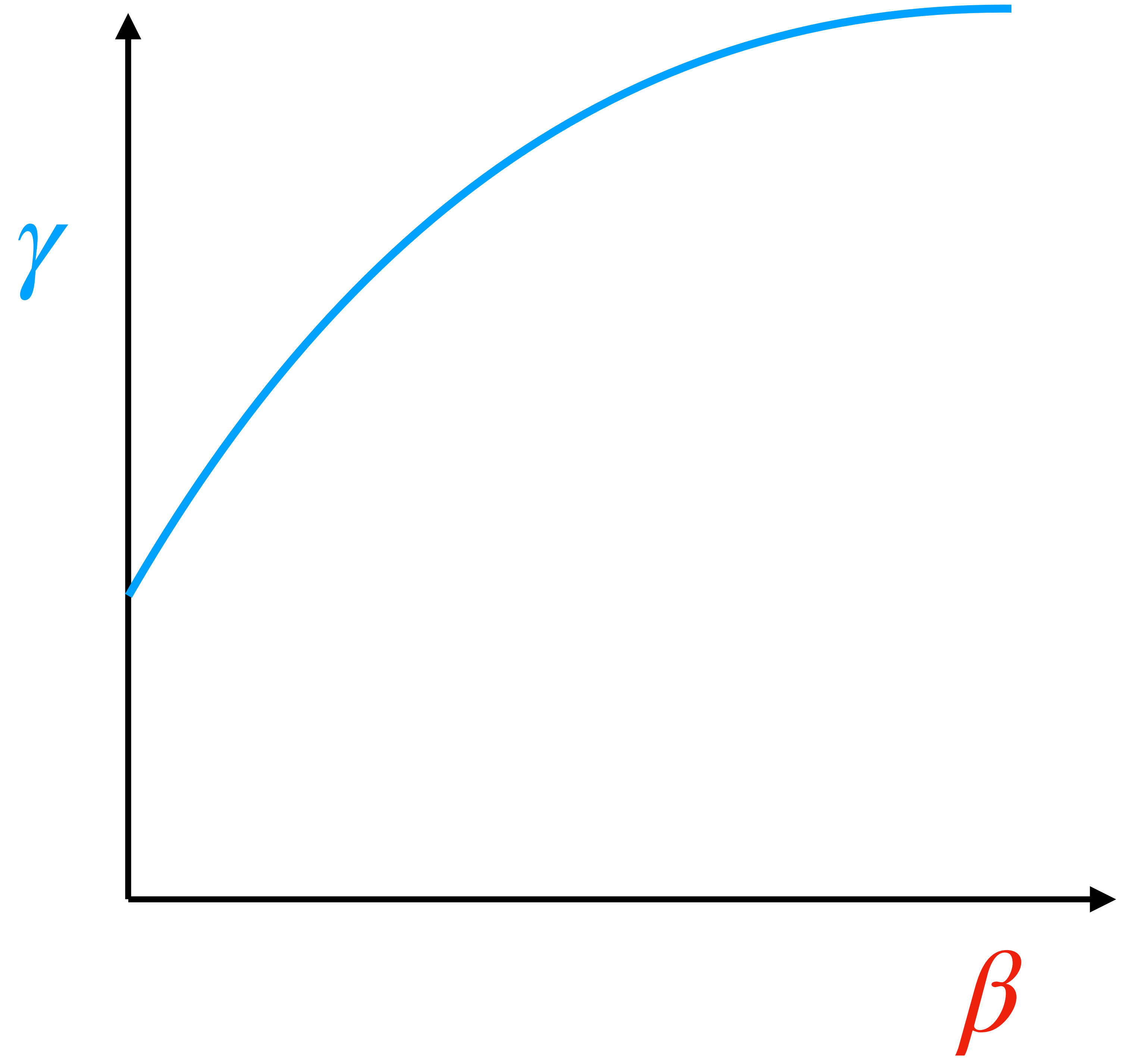
Maximize DICUT size

Approximation algorithms

- Exact computation of optimum maybe NP-Hard!
- $0 < \alpha < 1$, α approximation algorithm:
  - outputs T such that: $\alpha \cdot \text{OPT} \leq T \leq \text{OPT}$
  - outputs an "underestimate" that is not off by a factor more than α
  - Randomized algorithm: outputs such an estimate with probability at least $2/3$

