A sampling Lovász local lemma for large domain sizes

Chunyang Wang (Nanjing University)
Joint work with: Yitong Yin (Nanjing University)

FOCS 2024

(Atomic) Constraint Satisfaction Problem

$$\Phi = (V, \boldsymbol{Q}, \mathscr{C})$$

Variables: $V = \{v_1, v_2, \ldots, v_n\}$ with finite domains Q_v for each $v \in V$

Constraints: $\mathscr{C} = \{c_1, c_2, \ldots, c_m\}$ with each $c \in \mathscr{C}$ defined on $\mathsf{vbl}(c) \subseteq V$

$$c : \bigotimes_{v \in \mathsf{vbl}(c)} Q_v \to \{\text{True}, \text{False}\}$$

CSP solution: assignment $X \in \bigotimes_{v \in V} Q_v$ s.t. all constraints evaluate to True

(Atomic) Constraint Satisfaction Problem

$$\Phi = (V, \boldsymbol{Q}, \mathscr{C})$$

Variables: $V = \{v_1, v_2, \ldots, v_n\}$ with finite domains Q_v for each $v \in V$

Constraints: $\mathscr{C} = \{c_1, c_2, \ldots, c_m\}$ with each $c \in \mathscr{C}$ defined on $\text{vbl}(c) \subseteq V$

$$c : \bigotimes_{v \in \text{vbl}(c)} Q_v \to \{\text{True}, \text{False}\}$$

Atomic: $|\text{False}^{-1}(c)| = 1$ for each $c \in \mathscr{C}$

CSP solution: assignment $\boldsymbol{X} \in \bigotimes_{v \in V} Q_v$ s.t. all constraints evaluate to True

(Atomic) Constraint Satisfaction Problem

$$\Phi = (V, \boldsymbol{Q}, \mathscr{C})$$

Variables: $V = \{v_1, v_2, \ldots, v_n\}$ with finite domains Q_v for each $v \in V$

Constraints: $\mathscr{C} = \{c_1, c_2, \ldots, c_m\}$ with each $c \in \mathscr{C}$ defined on $\mathsf{vbl}(c) \subseteq V$

$c : \bigotimes_{v \in \mathsf{vbl}(c)} Q_v \to \{\text{True}, \text{False}\}$ **Atomic:** $|\text{False}^{-1}(c)| = 1$ for each $c \in \mathscr{C}$

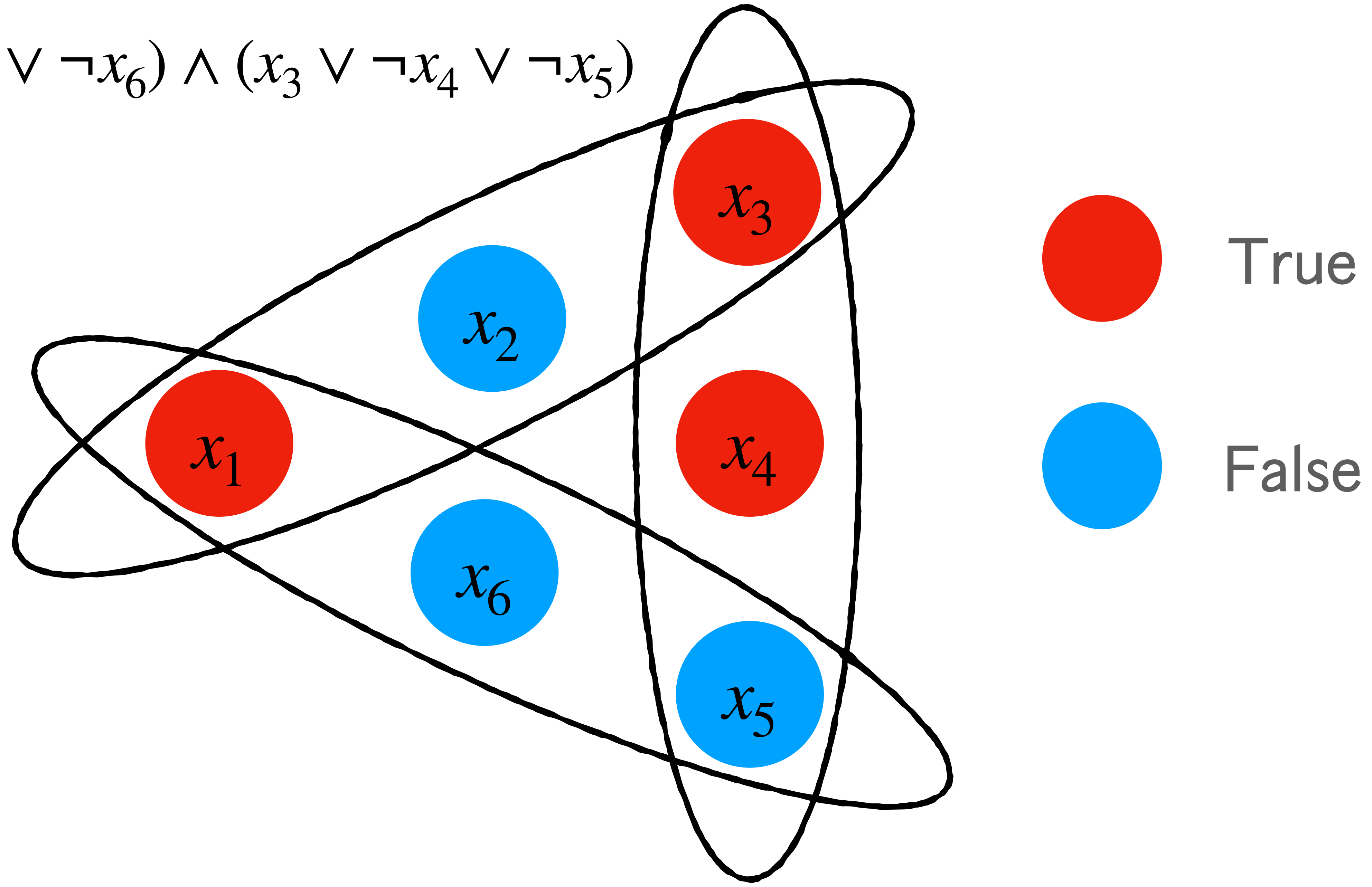
CSP solution: assignment $\boldsymbol{X} \in \bigotimes_{v \in V} Q_v$ s.t. all constraints evaluate to True

Decision: Can we efficiently decide if Φ has a solution?

Search: Can we efficiently find a solution of Φ?

Sampling: Can we efficiently sample an (almost) uniform random solution of Φ?

$$\Phi = (x_1 \vee \neg x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_5 \vee \neg x_6) \wedge (x_3 \vee \neg x_4 \vee \neg x_5)$$

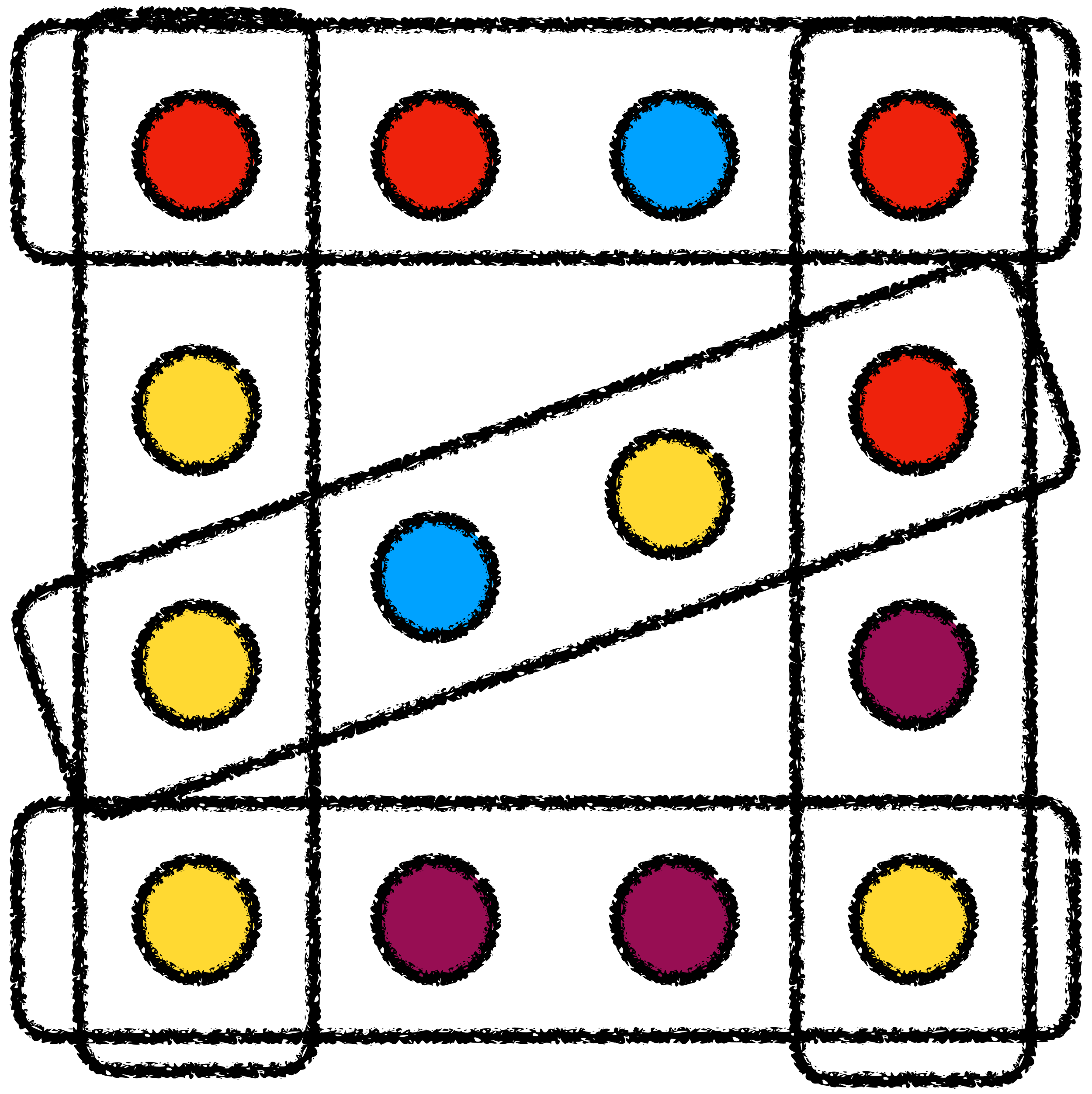


Example: hypergraph q-coloring

k-uniform hypergraph $H = (V, \mathscr{E})$

color set $[q]$ for each $v \in V$

Solution: an assignment such that no hyperedge (constraint) is monochromatic



Lovász Local Lemma

$$\Phi = (V, \boldsymbol{Q}, \mathscr{C})$$

Variable framework

- each $v \in V$ draws from Q_v uniformly and independently at random
- product distribution $\mathscr{P}$

Parameters

- **violation probability** $p = \max_{c \in \mathscr{C}} \Pr_{\mathscr{P}}[\neg c]$
- **dependency degree** $D = \max_{c \in \mathscr{C}} |\{c' \in \mathscr{C} \setminus \{c\} \mid \mathsf{vbl}(c) \cap \mathsf{vbl}(c') \neq \emptyset\}|$

Lovász Local Lemma

$$\Phi = (V, Q, \mathscr{C})$$

Variable framework

- each $v \in V$ draws from Q_v uniformly and independently at random
- product distribution $\mathscr{P}$

Parameters

- violation probability $p = \max_{c \in \mathscr{C}} \Pr_{\mathscr{P}}[\neg c]$
- dependency degree $D = \max_{c \in \mathscr{C}} |\{c' \in \mathscr{C} \setminus \{c\} \mid \mathsf{vbl}(c) \cap \mathsf{vbl}(c') \neq \emptyset\}|$

$$ep(D+1) \leq 1$$

Lovász Local Lemma [Erdos, Lovász '75]

Algorithmic Lovász Local Lemma [Moser, Tardos '10]

A CSP solution exists and can be efficiently found!

Sampling **L**ovász **L**ocal **L**emma

Sampling LLL

Input: a CSP formula $\Phi = (V, \boldsymbol{Q}, \mathscr{C})$ under LLL-like conditions $pD^c \lesssim 1$

Output: an (almost) uniform satisfying solution of Φ

Sampling Lovász Local Lemma

Sampling LLL

Input: a CSP formula $\Phi = (V, \boldsymbol{Q}, \mathscr{C})$ under LLL-like conditions $pD^c \lesssim 1$ → [BGGGS19,GGW22]: **NP-hard** if $pD^2 \gtrsim 1$!

Output: an (almost) uniform satisfying solution of Φ

Sampling Lovász Local Lemma

Sampling LLL

Input: a CSP formula $\Phi = (V, \boldsymbol{Q}, \mathscr{C})$ under LLL-like conditions $pD^c \lesssim 1$

Output: an (almost) uniform satisfying solution of Φ

[BGGGS19,GGW22]:
NP-hard if $pD^2 \gtrsim 1$!

Applications:

Approximate counting CSP solutions (Counting LLL)

Almost Uniform Sampling → **Approximate Counting**

self-reduction
[Jerrum, Valiant, Vazirani 1986]

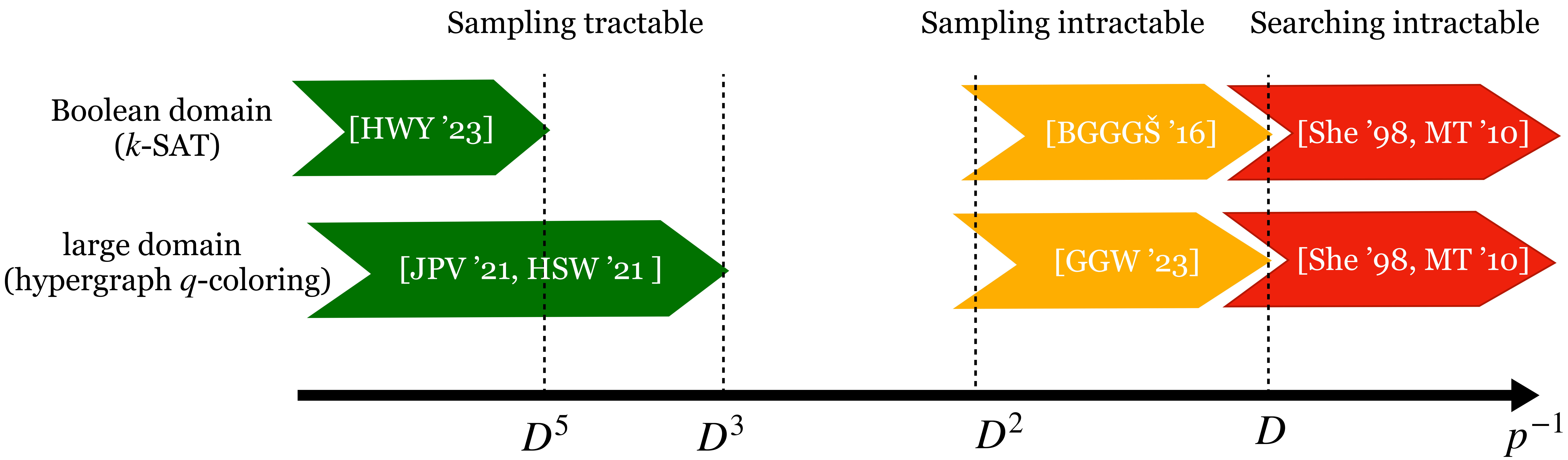
adaptive simulated annealing
[Štefankovič, Vempala, Vigoda 2009]

Inference in probabilistic graphical models

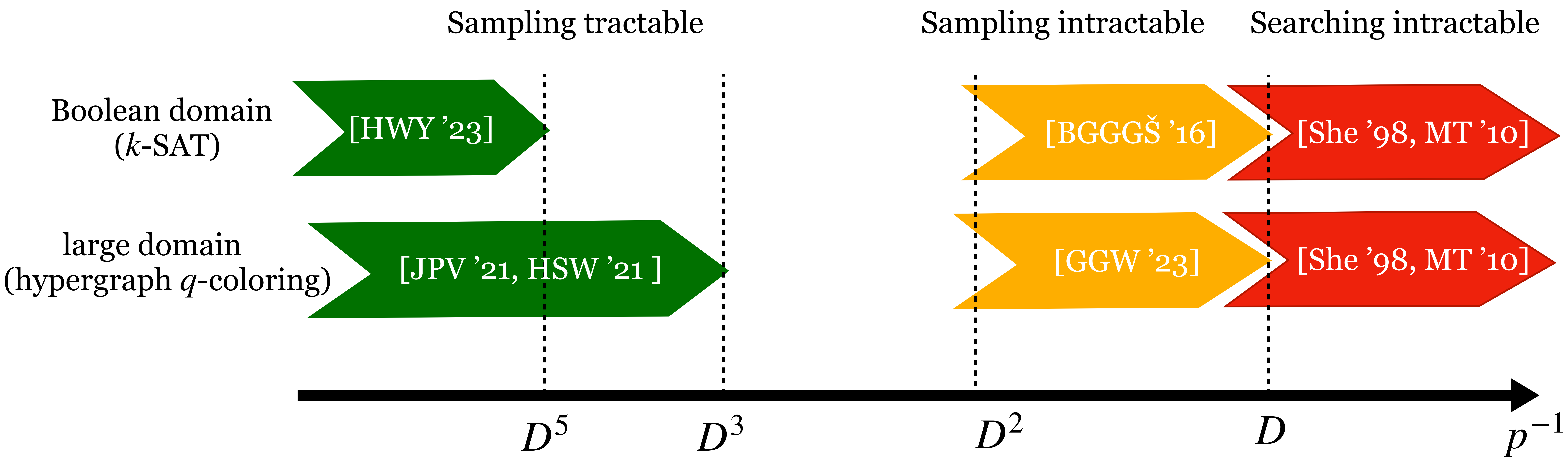
Gibbs distribution μ: uniform distribution over all solutions to Φ

Inference: $\Pr_{X\sim\mu}[X_{v_i} = \cdot \mid X_S = x_s]$

Sampling Lovász Local Lemma

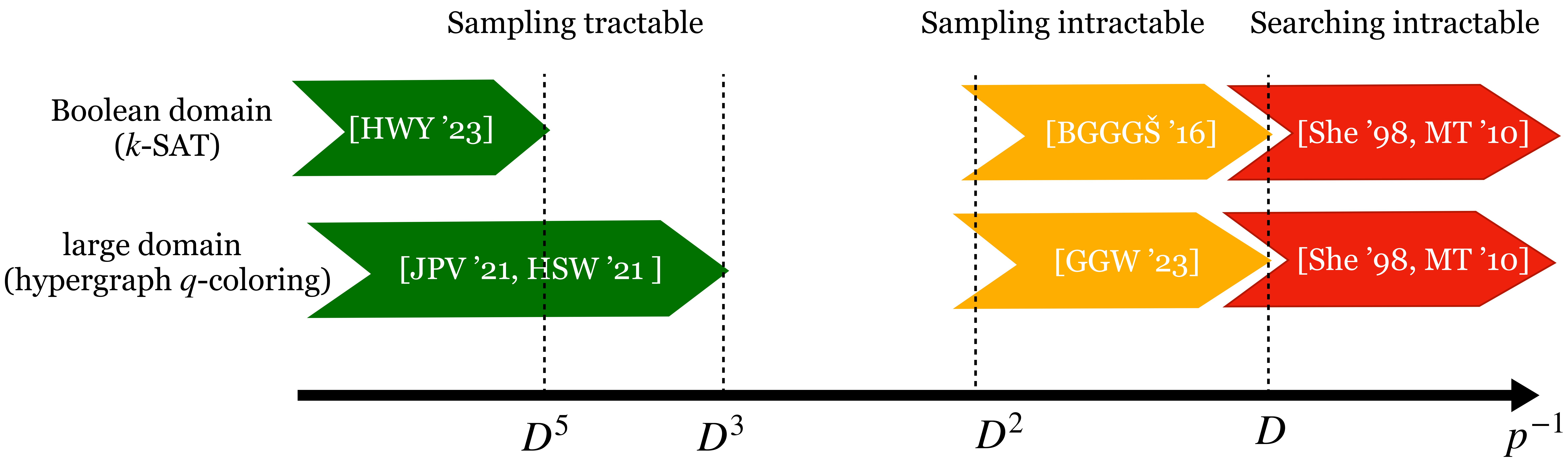


Sampling Lovász Local Lemma



Open problem: Is $pD^2 \lesssim 1$ the correct threshold?

Sampling Lovász Local Lemma



Our result. (sampling/counting atomic CSPs)

We give poly-time (approx) sampling/counting algorithms for atomic CSPs satisfying

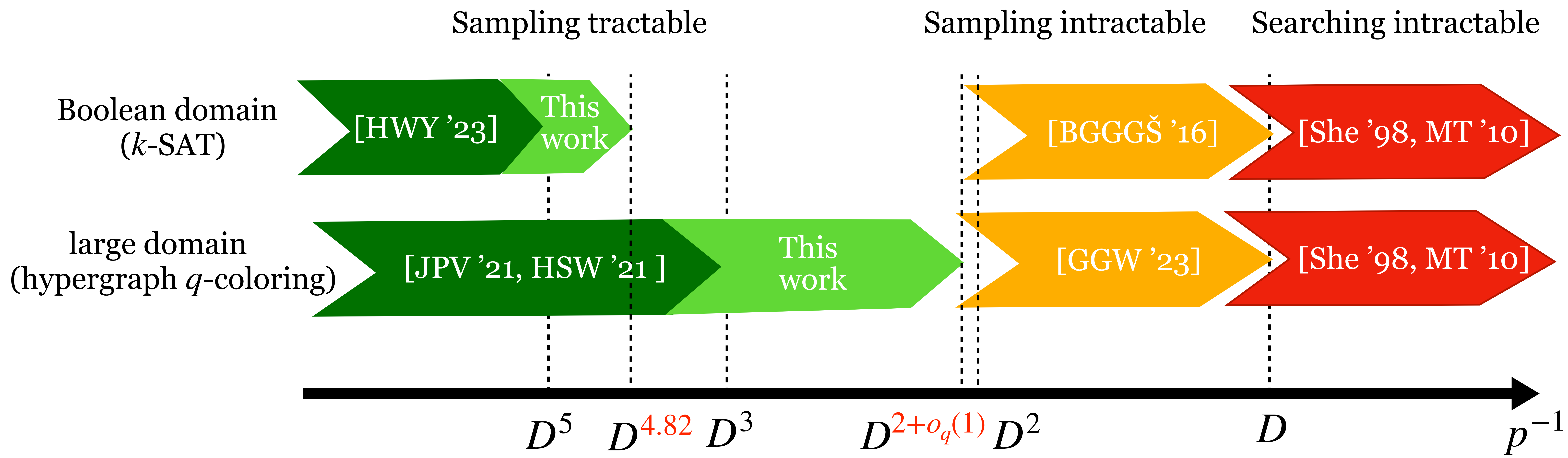
$$(8e)^3 \cdot p \cdot (D+1)^{2+\zeta} \leq 1,$$

where $\zeta = \dfrac{2\ln(2 - 1/q_{\min})}{\ln(q_{\min}) - \ln(2 - 1/q_{\min})}$

$\xi \to 0$ as $q_{\min} \to \infty$!

min domain size $q_{\min} = 2$: $\zeta = 4.82$

Sampling Lovász Local Lemma



Our result. (sampling/counting atomic CSPs)

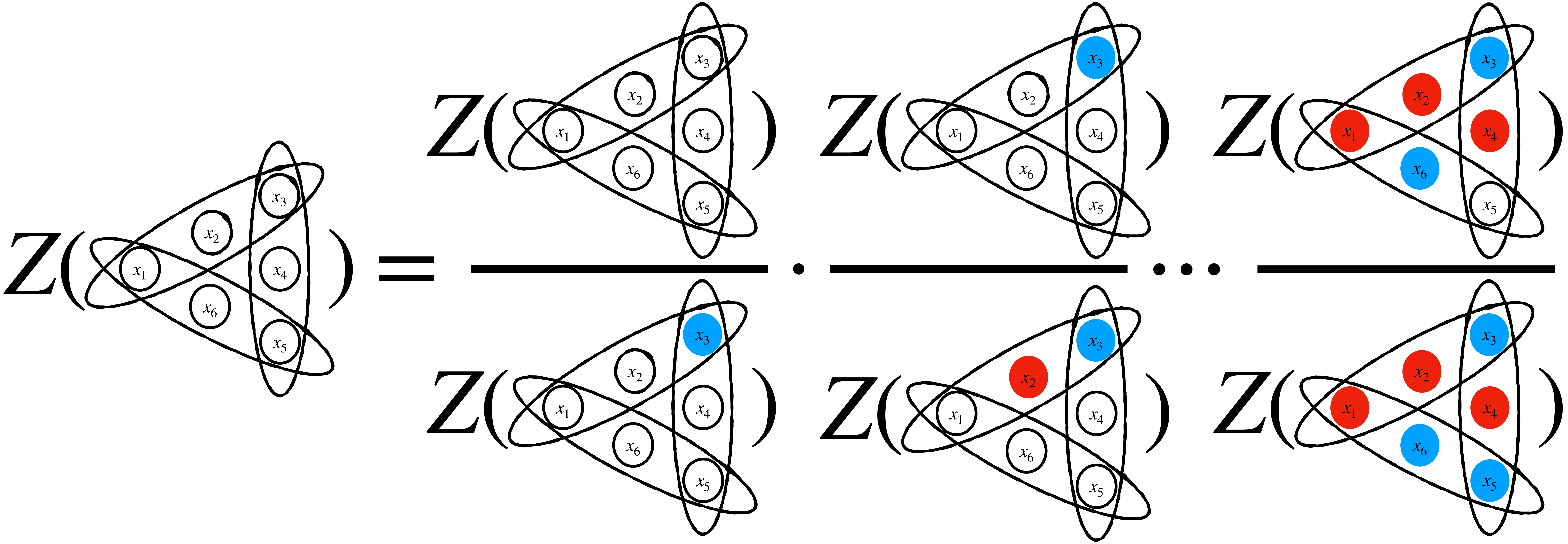
We give poly-time (approx) sampling/counting algorithms for atomic CSPs satisfying

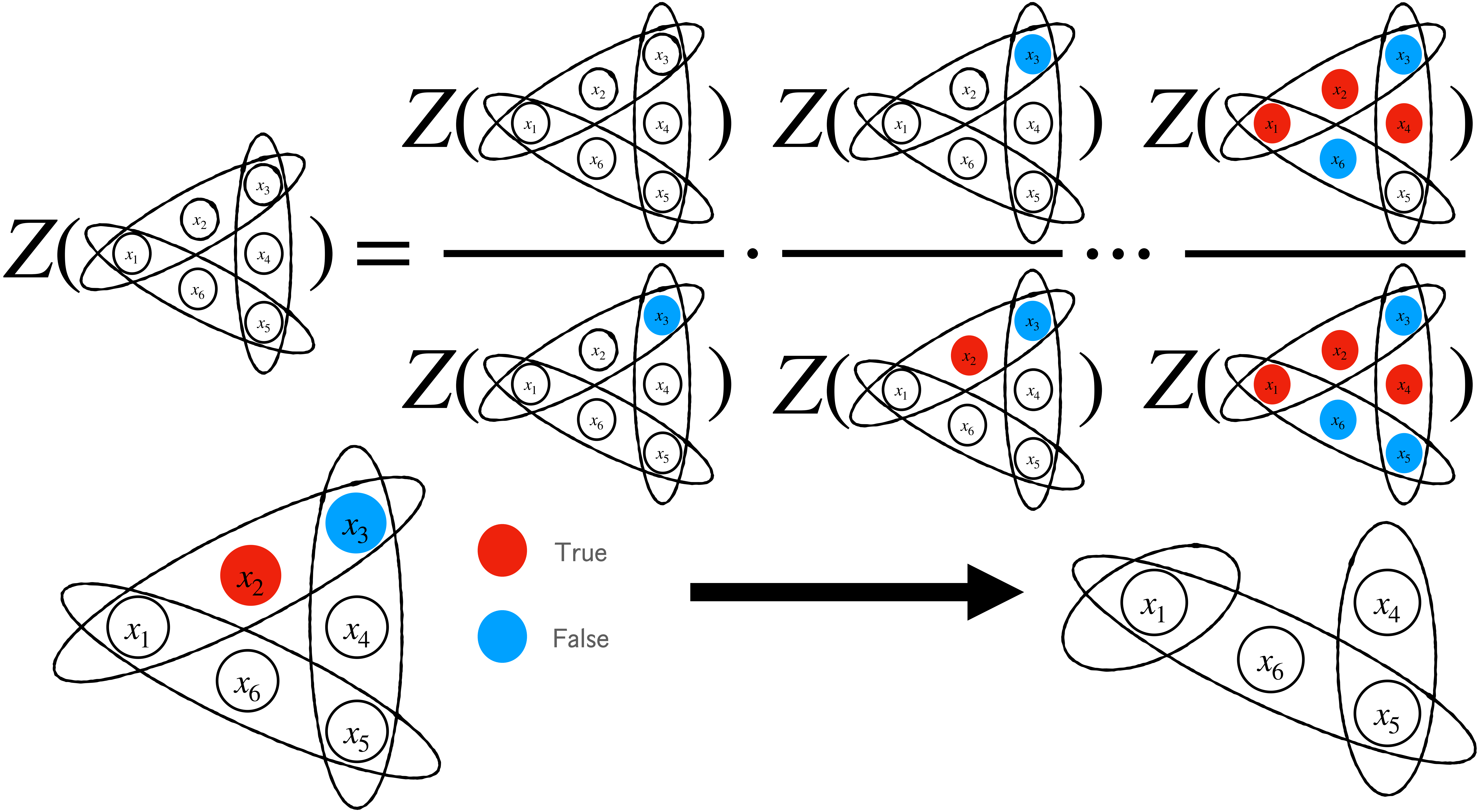
$$(8e)^3 \cdot p \cdot (D+1)^{2+\zeta} \le 1,$$

where $\zeta = \dfrac{2\ln(2-1/q_{\min})}{\ln(q_{\min}) - \ln(2-1/q_{\min})}$

$\xi \to 0$ as $q_{\min} \to \infty$!

min domain size $q_{\min} = 2$: $\zeta = 4.82$


Z
x_1
x_2
x_3
x_4
x_5
x_6



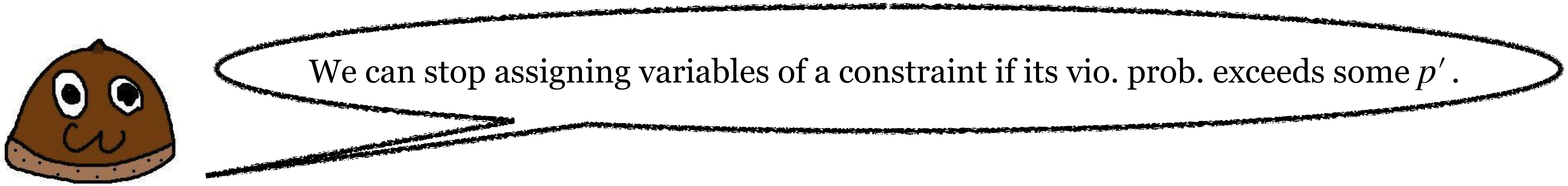
$\Phi = (x_1 \vee \neg x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_5 \vee \neg x_6) \wedge (x_3 \vee \neg x_4 \vee \neg x_5)$

$\Phi' = (x_1) \wedge (\neg x_1 \vee \neg x_5 \vee \neg x_6) \wedge (\neg x_4 \vee \neg x_5)$

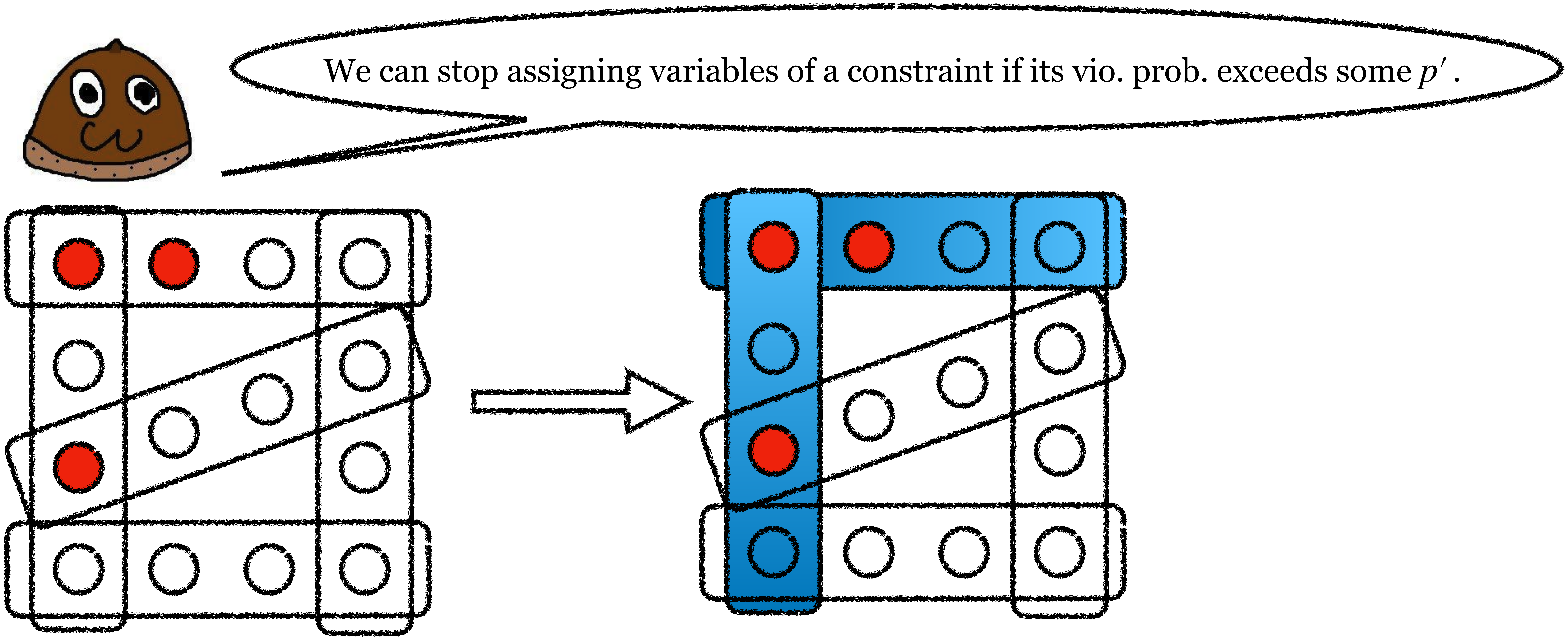
Non self-reducibility: LLL condition may degrade after pinning!

Previous approach: freezing

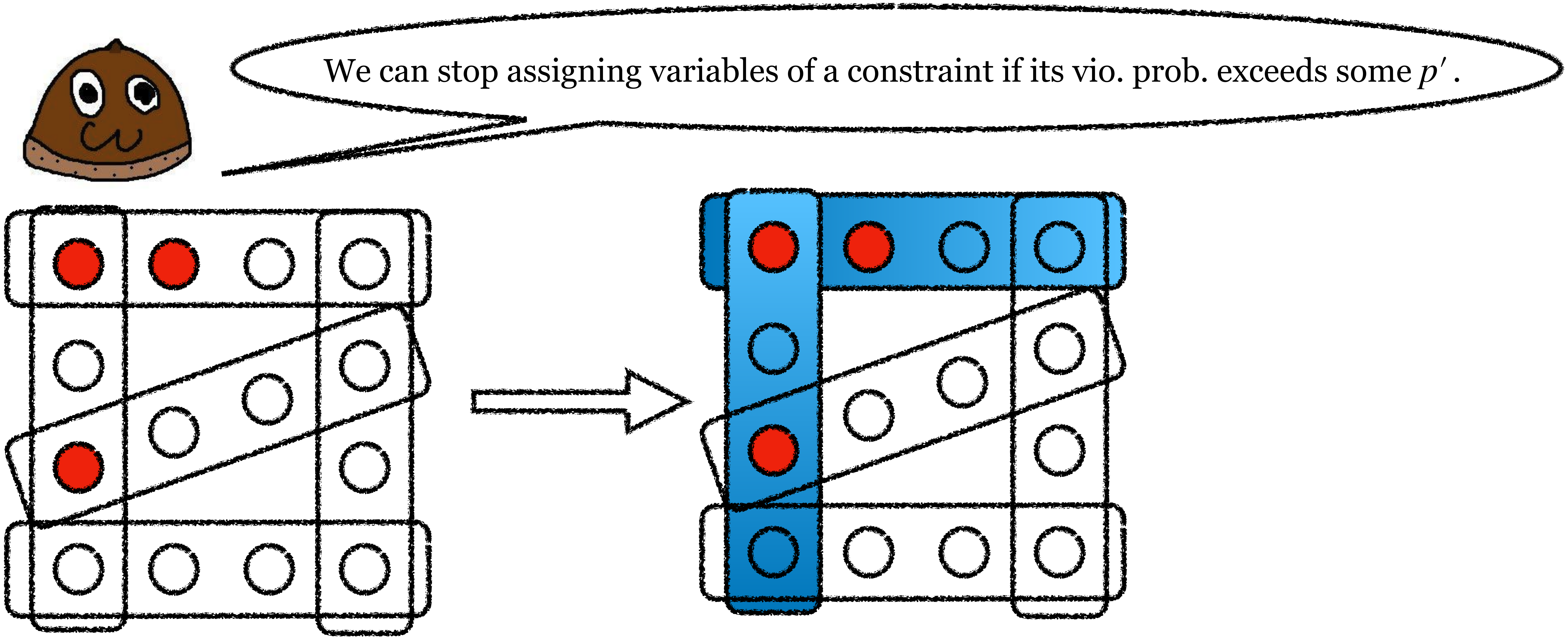
Previous approach: freezing



Previous approach: freezing

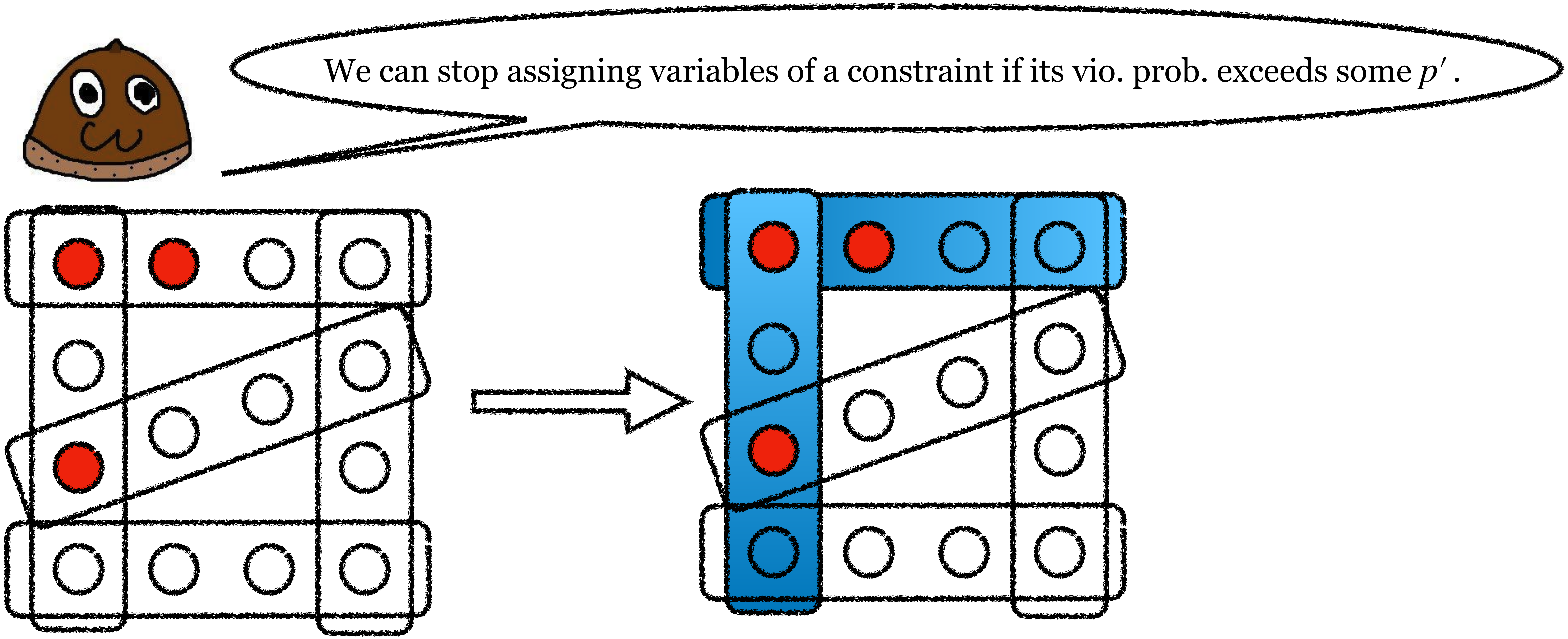


Previous approach: freezing



First use: [Bec '91] for algorithmic LLL, finally lead to $pD^4 \lesssim 1$ [Alon '91, MR '99, Sri '09]

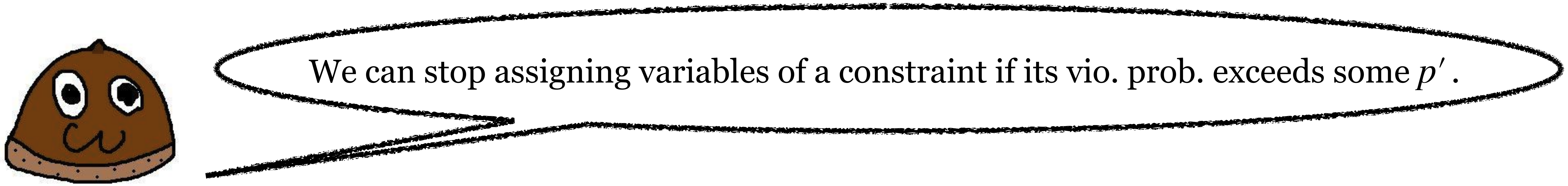
Previous approach: freezing

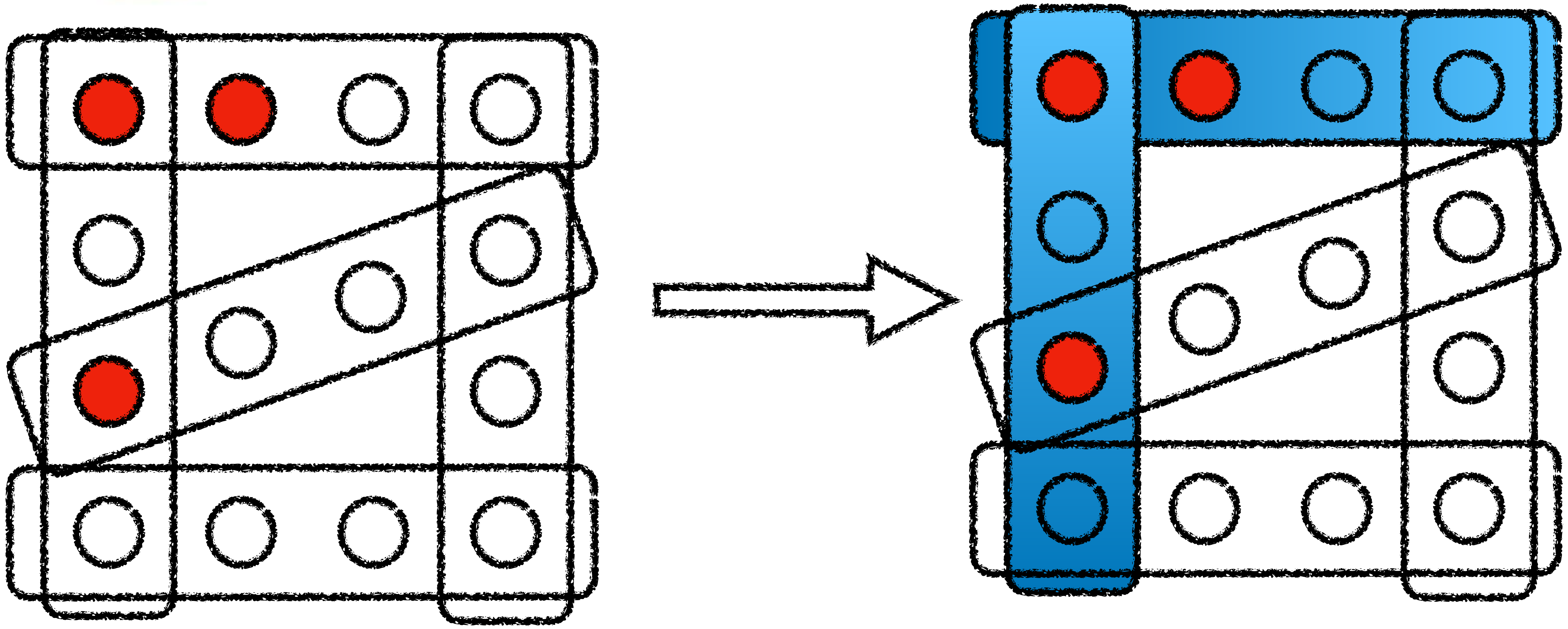


First use: [Bec '91] for algorithmic LLL, finally lead to $pD^4 \lesssim 1$ [Alon '91, MR '99, Sri '09]

In sampling LLL: freezing [JPV '21b, HWY '23]
marking (static variant of freezing) [Moi '19, GLLZ '19, FGYZ '20]
state compression (large domain variant of marking) [FHY '21, JPV '21a ,HSW '21]

Previous approach: freezing





LLL condition: need small p'

"Factorization": need small p/p'

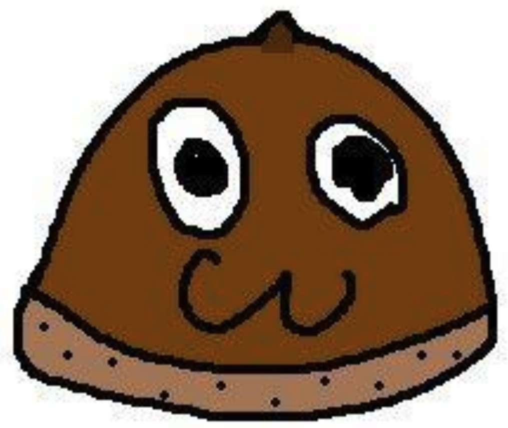
inevitably leads to suboptimal conditions

First use: [Bec '91] for algorithmic LLL, finally lead to $pD^4 \lesssim 1$ [Alon '91, MR '99, Sri '09]

In sampling LLL: freezing [JPV '21b, HWY '23]
marking (static variant of freezing) [Moi '19, GLLZ '19, FGYZ '20]
state compression (large domain variant of marking) [FHY '21, JPV '21a ,HSW '21]

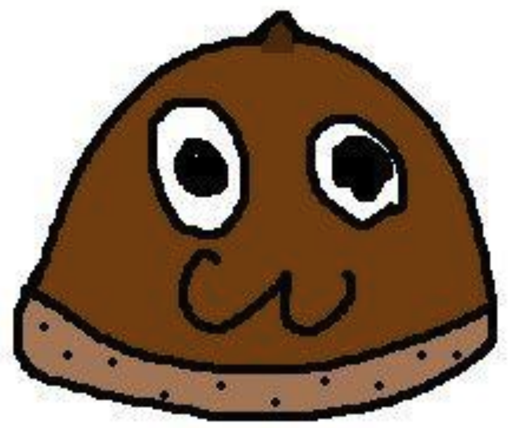
Decay of correlation

Decay of correlation



Dependencies (between variables) decays as the distance grows.