Promise problems

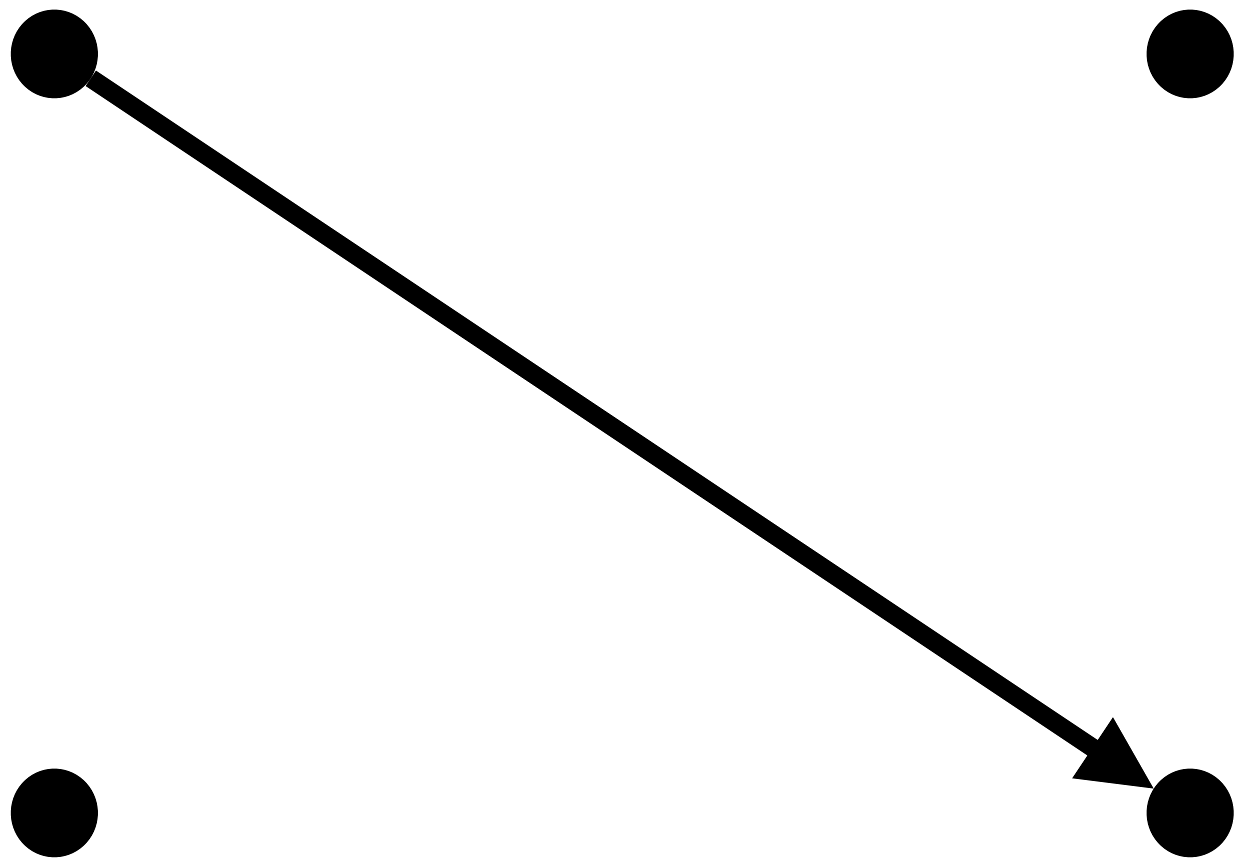
- For $\beta < \gamma$, can you distinguish instances with OPT $\geq \gamma$ from OPT $\leq \beta$?
- Finer study of approximation:
  - $\alpha = \inf_{\text{indistinguishable } (\gamma,\beta)} \beta/\gamma$