Decay of correlation

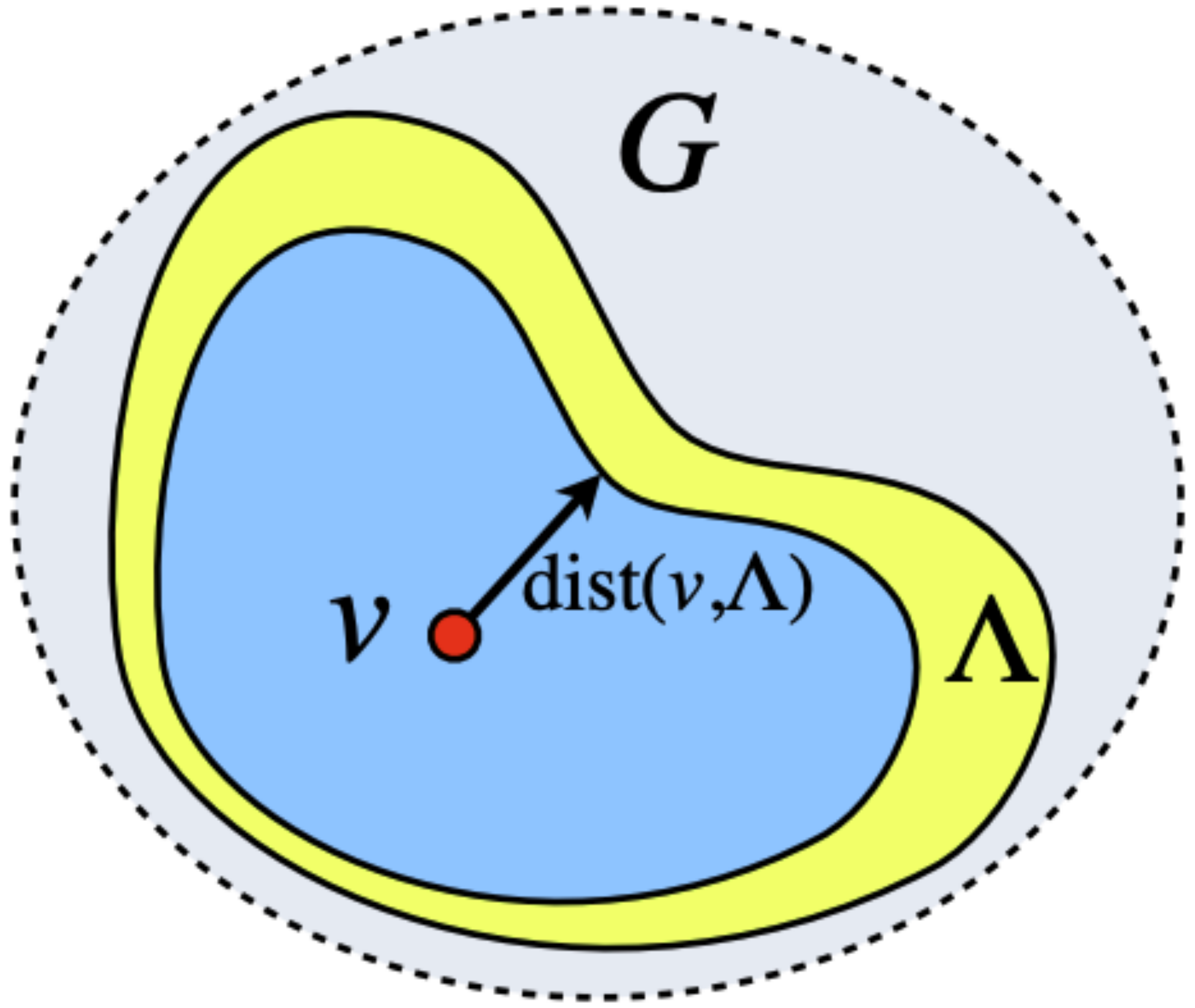


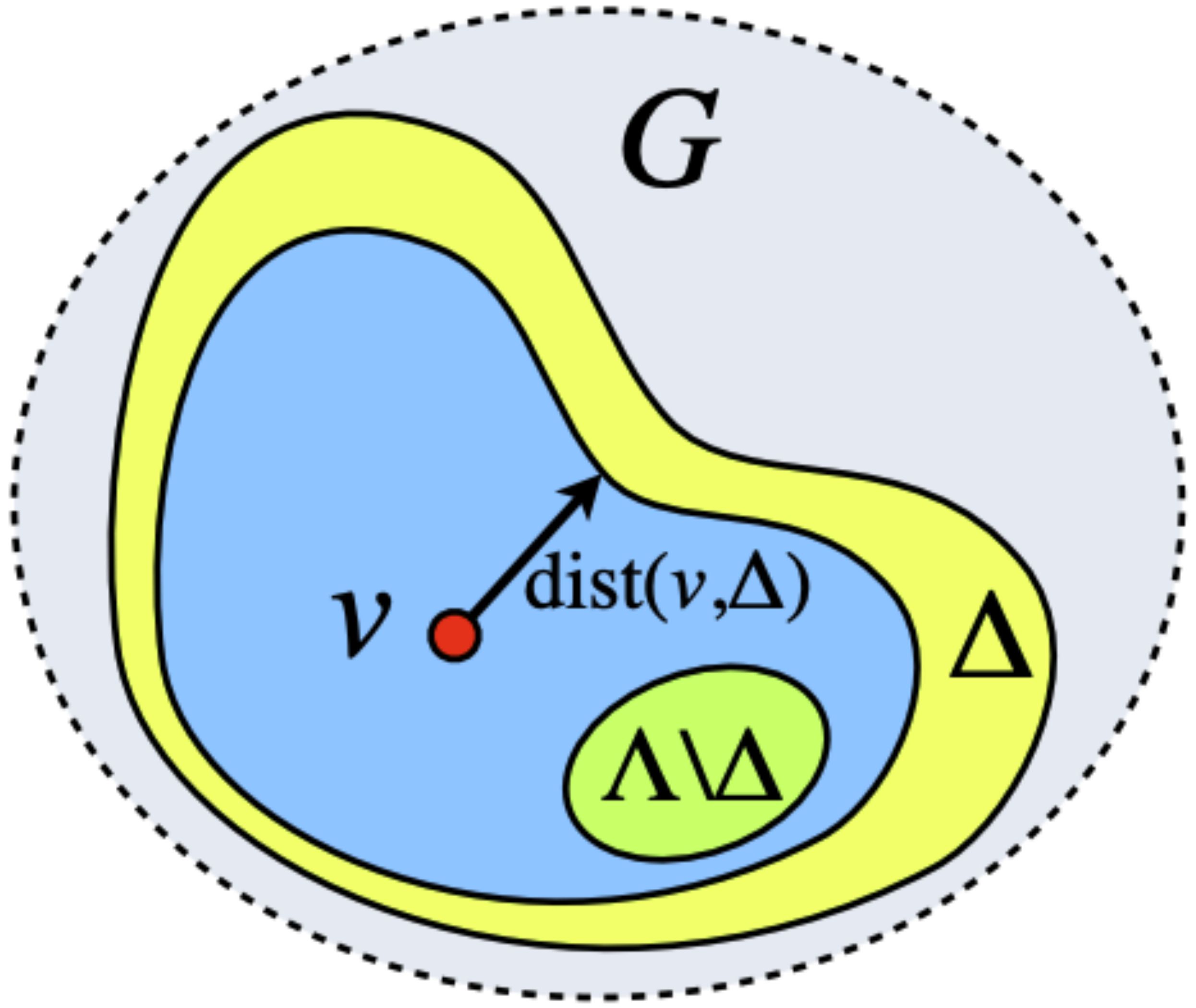
Dependencies (between variables) decays as the distance grows.

μ_v^σ: marginal probability of v conditioning on σ

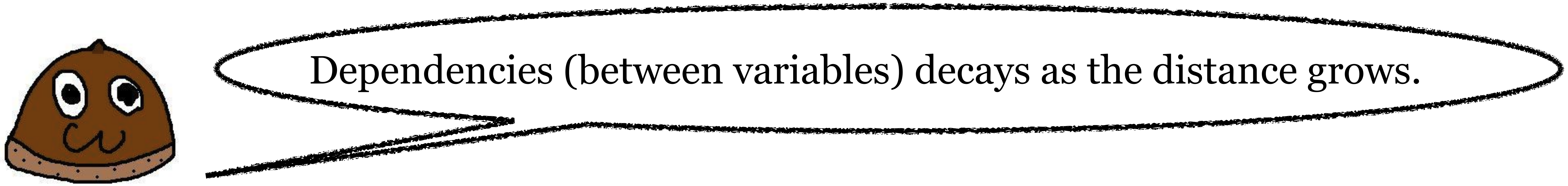
Weak Spatial Mixing (WSM): $\forall \sigma, \tau \in \mathcal{Q}_\Lambda : |\mu_v^\sigma - \mu_v^\tau|_{\mathrm{TV}} \le \delta(\mathrm{dist}_{\mathrm{G}}(\mathrm{v}, \Lambda))$

Strong Spatial Mixing (SSM): $\forall \sigma, \tau \in \mathcal{Q}_\Lambda$ that differ on Δ : $|\mu_v^\sigma - \mu_v^\tau|_{\mathrm{TV}} \le \delta(\mathrm{dist}_{\mathrm{G}}(\mathrm{v}, \Delta))$





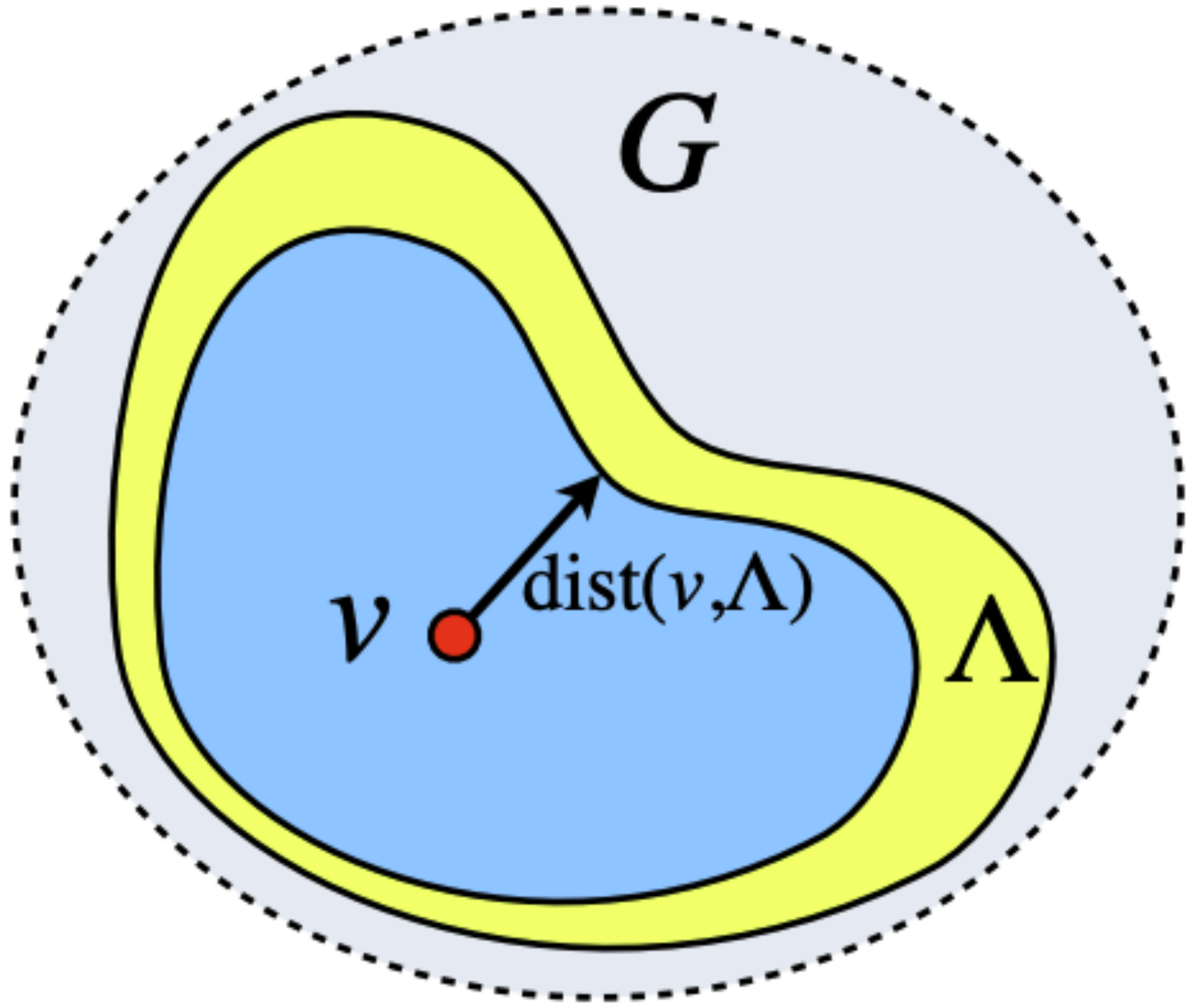
Decay of correlation



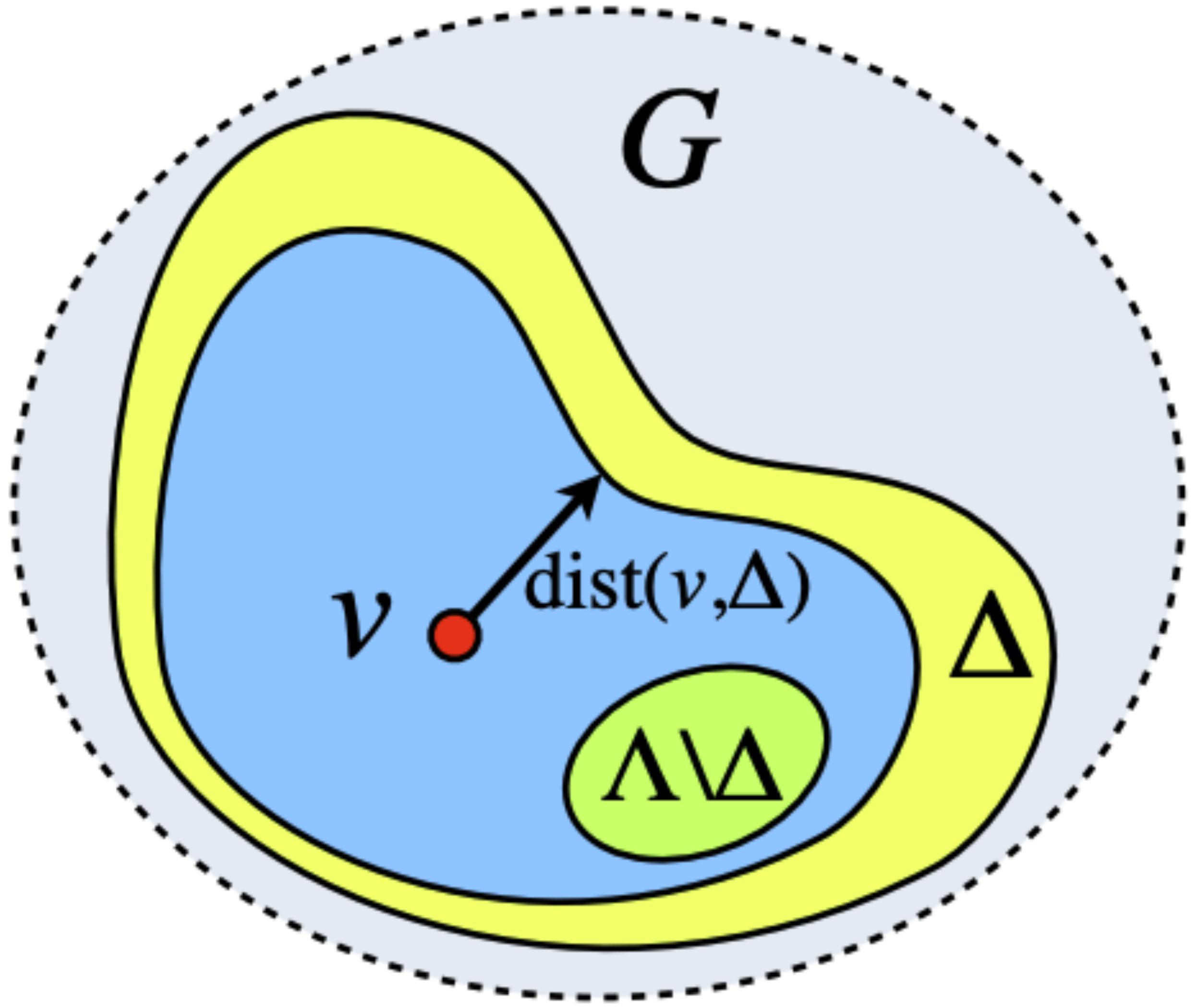
μ_v^σ: marginal probability of v conditioning on σ

Weak Spatial Mixing (WSM): $\forall \sigma, \tau \in \mathcal{Q}_\Lambda : |\mu_v^\sigma - \mu_v^\tau|_{\mathrm{TV}} \leq \delta(\mathrm{dist_G}(\mathrm{v}, \Lambda))$

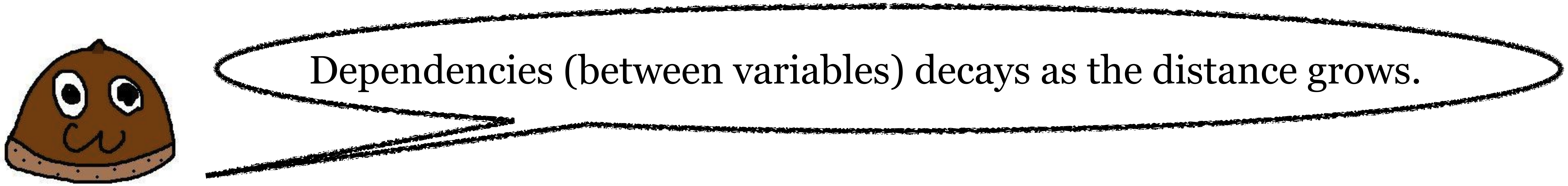
Strong Spatial Mixing (SSM): $\forall \sigma, \tau \in \mathcal{Q}_\Lambda$ that differ on Δ : $|\mu_v^\sigma - \mu_v^\tau|_{\mathrm{TV}} \leq \delta(\mathrm{dist_G}(\mathrm{v}, \Delta))$



Both notions fail for CSPs !



Decay of correlation

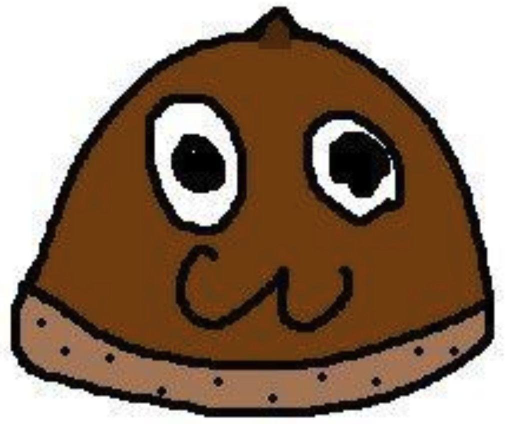


μ_v^σ: marginal probability of v conditioning on σ

Weak Spatial Mixing (WSM): $\forall \sigma, \tau \in \mathcal{Q}_\Lambda : |\mu_v^\sigma - \mu_v^\tau|_{\mathrm{TV}} \leq \delta(\mathrm{dist_G(v, \Lambda)})$

Strong Spatial Mixing (SSM): $\forall \sigma, \tau \in \mathcal{Q}_\Lambda$ that differ on $\Delta : |\mu_v^\sigma - \mu_v^\tau|_{\mathrm{TV}} \leq \delta(\mathrm{dist_G(v, \Delta)})$

Decay of correlation

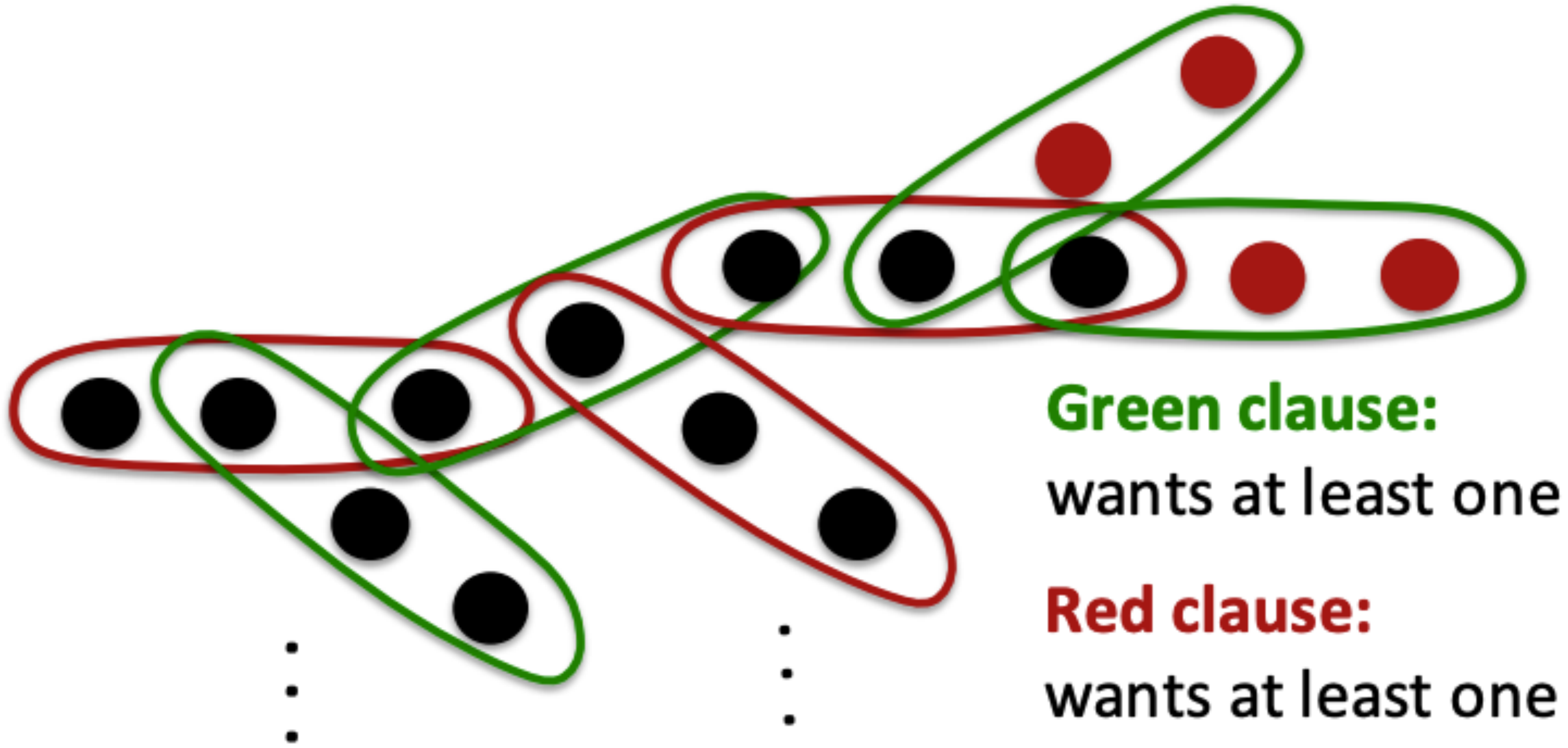


Dependencies (between variables) decays as the distance grows.

μ_v^σ: marginal probability of v conditioning on σ

Weak Spatial Mixing (WSM): $\forall \sigma, \tau \in \mathcal{Q}_\Lambda : |\mu_v^\sigma - \mu_v^\tau|_{\mathrm{TV}} \leq \delta(\mathrm{dist}_G(v, \Lambda))$

Strong Spatial Mixing (SSM): $\forall \sigma, \tau \in \mathcal{Q}_\Lambda$ that differ on $\Delta : |\mu_v^\sigma - \mu_v^\tau|_{\mathrm{TV}} \leq \delta(\mathrm{dist}_G(v, \Delta))$



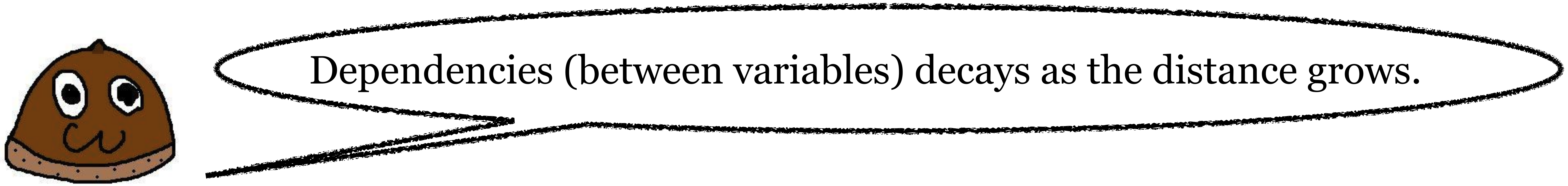
Green clause:
wants at least one green

Red clause:
wants at least one red

long-range dependencies exist
when $D = O(k)$

Credit: Ankur Moitra's
talk at STOC 2017

Decay of correlation



Theorem. (Decay of correlation, informal)

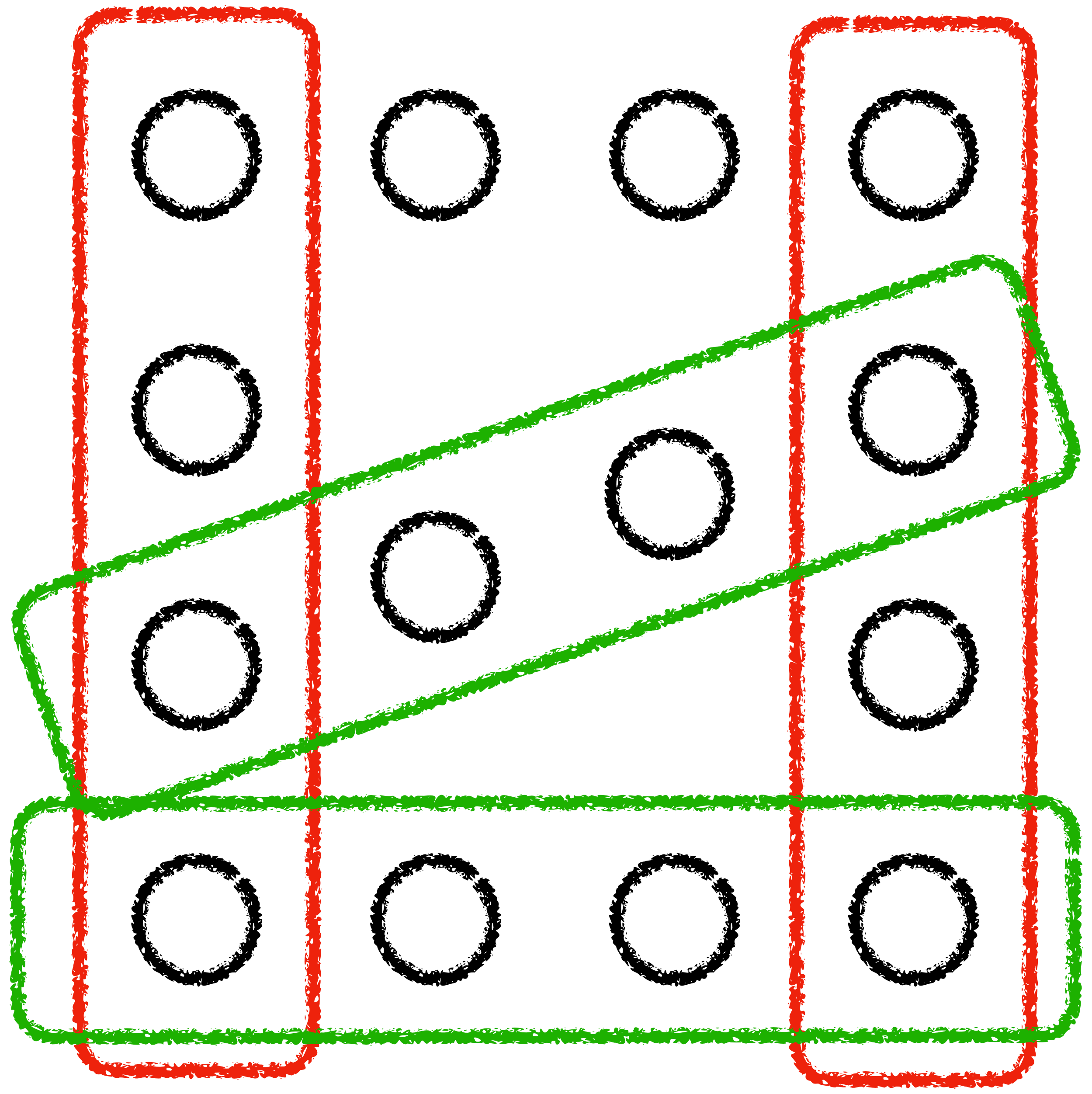
For two CSPs $(V, \mathcal{Q}, \mathscr{C})$ and $(V, \mathcal{Q}, \mathscr{C} \setminus \{c_0\})$ (differ in one constraint) under our condition, there exists a coupling (X, Y) of $\mu_{\mathscr{C} \setminus \{c_0\}}$ and $\mu_{\mathscr{C}}$ such that

$$\Pr[d_{\mathrm{Ham}}(X, Y) \geq K] \leq \exp(-O(K)) .$$

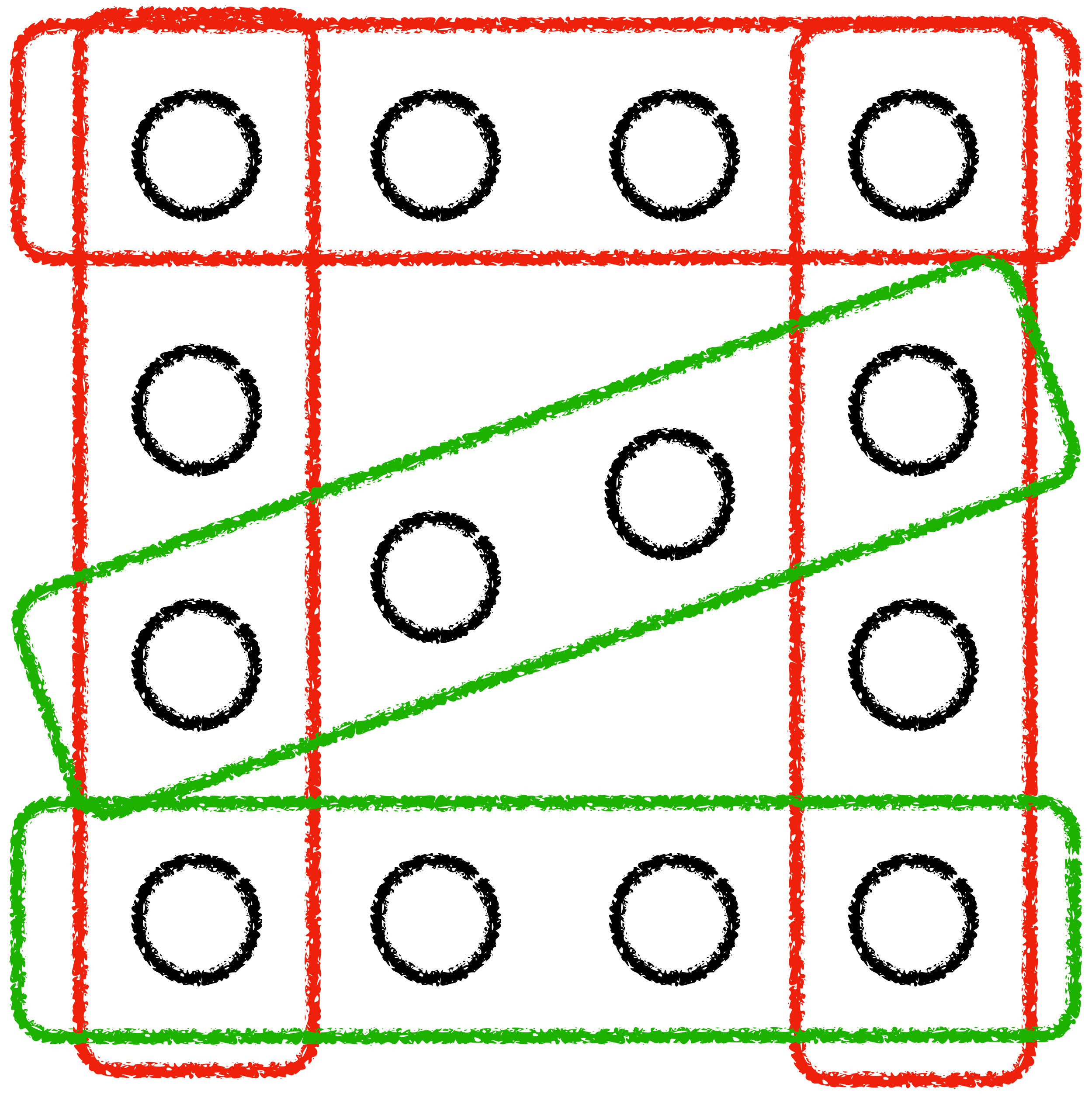
$\mu_{\mathscr{C} \setminus \{c_0\}}$: uniform distribution over solutions of $(V, \mathcal{Q}, \mathscr{C})$

$\mu_{\mathscr{C}}$: uniform distribution over solutions of $(V, \mathcal{Q}, \mathscr{C} \setminus \{c_0\})$

A constraint-wise coupling



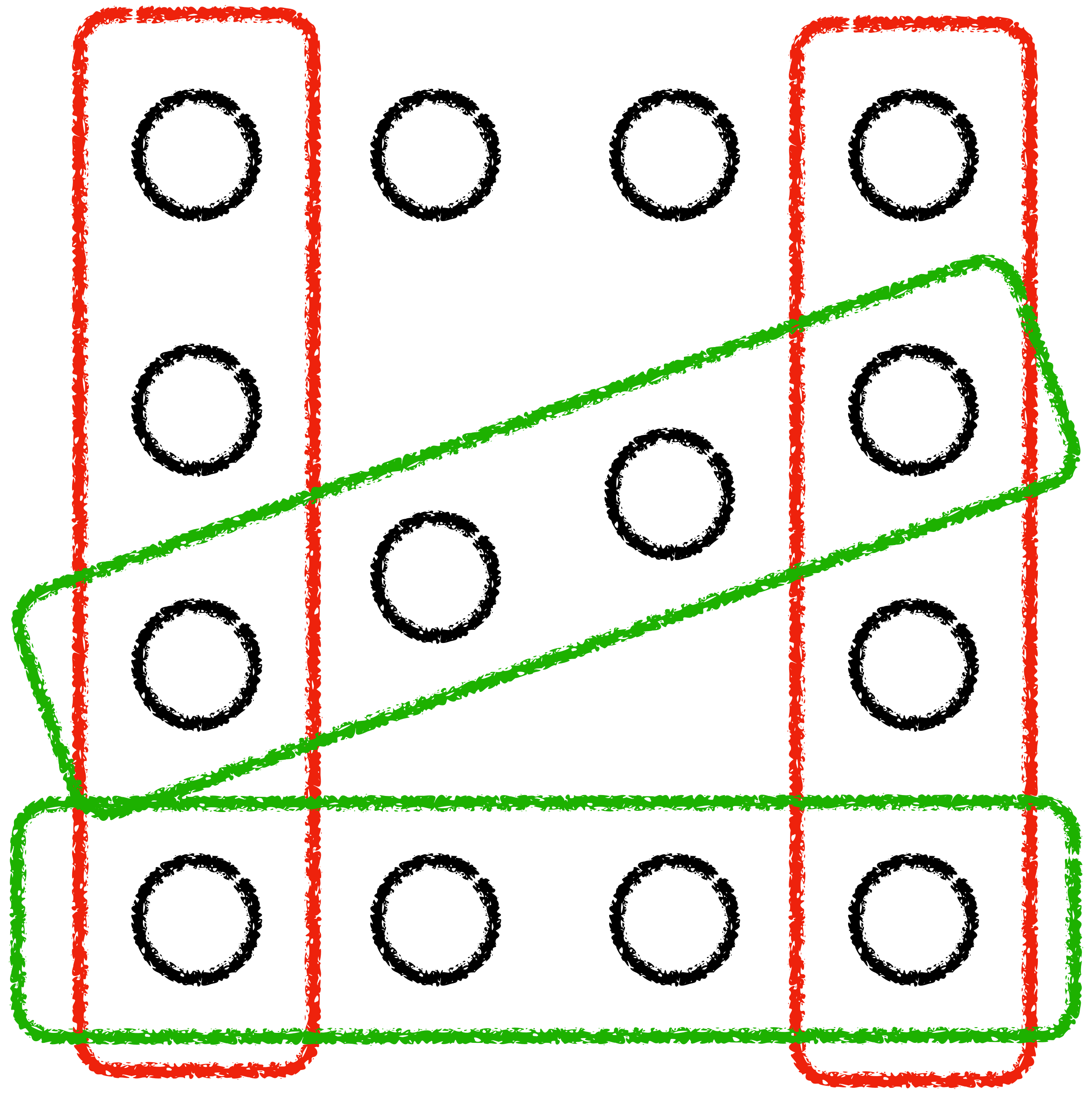
$(V, \mathcal{Q}, \mathscr{C}\backslash\{c_0\})$



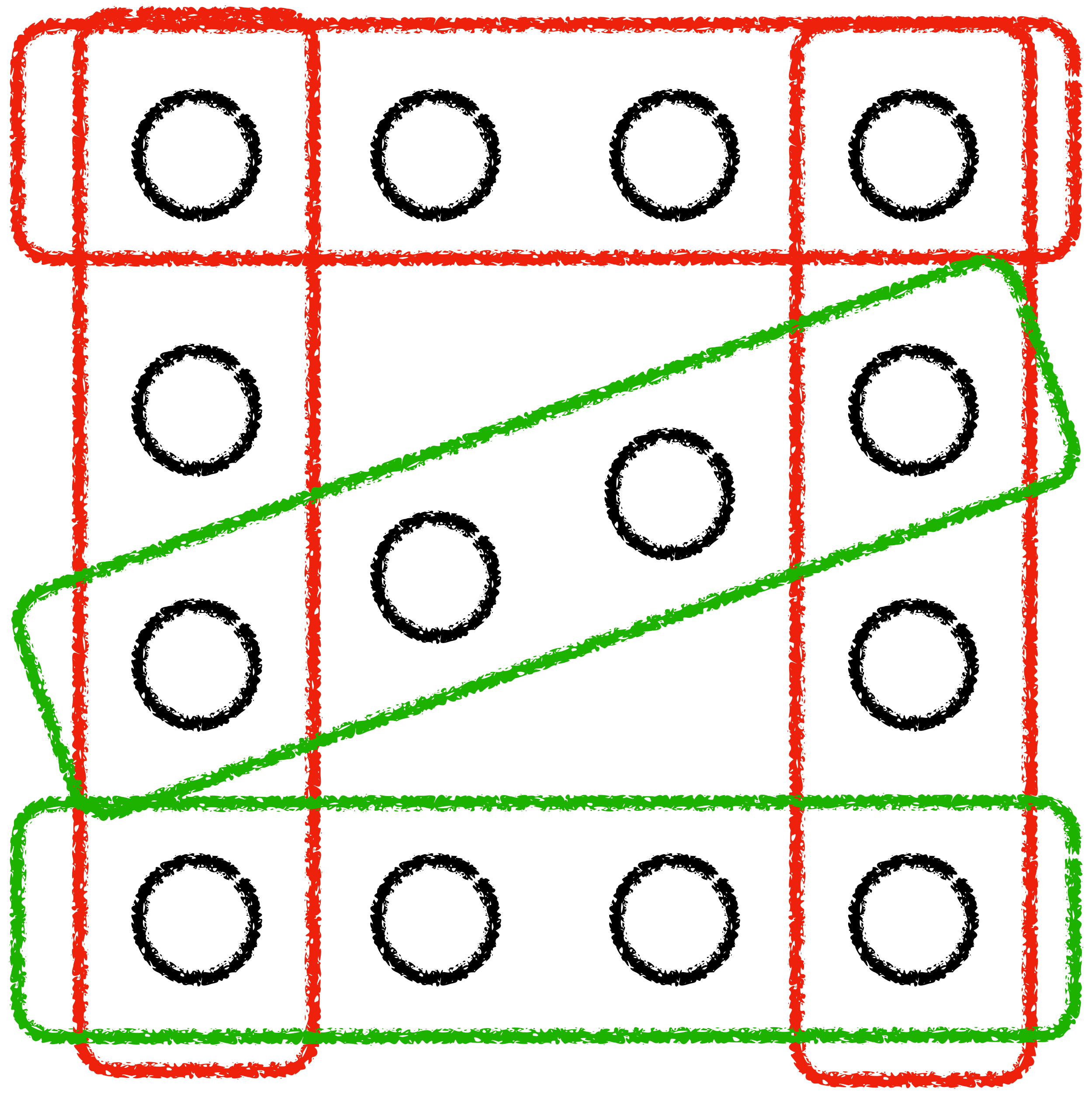
$(V, \mathcal{Q}, \mathscr{C})$

We want to couple $\mu_{\mathscr{C}\backslash\{c_0\}}$ with $\mu_{\mathscr{C}}$.

A constraint-wise coupling



$(V, \mathcal{Q}, \mathscr{C} \setminus \{c_0\})$



$(V, \mathcal{Q}, \mathscr{C})$

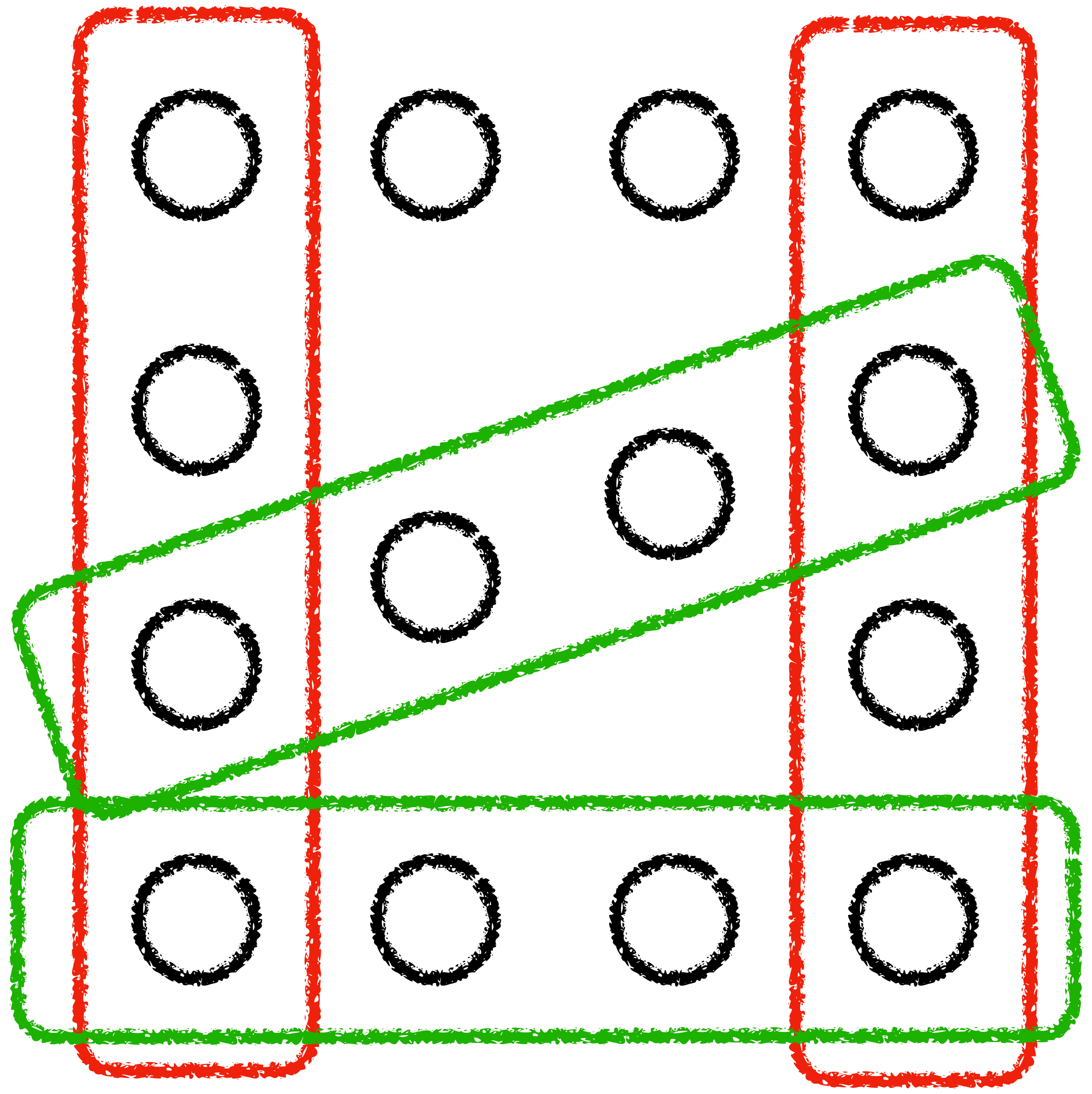
$$\mu_{\mathscr{C}\setminus\{c_0\}} = \mu_{\mathscr{C}\setminus\{c_0\}}(c_0) \cdot \mu_{\mathscr{C}} + \mu_{\mathscr{C}\setminus\{c_0\}}(\neg c_0) \cdot \mu_{\mathscr{C}\setminus\{c_0\}}(\cdot \mid \neg c_0)$$