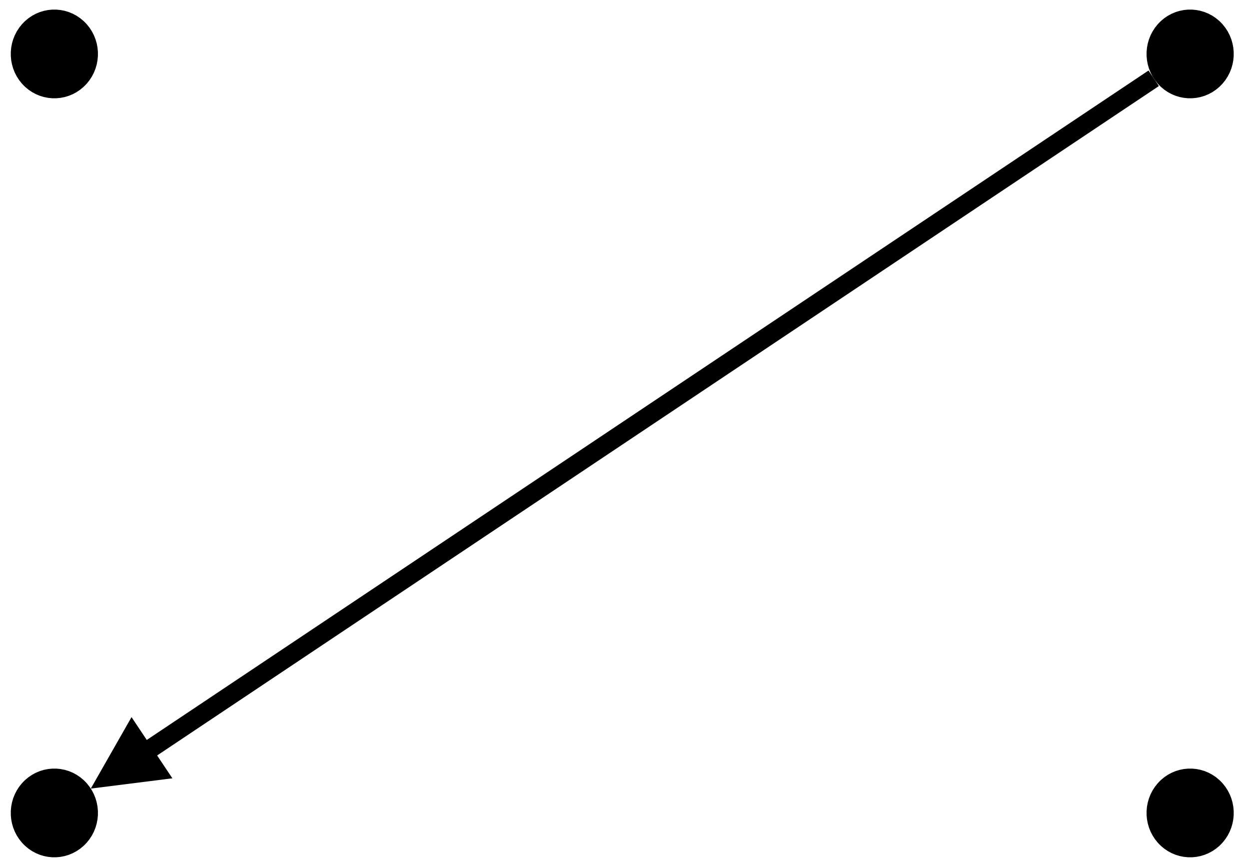
Approximability curve

Streaming setting

Streaming setting



Streaming setting



Streaming setting

Streaming setting

Compute the size of the largest CUT/DICUT in small space

Folklore approximations

- Random CUT achieves $1/2$ approximation for Max-CUT
  - [Kapralov-Krachun'19] Beating this requires at least $\Omega(n)$ space!
- Random DICUT achieves $1/4$ approximation for Max-DICUT
  - We can beat this in $\log(n)$ space
- Sample $O(n/\epsilon^2)$ random edges and compute Max-CUT/Max-DICUT value (possibly in exponential time) to obtain $(1-\epsilon)$ approximation!