A constraint-wise coupling

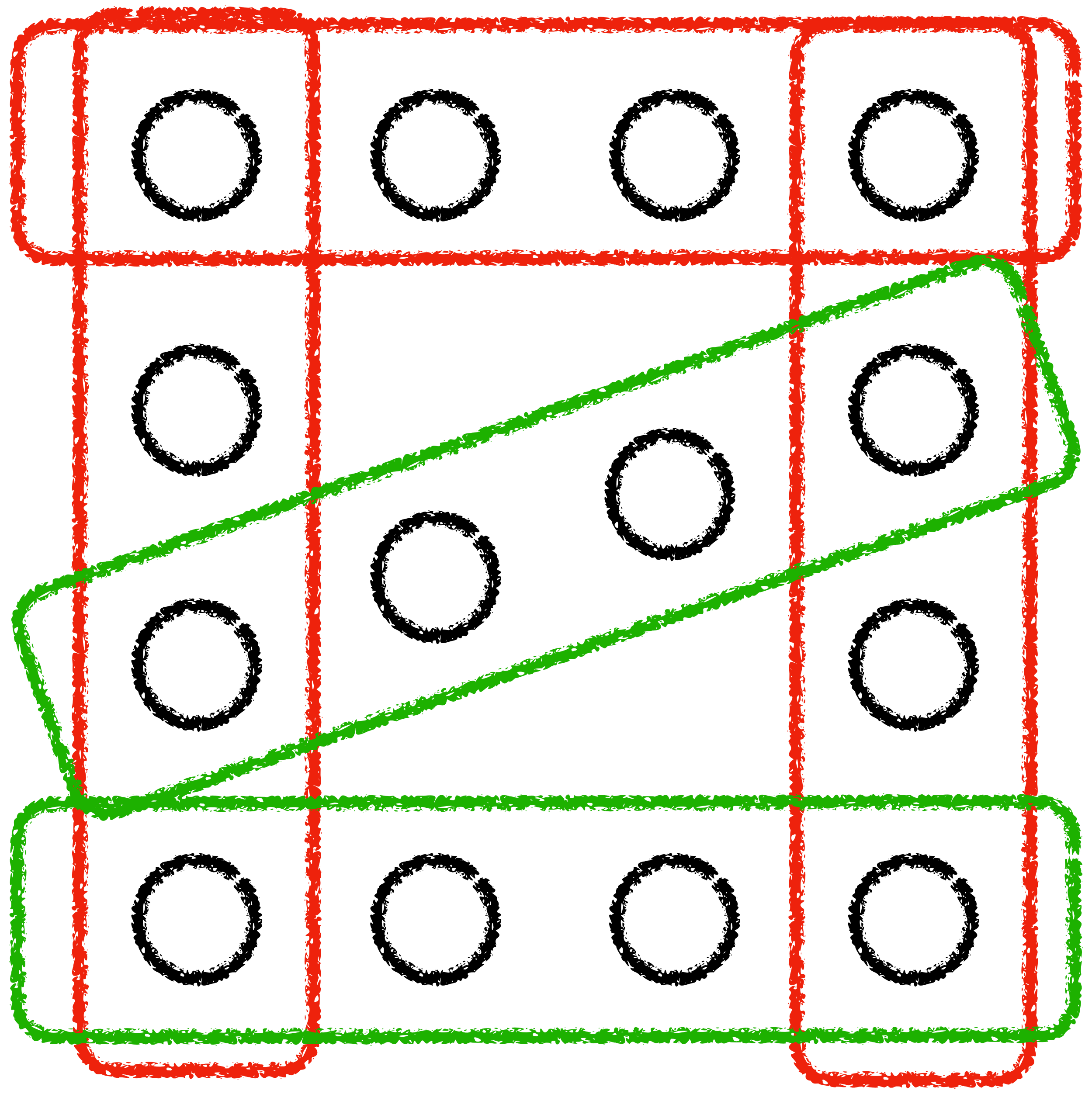
$(V, \mathcal{Q}, \mathcal{C} \backslash \{c_0\})$ $(V, \mathcal{Q}, \mathcal{C})$

with prob. $\mu_{\mathcal{C}\backslash\{c_0\}}(c_0)$, couple $\mu_{\mathcal{C}}$ with $\mu_{\mathcal{C}}$;
with prob. $\mu_{\mathcal{C}\backslash\{c_0\}}(\neg c_0)$, couple $\mu_{\mathcal{C}\backslash\{c_0\}}(\cdot \mid \neg c_0)$ with $\mu_{\mathcal{C}}$.

A constraint-wise coupling



$(V, \mathcal{Q}, \mathcal{C} \setminus \{c_0\})$

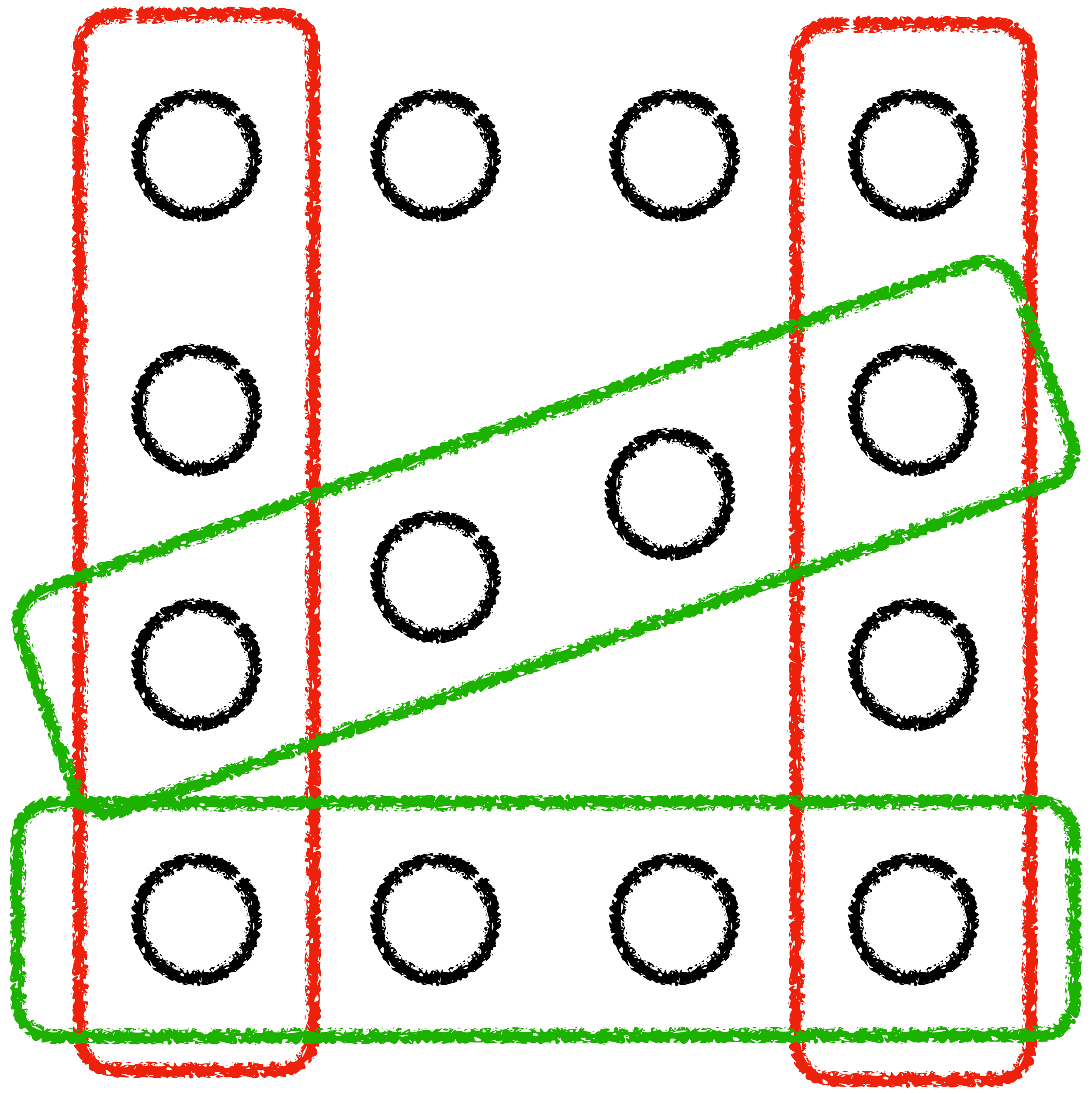


$(V, \mathcal{Q}, \mathcal{C})$

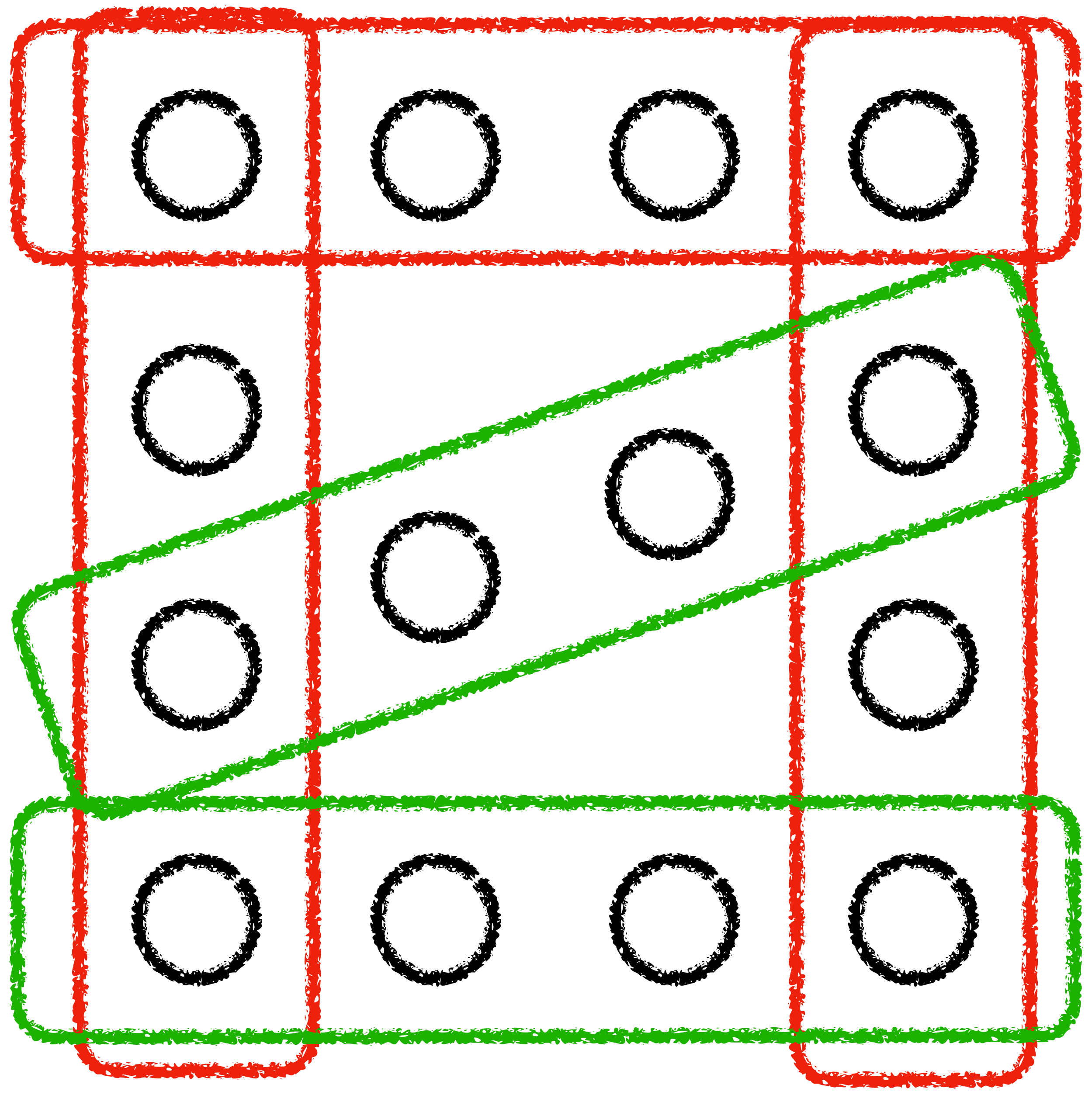
with prob. $\mu_{\mathcal{C}\setminus\{c_0\}}(c_0)$, couple $\mu_{\mathcal{C}}$ with $\mu_{\mathcal{C}}$; **can be perfectly coupled!**

with prob. $\mu_{\mathcal{C}\setminus\{c_0\}}(\neg c_0)$, couple $\mu_{\mathcal{C}\setminus\{c_0\}}(\cdot \mid \neg c_0)$ with $\mu_{\mathcal{C}}$.

A constraint-wise coupling



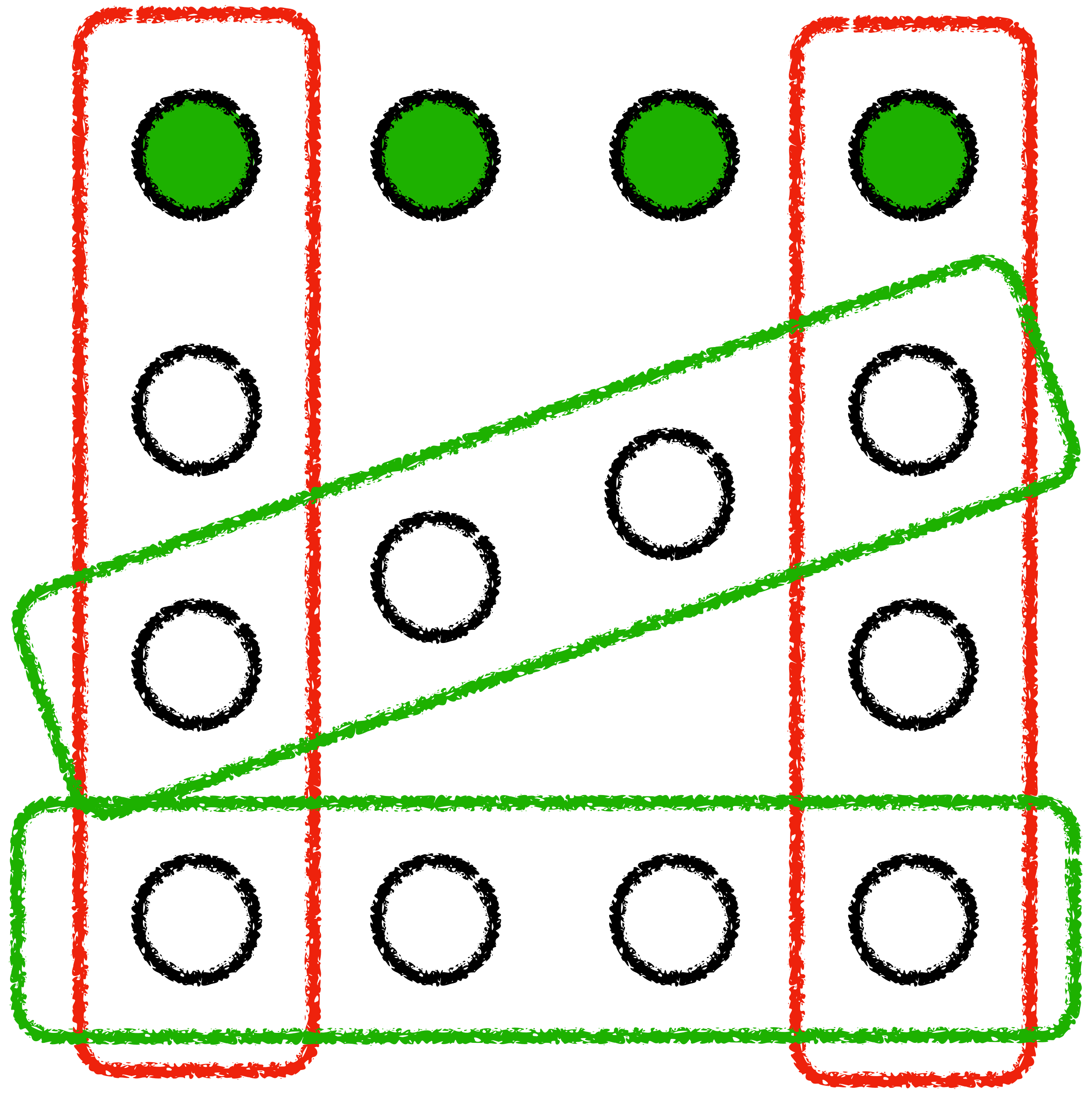
$(V, \mathcal{Q}, \mathcal{C} \setminus \{c_0\})$



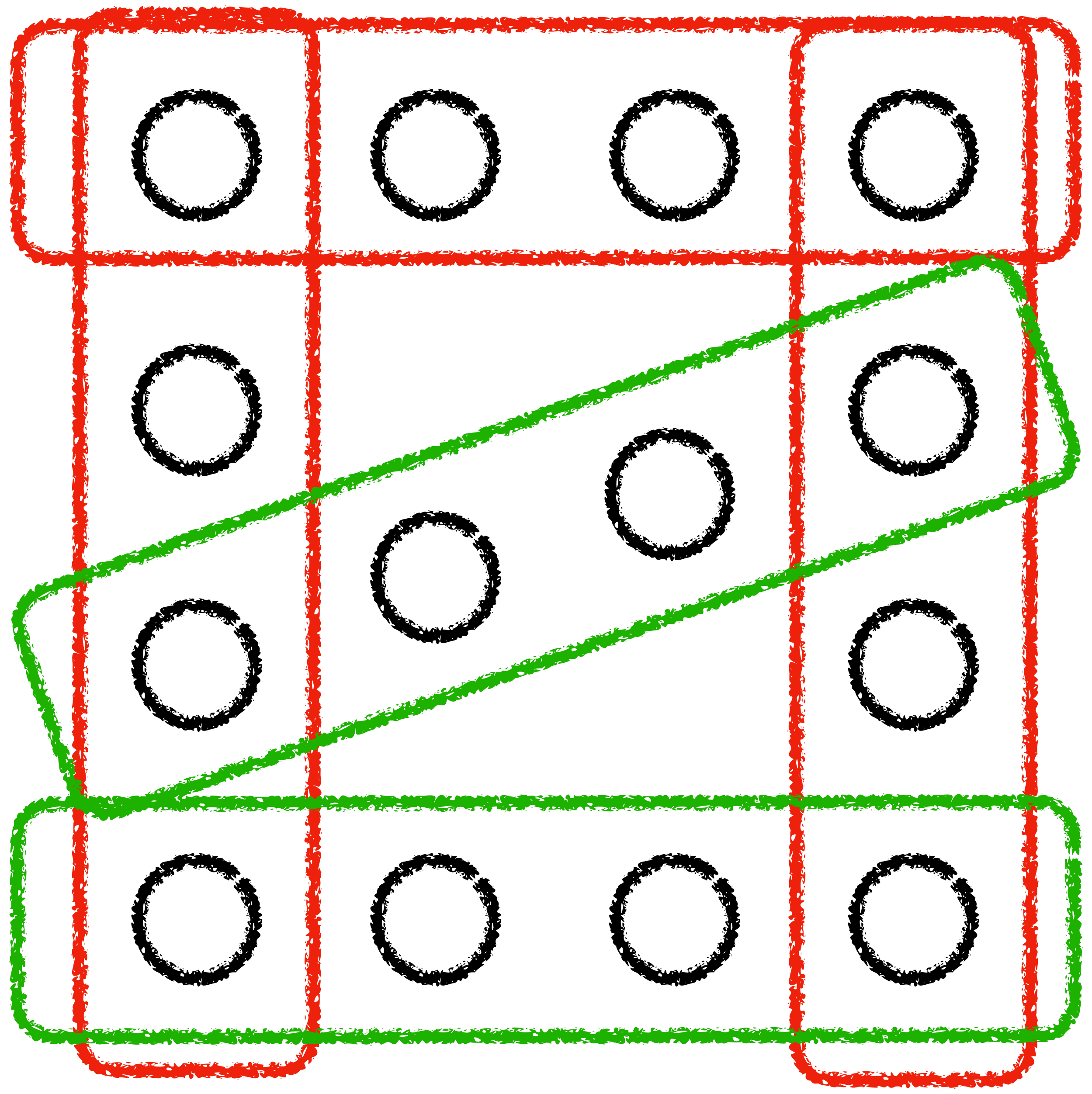
$(V, \mathcal{Q}, \mathcal{C})$

We now couple $\mu_{\mathcal{C} \setminus \{c_0\}}(\,\cdot \mid \neg c_0)$ with $\mu_{\mathcal{C}}$.