Streaming Approximability of CSPs in $o(n)$ space

Streaming Approximability of CSPs in $o(n)$ space

(bounded-degree instances)

Dichotomy theorem: $\log n$ vs $\sqrt{n}$

[Chou-Golovnev-Sudan-V'21]

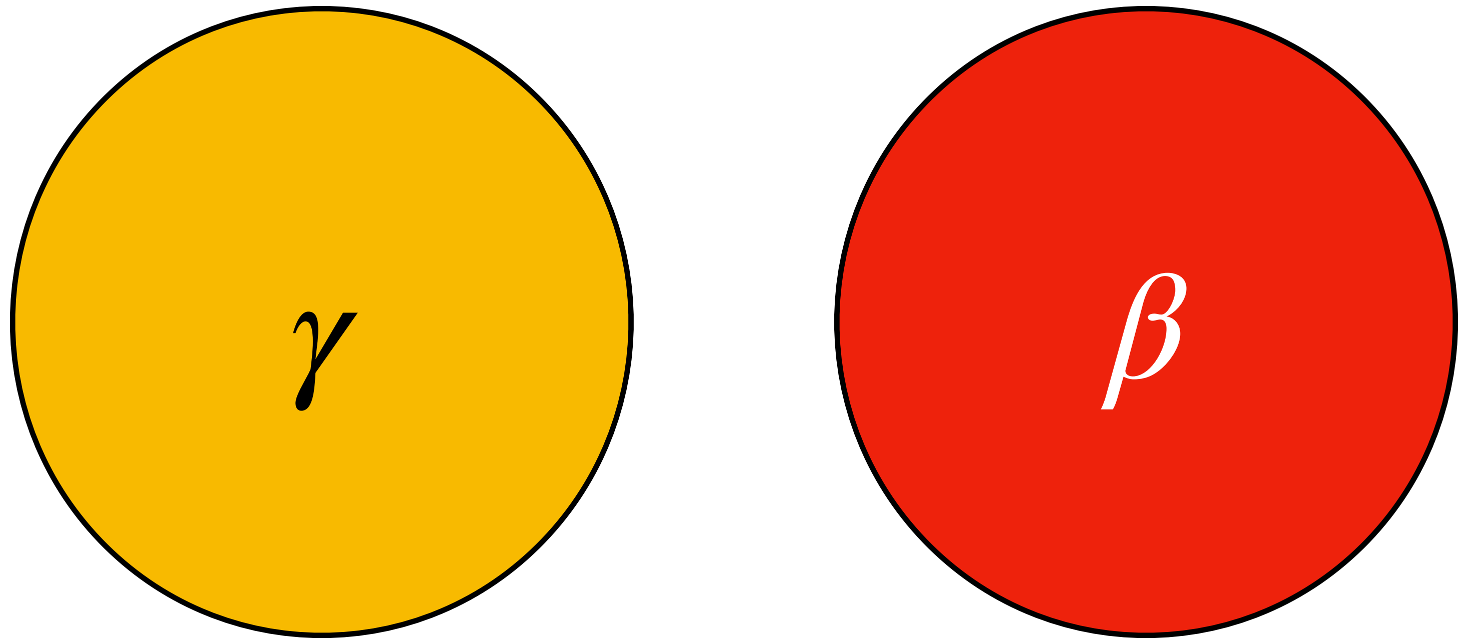
- Bias of a vertex $v = \dfrac{\text{out}(v) - \text{in}(v)}{\text{deg}(v)} \in [-1,1]$
- Bias intervals: $-1 \leq b_0 \leq \cdots \leq b_{r-1} \leq 1$ where $b_{i+1}/b_i \leq \epsilon$
- Vertex-bias statistics: number of vertices of of degree $i \in [d]$ in each bias class
- Can be computed in $O(\log n)$ space

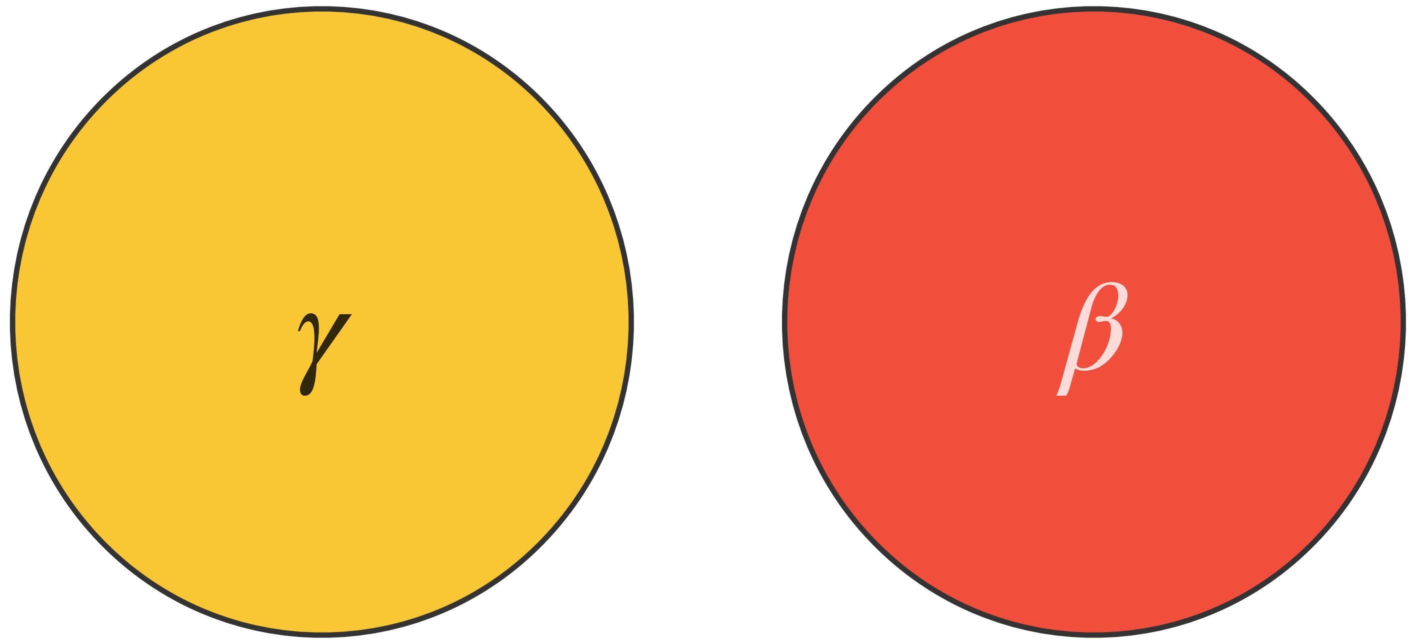
Dichotomy theorem: $\log n$ vs $\sqrt{n}$

[Chou-Golovnev-Sudan-V'21]