A constraint-wise coupling



$(V, \mathcal{Q}, \mathscr{C} \backslash \{c_0\})$

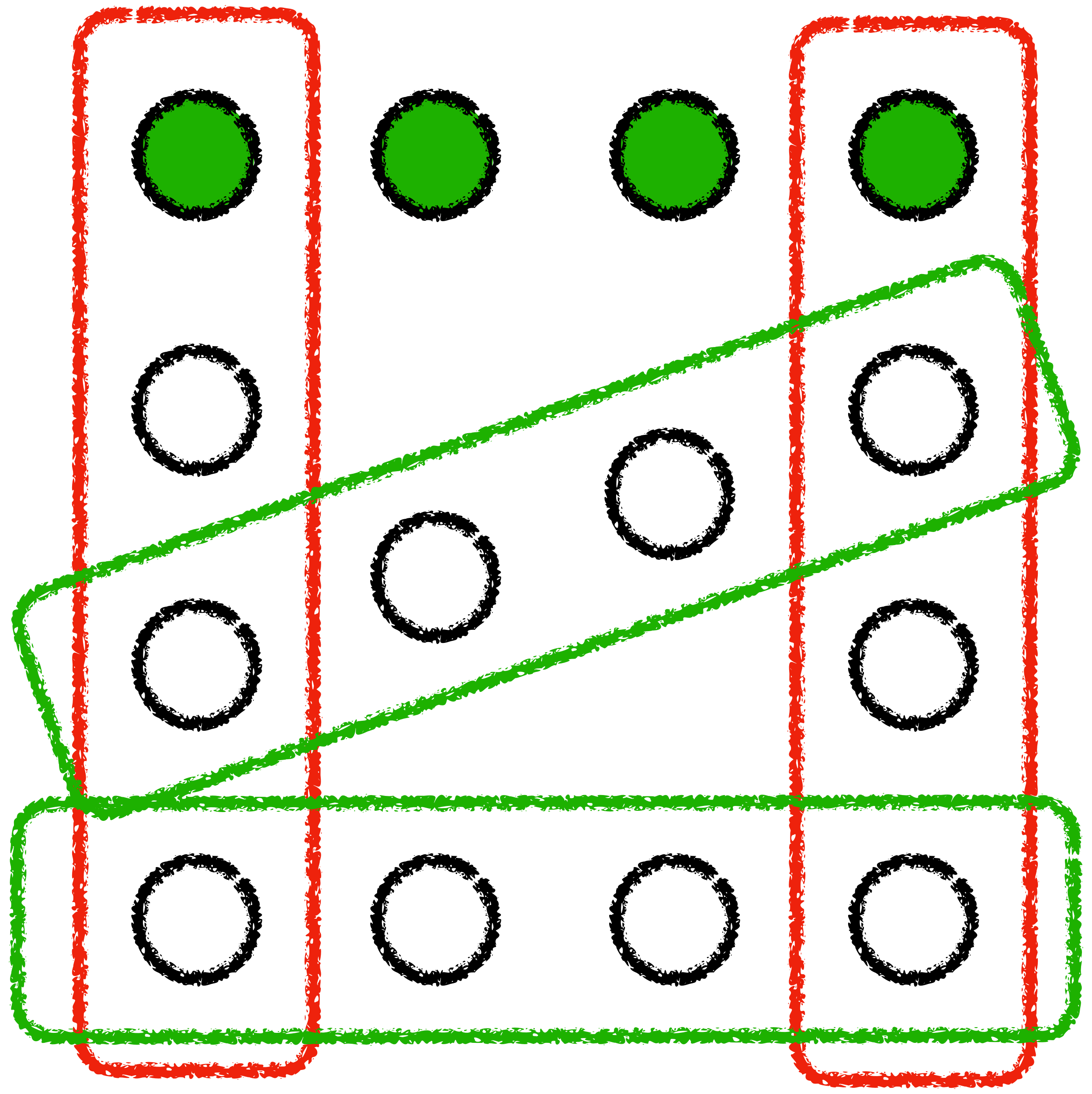


$(V, \mathcal{Q}, \mathscr{C})$

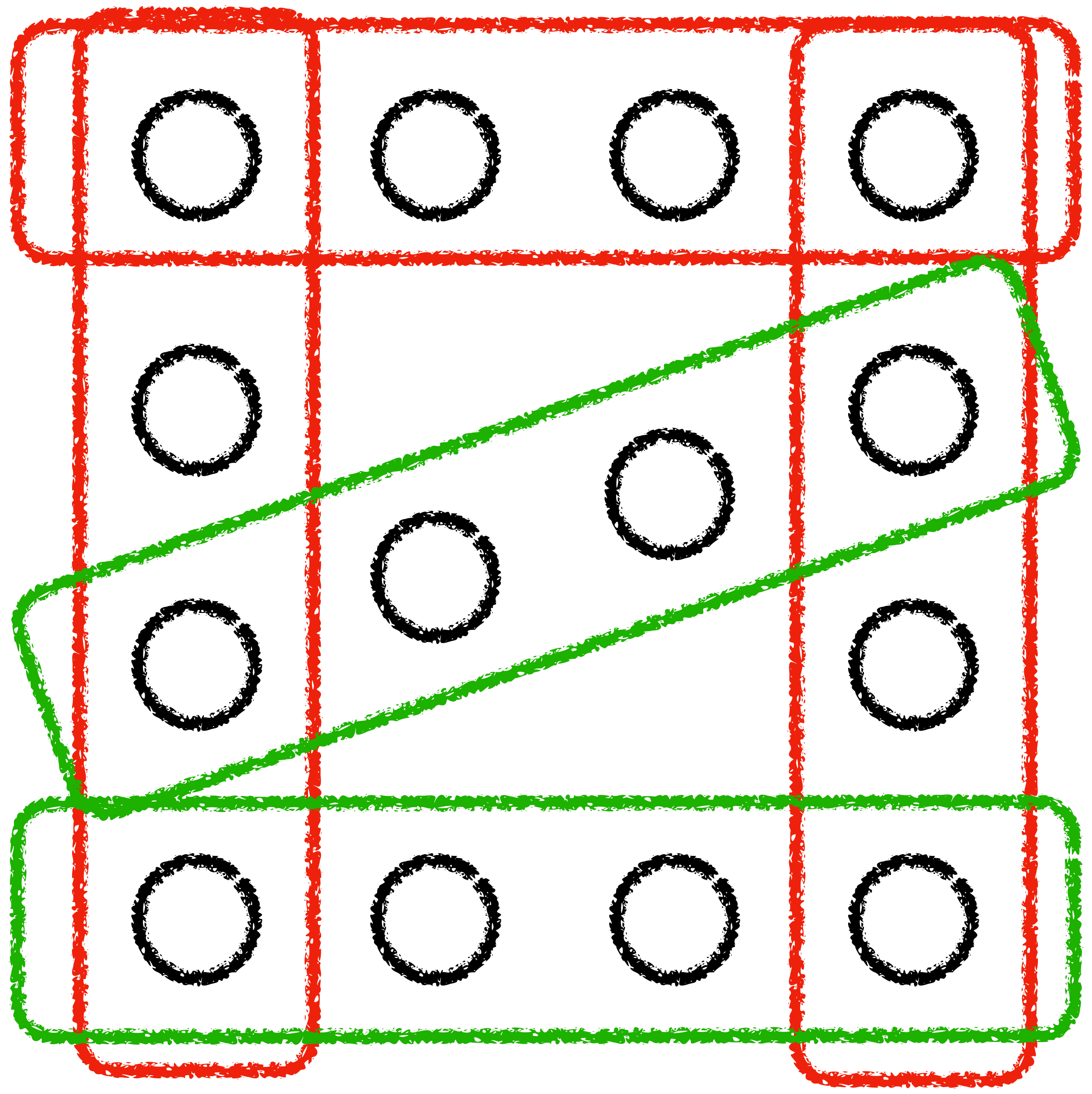
forced assignment !

We now couple $\mu_{\mathscr{C} \backslash \{c_0\}}(\,\cdot \mid \neg c_0)$ with $\mu_{\mathscr{C}}$.

A constraint-wise coupling



$(V, \mathcal{Q}, \mathscr{C} \setminus \{c_0\})$



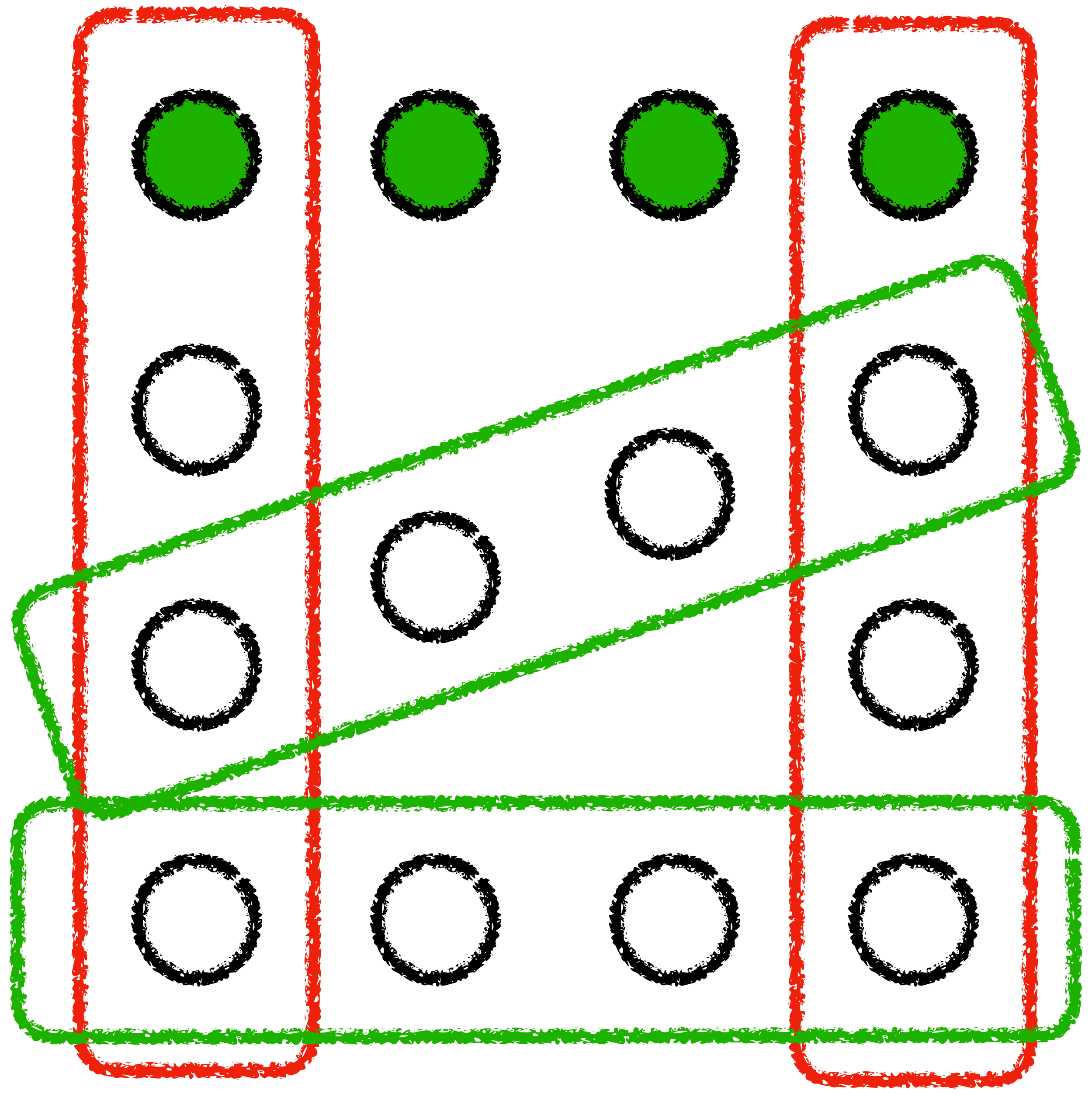
$(V, \mathcal{Q}, \mathscr{C})$

forced assignment !

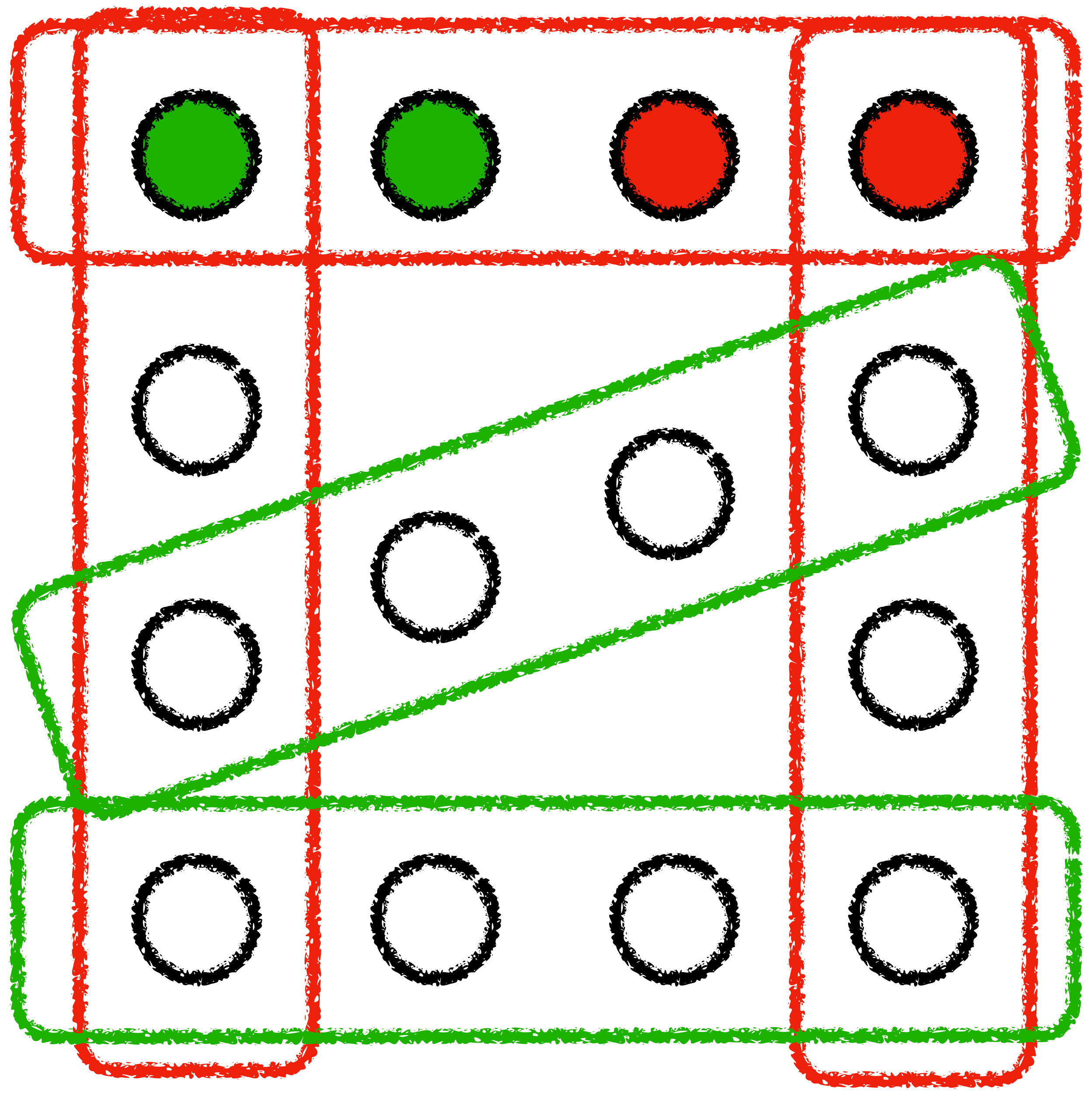
We now couple $\mu_{\mathscr{C}\setminus\{c_0\}}(\cdot \mid \neg c_0)$ with $\mu_{\mathscr{C}}$.

We further decompose $\mu_{\mathscr{C}} = \sum_{\rho \in \mathcal{Q}_{\mathrm{vbl}(c_0)}} \mu_{\mathscr{C}}(\rho) \cdot \mu_{\mathscr{C}}(\cdot \mid \rho)$.

A constraint-wise coupling



$(V, \mathcal{Q}, \mathcal{C} \backslash \{c_0\})$

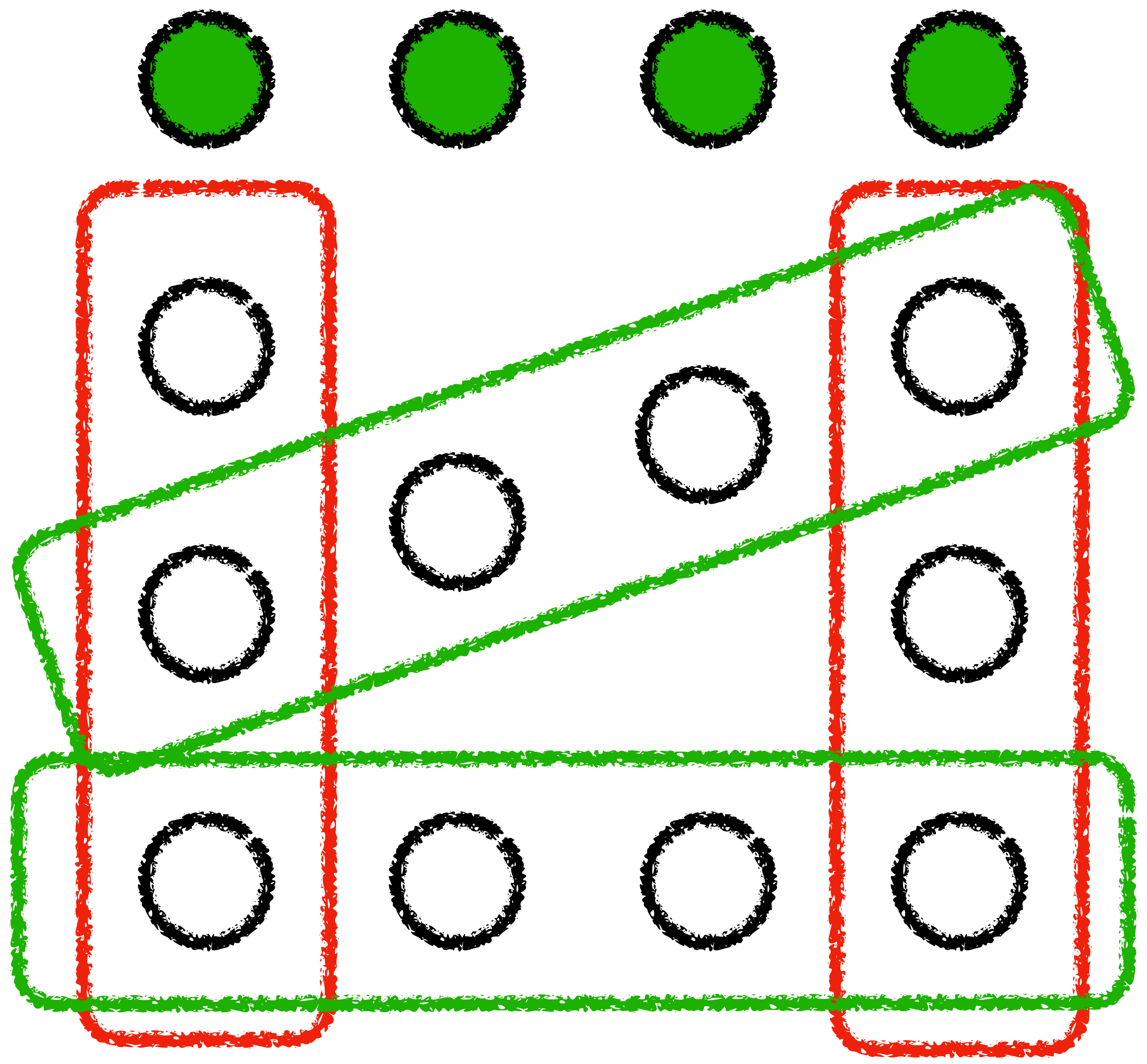


$(V, \mathcal{Q}, \mathcal{C})$

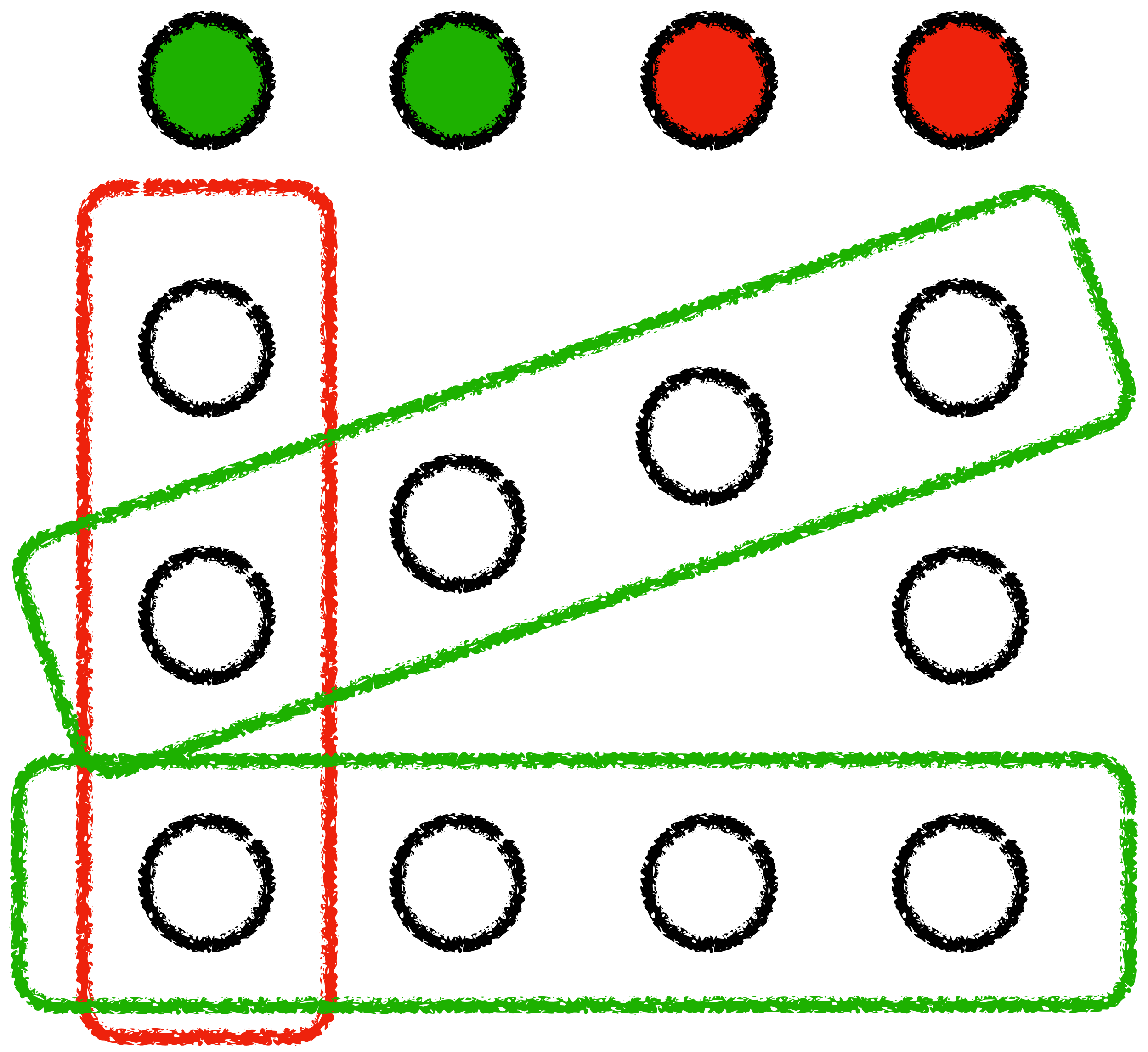
We now couple $\mu_{\mathcal{C}\backslash\{c_0\}}(\,\cdot \mid \neg c_0)$ with $\mu_{\mathcal{C}}$.

We further decompose $\mu_{\mathcal{C}} = \sum_{\rho \in \mathcal{Q}_{\text{vbl}(c_0)}} \mu_{\mathcal{C}}(\rho) \cdot \mu_{\mathcal{C}}(\,\cdot \mid \rho)$.