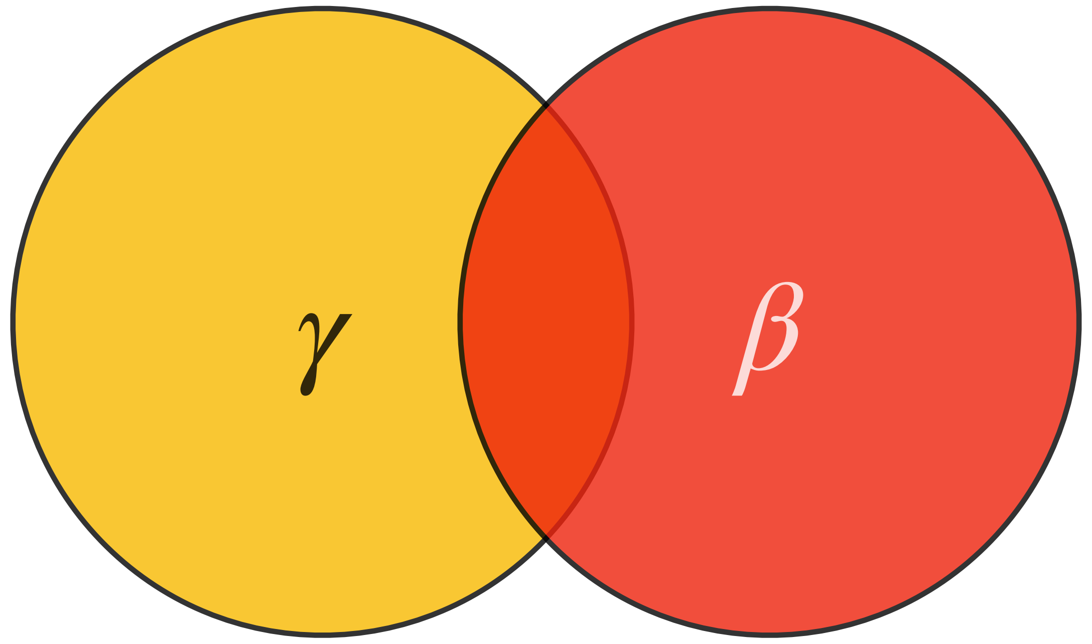
For $\beta < \gamma$, when can you distinguish instances with OPT $\geq \gamma$ from OPT $\leq \beta$?

vertex-bias statistics





distinguishable in $O(\log n)$ space



indistinguishable in $o(\sqrt{n})$ space

Dichotomy theorem: $\log n$ vs $\sqrt{n}$

[Chou-Golovnev-Sudan-V'21]

- Decidable in PSPACE
- Decidable characterization of CSPs that are "approximation resistant" in $o(\sqrt{n})$ space
- Incomplete: only proves $\sqrt{n}$ space lower bounds
- Linear space lower bounds for an infinite subset of approximation resistant CSPs [Chou-Golovnev-Sudan-Velingker-V'22]

$\tilde{O}(\sqrt{n})$ space algorithm for Max-DICUT

[Saxena-Singer—Sudan-V'23]

- Edge-bias statistics
- Compute the number of edges between each pair of bias classes
- Sample $O_d(\sqrt{n})$ vertices and record their induced subgraph
- Oblivious algorithms: "round" all variables in each bias class with same probability [Feige-Jozeph'15]
- Better approximation-ratio for Max-DICUT!
- $\exists \delta > 0$ such that oblivious algorithm cannot beat $1/2 - \delta$ approximation

Reaching $1/2$-approximation

[Saxena-Singer—Sudan-V'upcoming]

- Radius-k-subgraph statistics $\implies 1/2 - 1/k^2$ approximation
- Computable in $\tilde{O}(n^{1-f(d,k)})$ space
- Sample $O(n^{1-f(d,k)})$ vertices and record their induced subgraph

Open problems

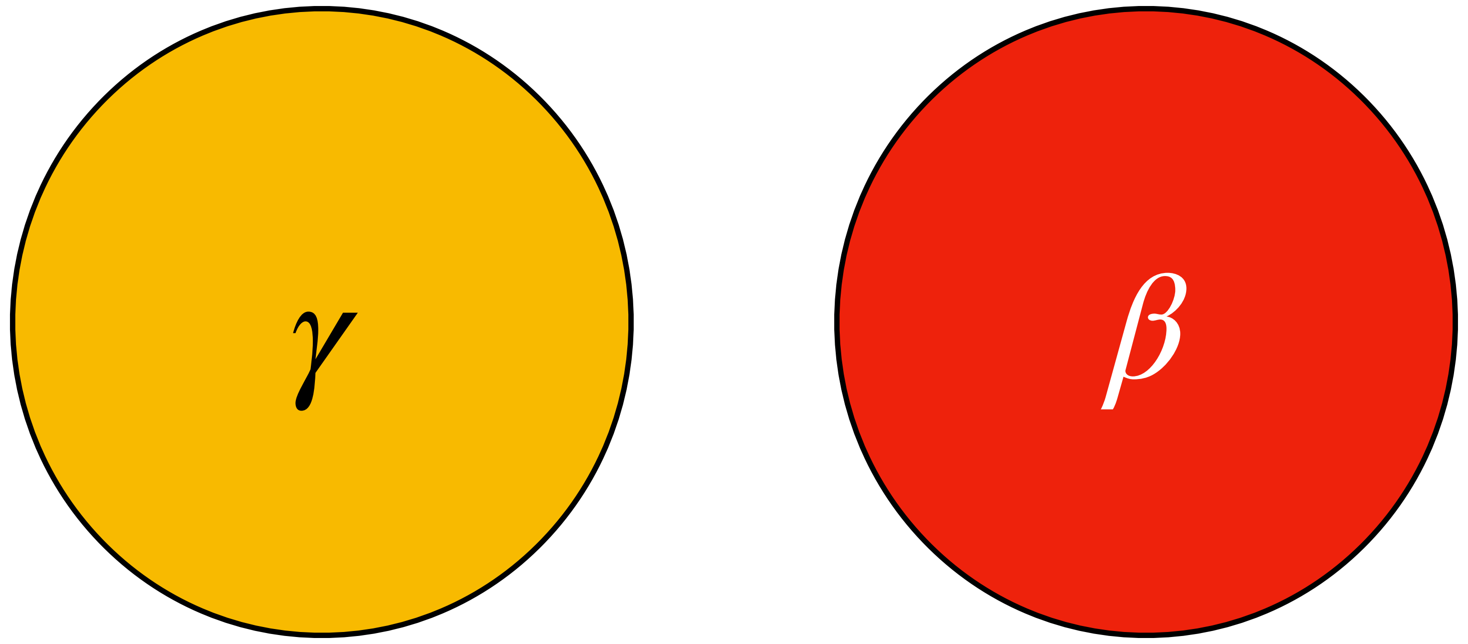
Do edge-bias statistics give the best algorithm for Max-DICUT in $\tilde{O}(\sqrt{n})$ space?

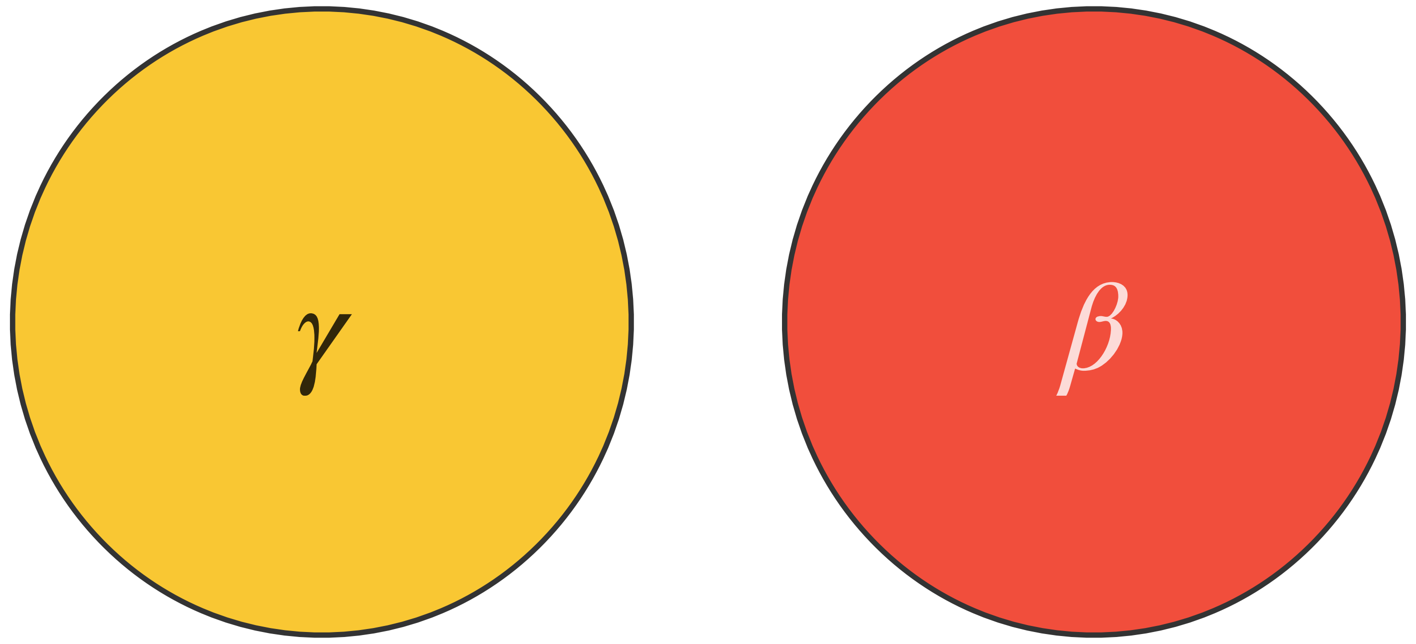
Do edge-bias statistics give the best algorithm for Max-CSP in $\tilde{O}(\sqrt{n})$ space?