A constraint-wise coupling

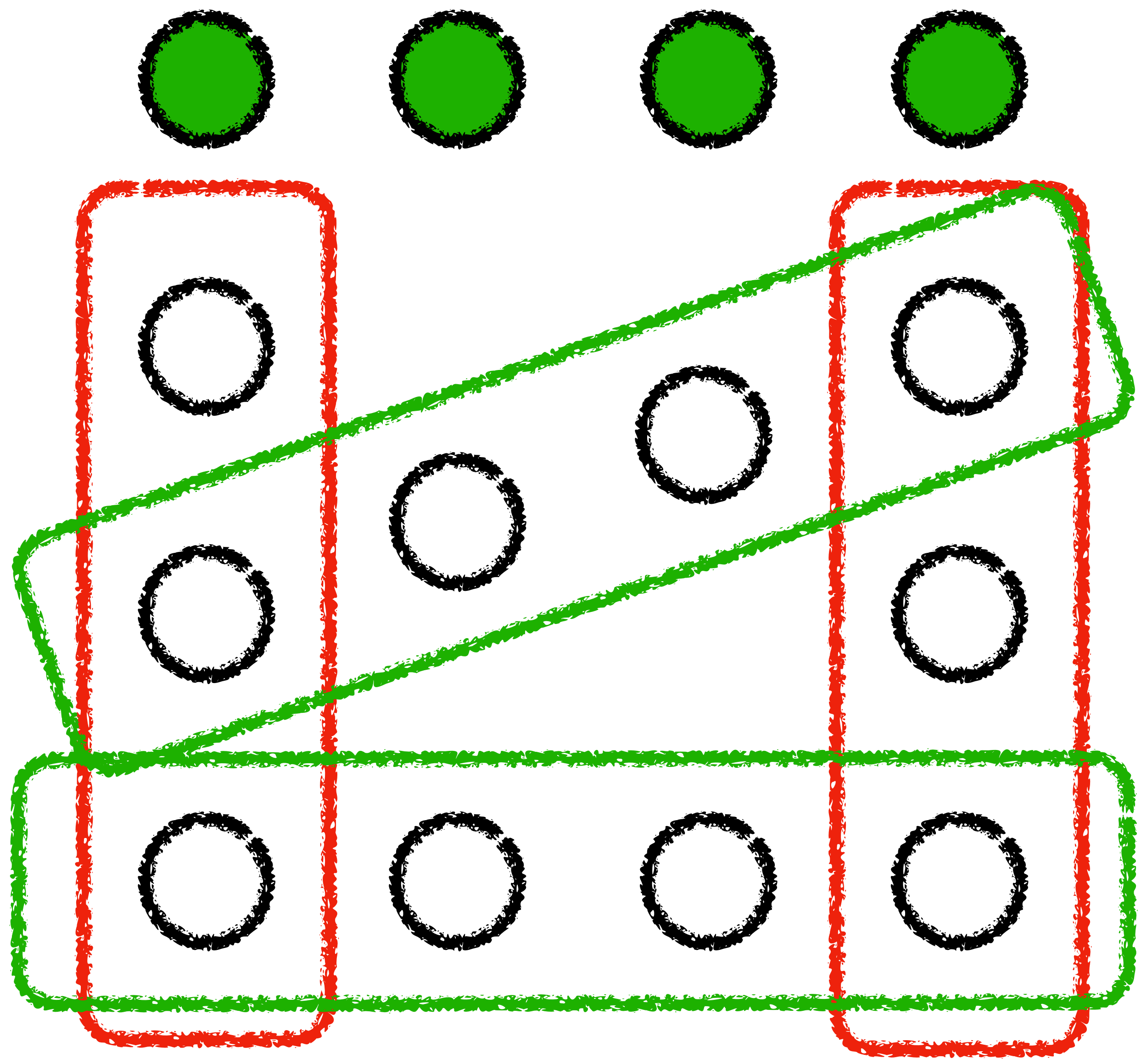


$(V, \mathscr{Q}, \mathscr{C} \backslash \{c_0\})$

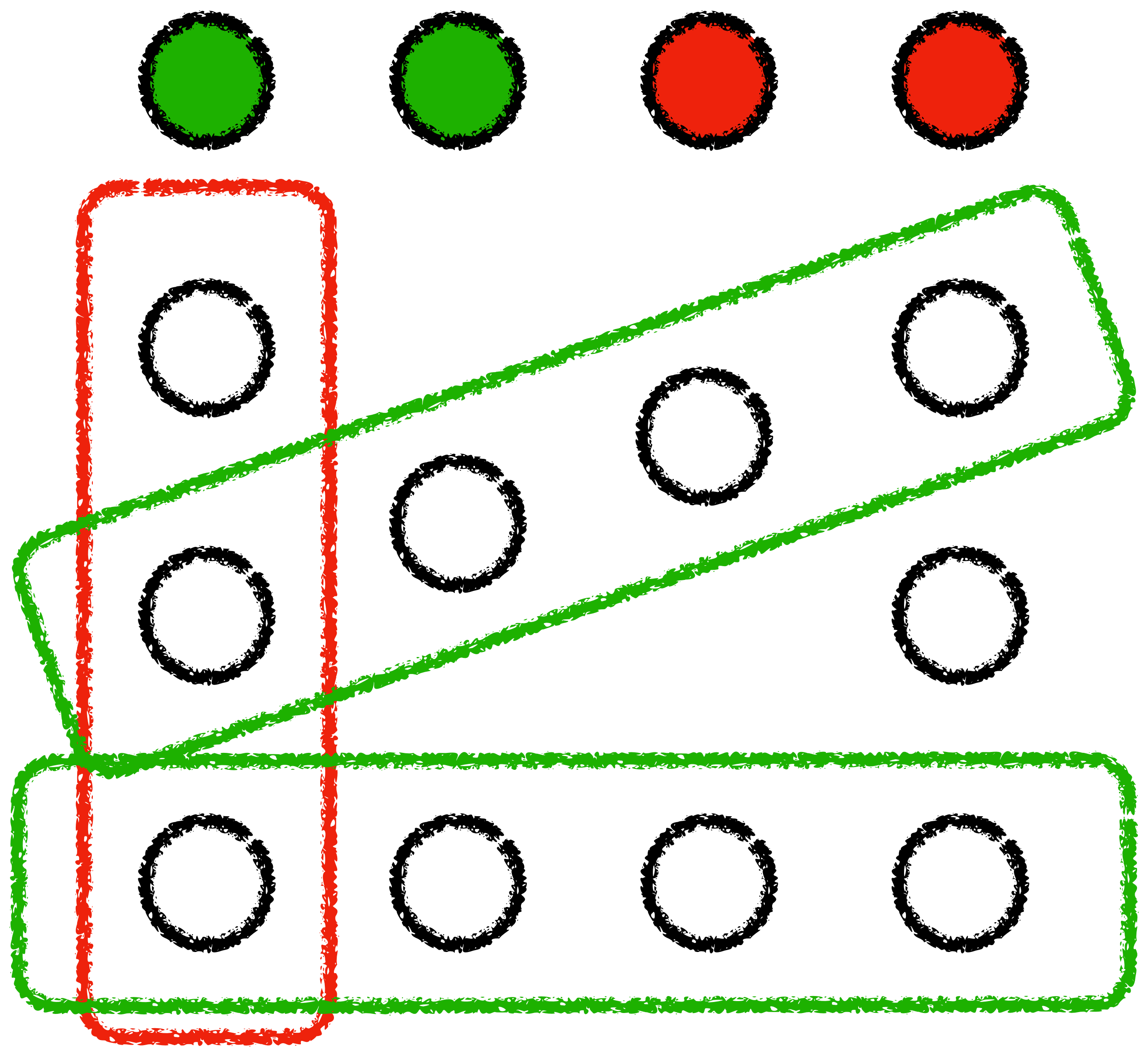


$(V, \mathscr{Q}, \mathscr{C})$

A constraint-wise coupling



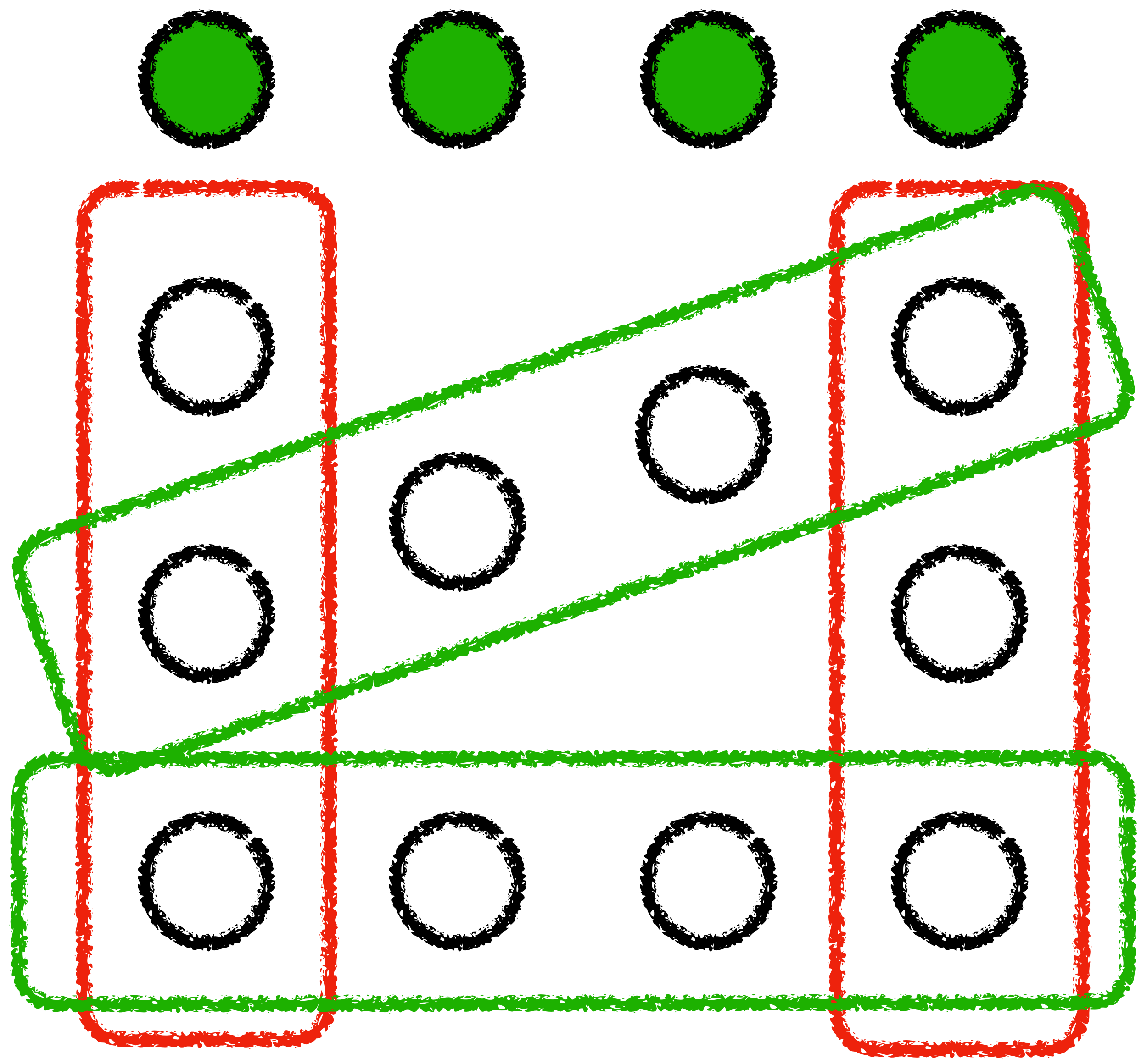
$(V, \mathcal{Q}, \mathcal{C} \backslash \{c_0\})$



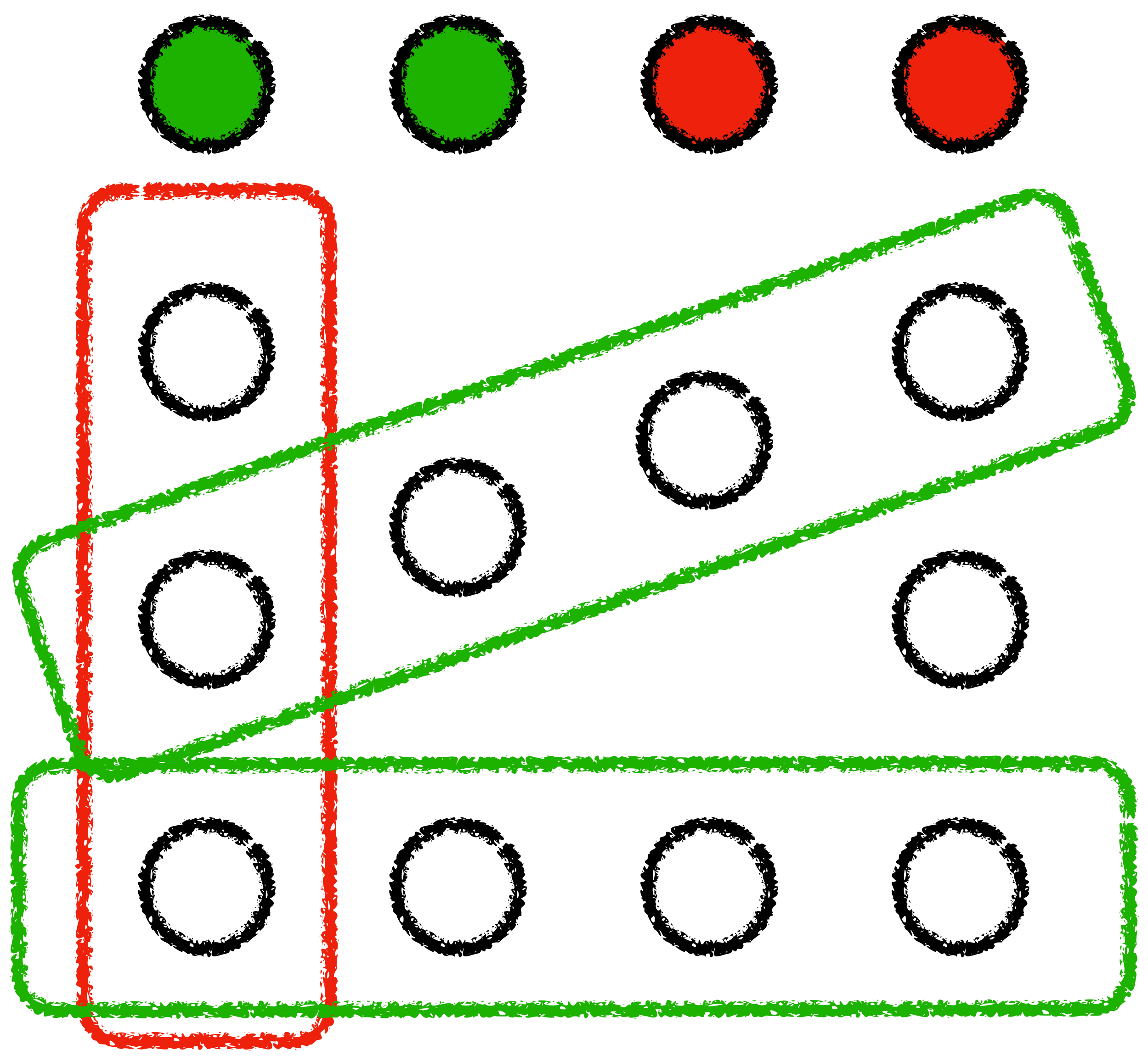
$(V, \mathcal{Q}, \mathcal{C})$

Simplify the formula, we are done if the set of constraints are the same.

A constraint-wise coupling



$(V, \mathcal{Q}, \mathscr{C} \backslash \{c_0\})$

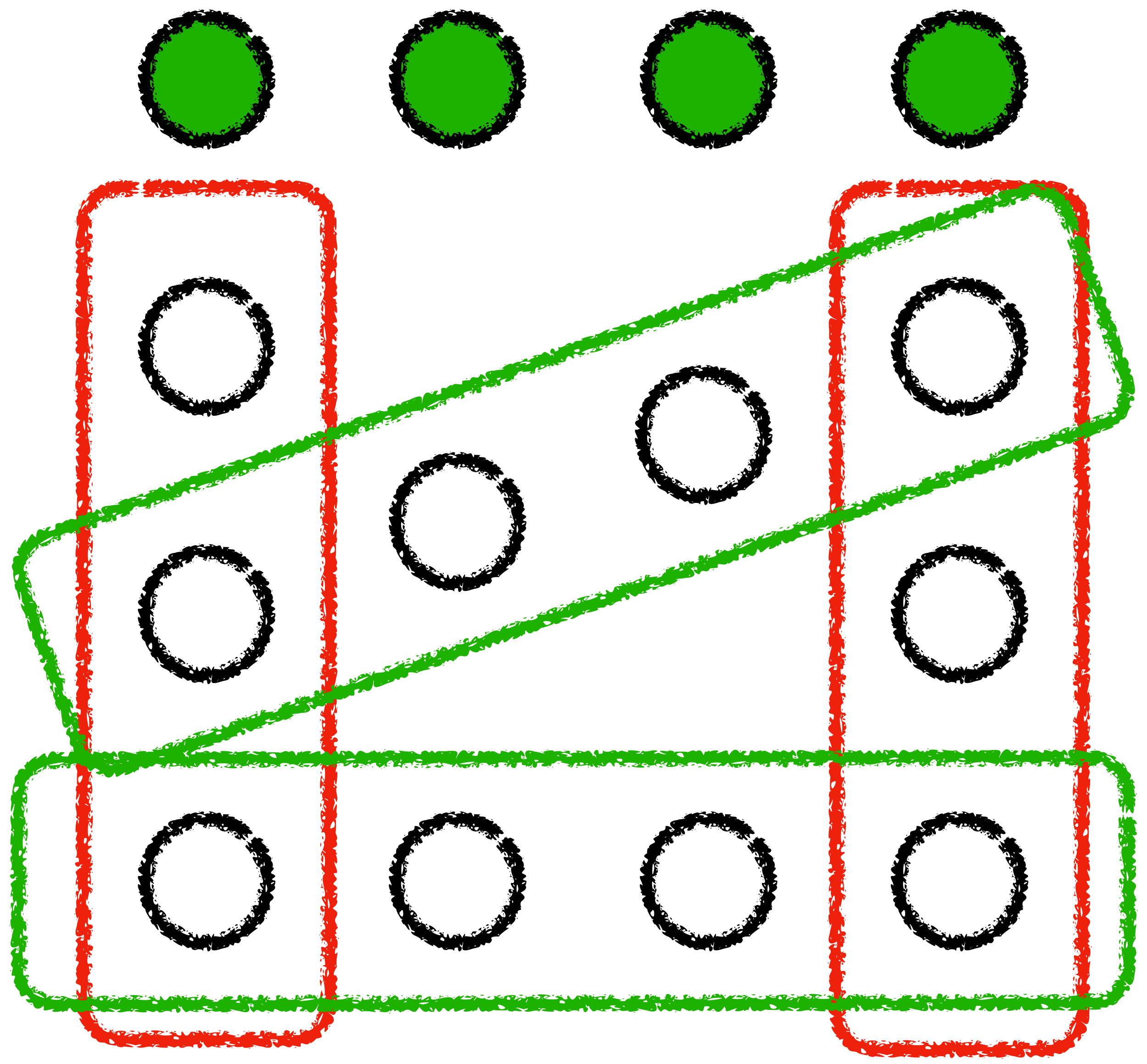


$(V, \mathcal{Q}, \mathscr{C})$

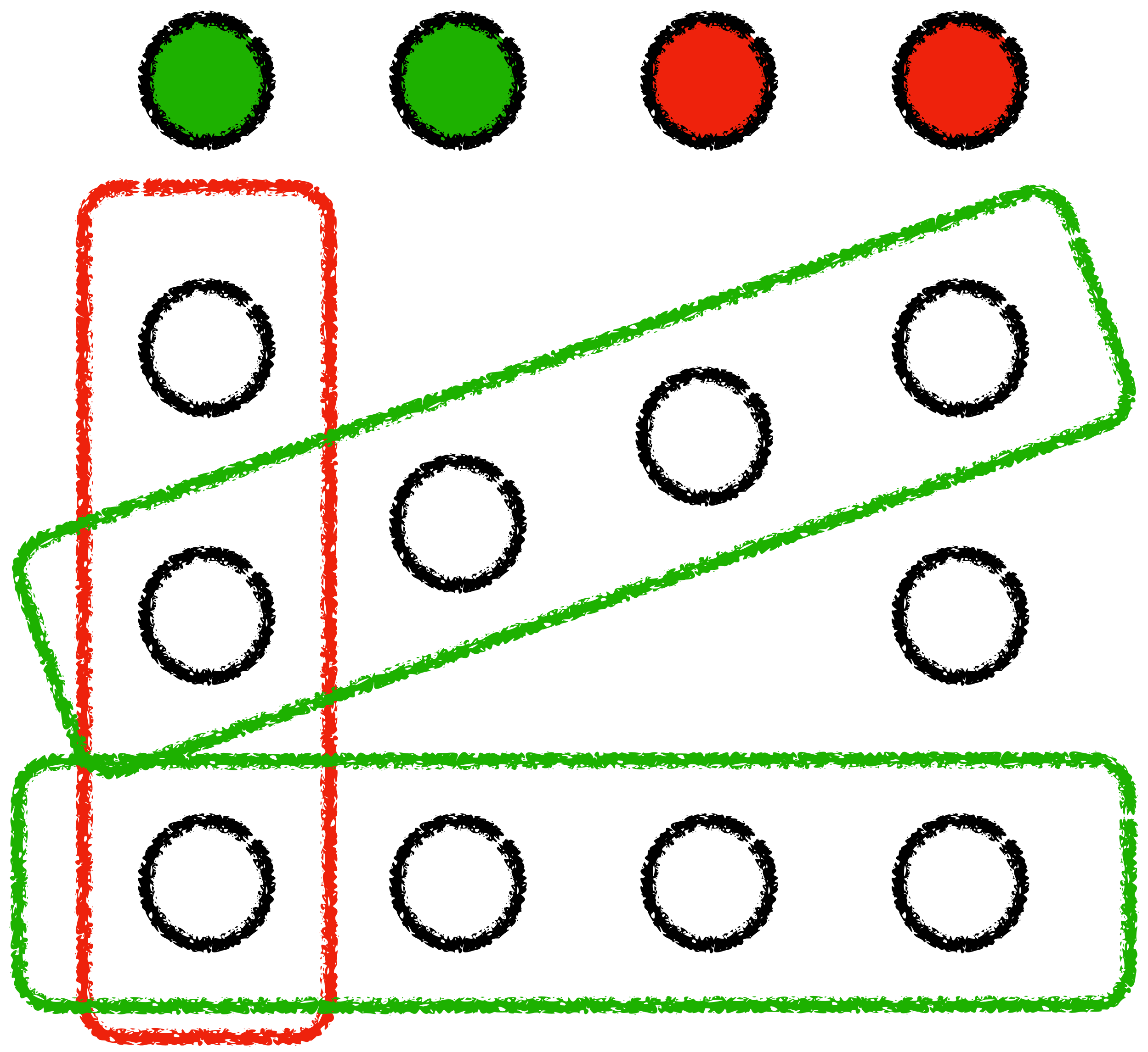
Simplify the formula, we are done if the set of constraints are the same.

Otherwise, we pick any constraint in the discrepancy set and recurse!