Dichotomy theorem??

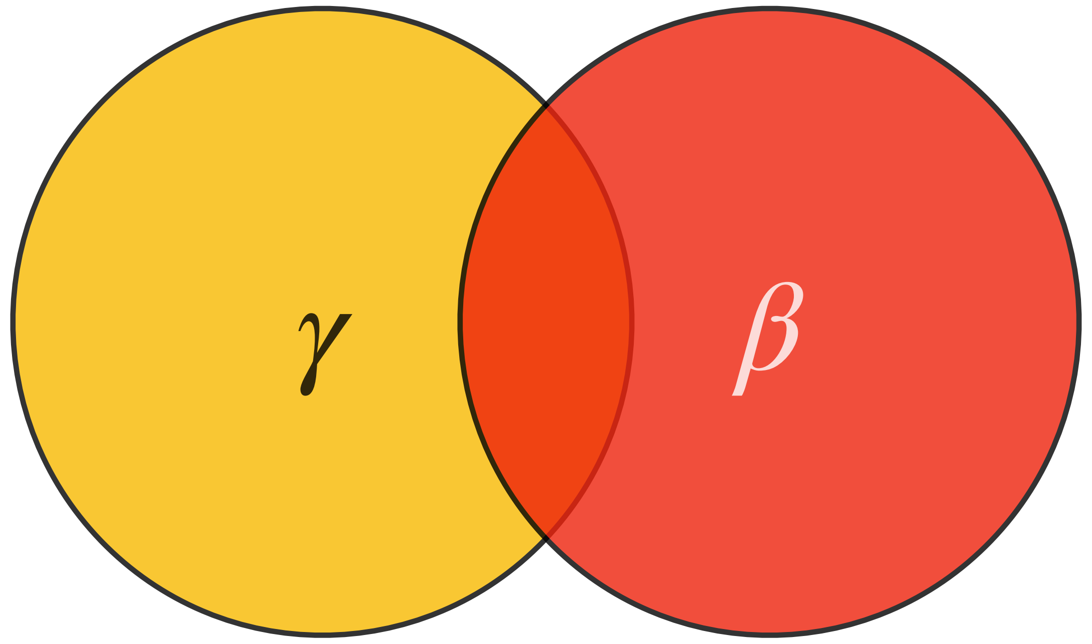
For $\beta < \gamma$, when can you distinguish instances with OPT $\geq \gamma$ from OPT $\leq \beta$?

edge-bias statistics





distinguishable in $\tilde{O}(\sqrt{n})$ space



indistinguishable in $o(n^c)$ space, $c > 1/2$

Is there a CSP that is approximation resistant in $o(\sqrt{n})$ space but approximable in $o(n)$ space?

Is every CSP that is approximation resistant in P also approximation resistant in $o(\sqrt{n})$ space?

Thank you!