Analysis of the coupling



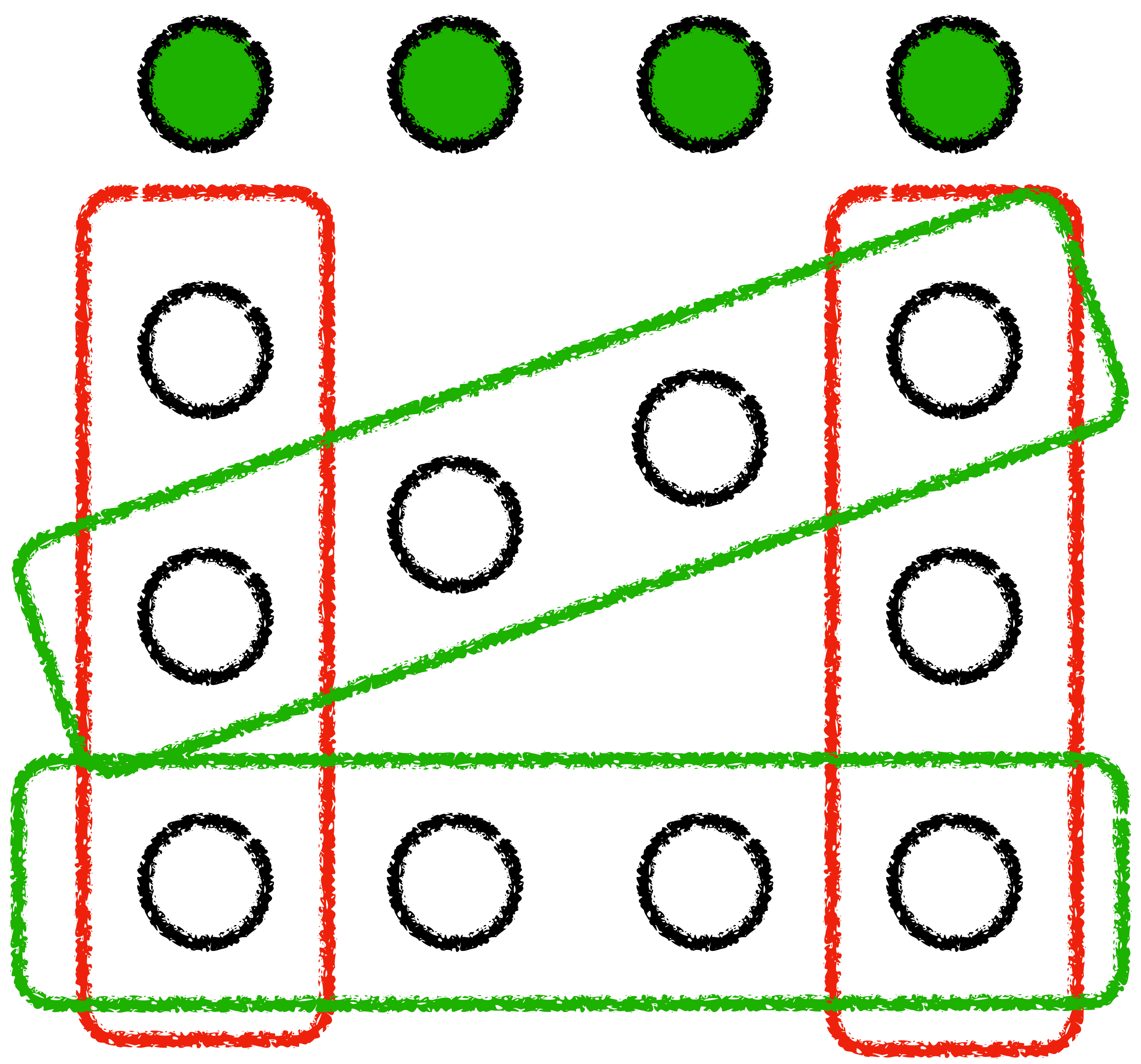
$(V, \mathcal{Q}, \mathscr{C} \backslash \{c_0\})$



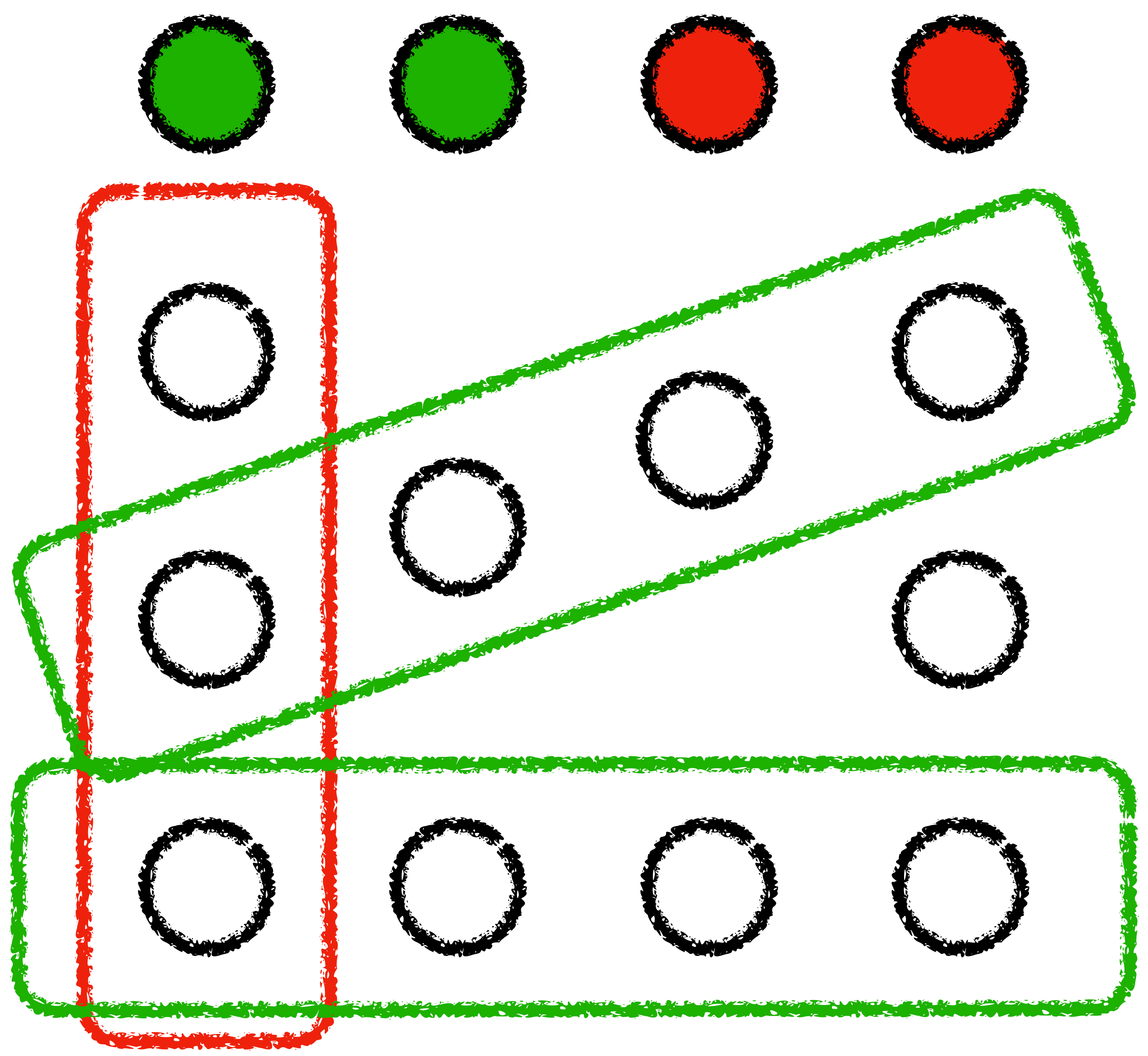
$(V, \mathcal{Q}, \mathscr{C})$

Challenge for the analysis: LLL condition still may degrade after each step

Analysis of the coupling



$(V, \mathcal{Q}, \mathscr{C} \backslash \{c_0\})$

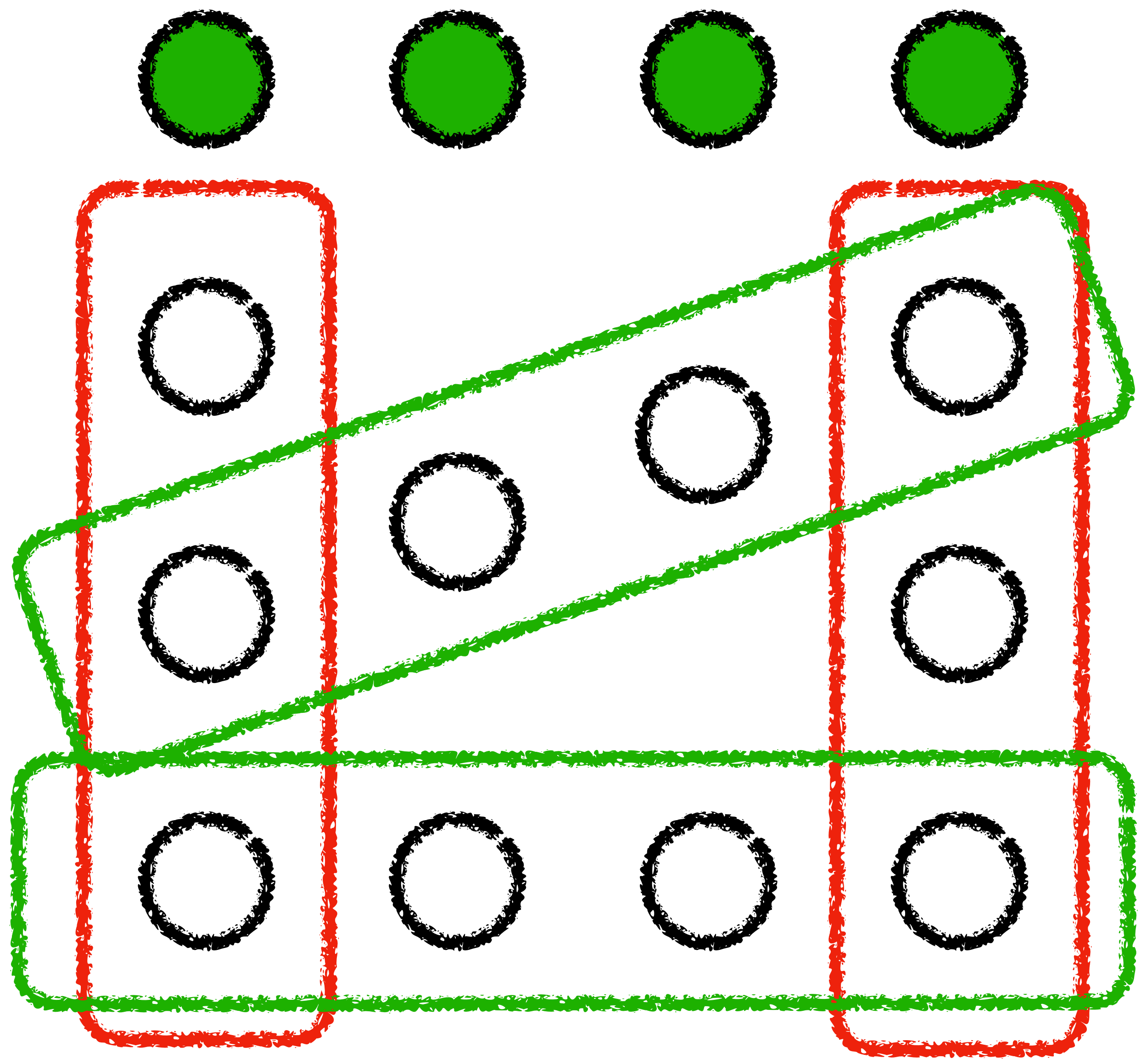


$(V, \mathcal{Q}, \mathscr{C})$

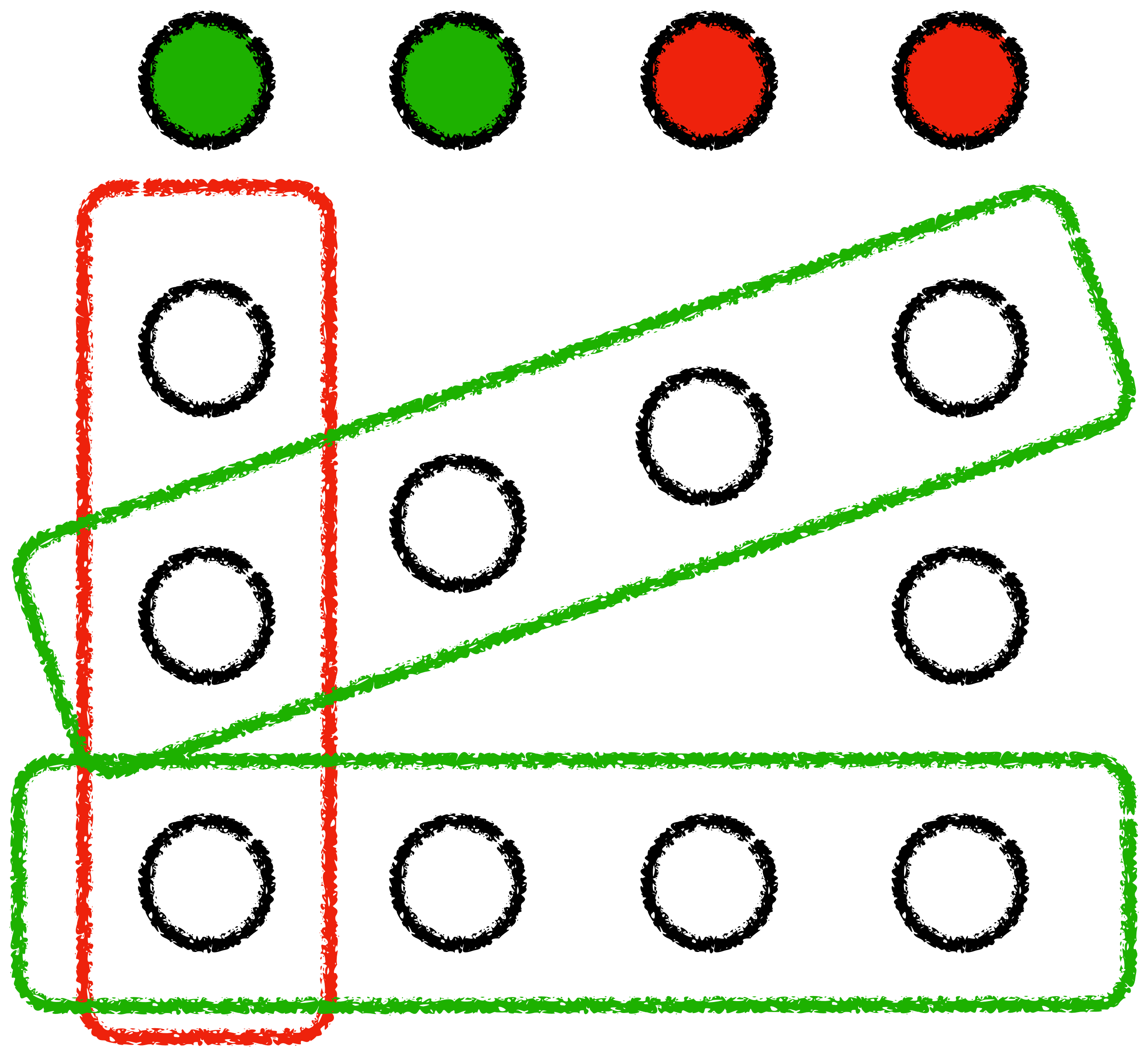
All randomness by the procedure can be identified by two independent samples:

$$\mathfrak{X} \sim \mu_{\mathscr{C}\backslash\{c_0\}}, \quad \mathfrak{Y} \sim \mu_{\mathscr{C}}.$$

Analysis of the coupling



$(V, \mathcal{Q}, \mathscr{C} \setminus \{c_0\})$



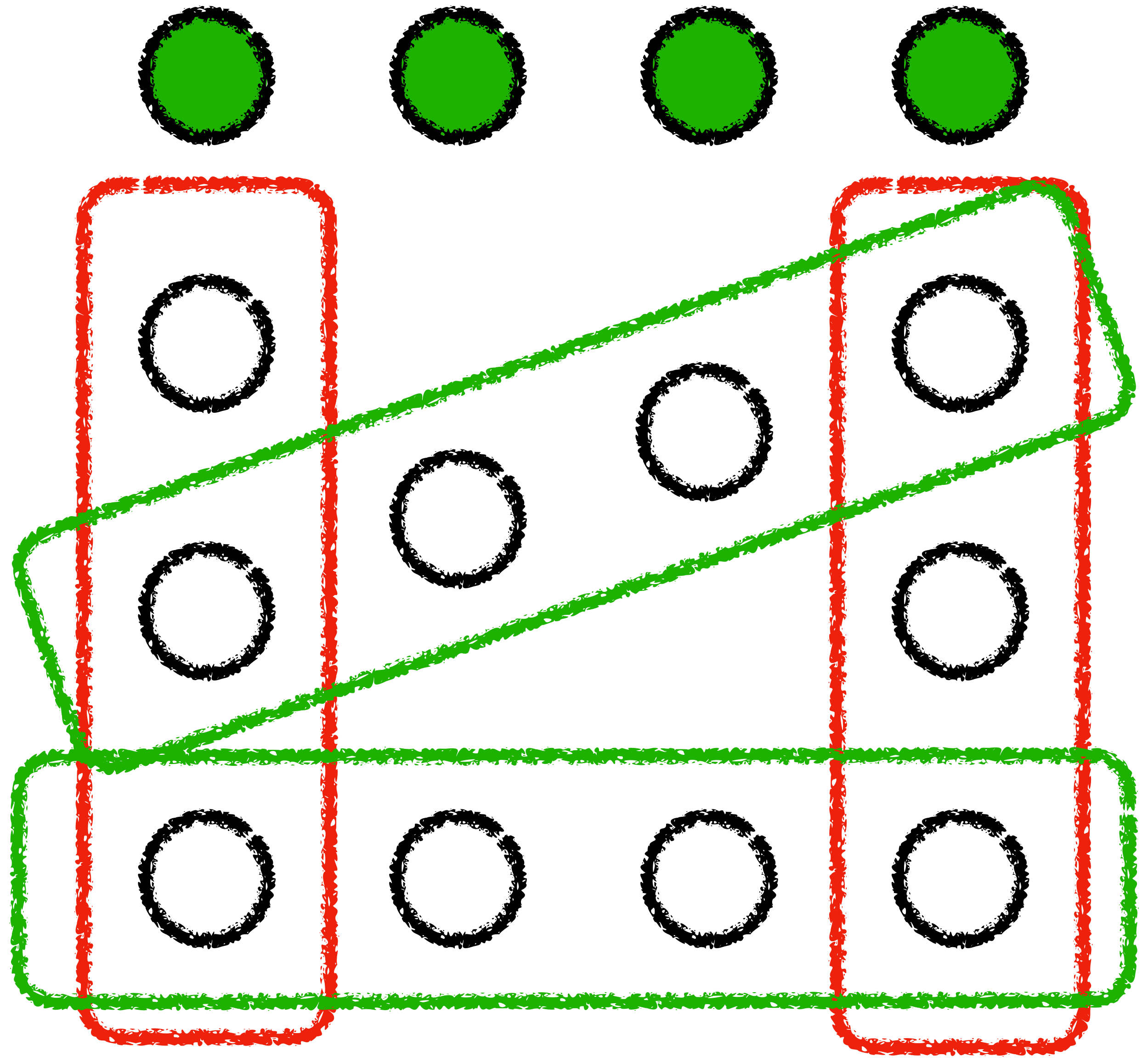
$(V, \mathcal{Q}, \mathscr{C})$

All randomness by the procedure can be identified by two independent samples:

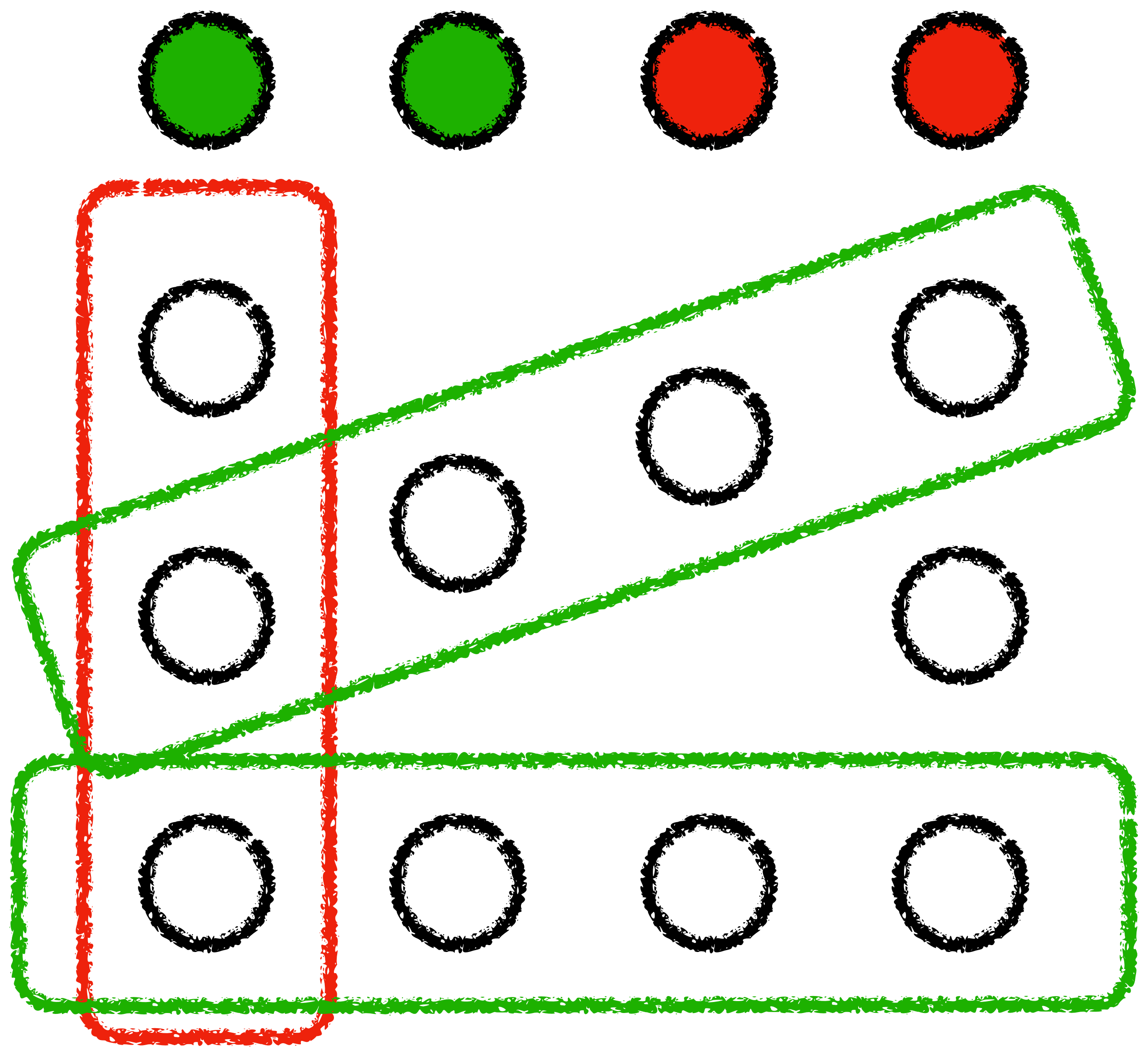
$$\mathfrak{X} \sim \mu_{\mathscr{C} \setminus \{c_0\}}, \quad \mathfrak{Y} \sim \mu_{\mathscr{C}} .$$

Sampling by marginal distribution = Revealing local information of $\mathfrak{X}$ and $\mathfrak{Y}$

Analysis of the coupling



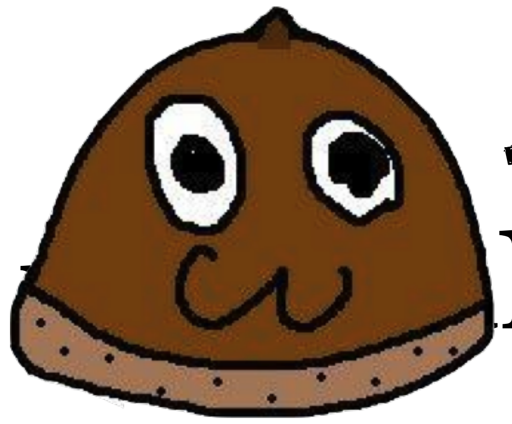
$(V, \mathcal{Q}, \mathscr{C}\setminus\{c_0\})$



$(V, \mathcal{Q}, \mathscr{C})$

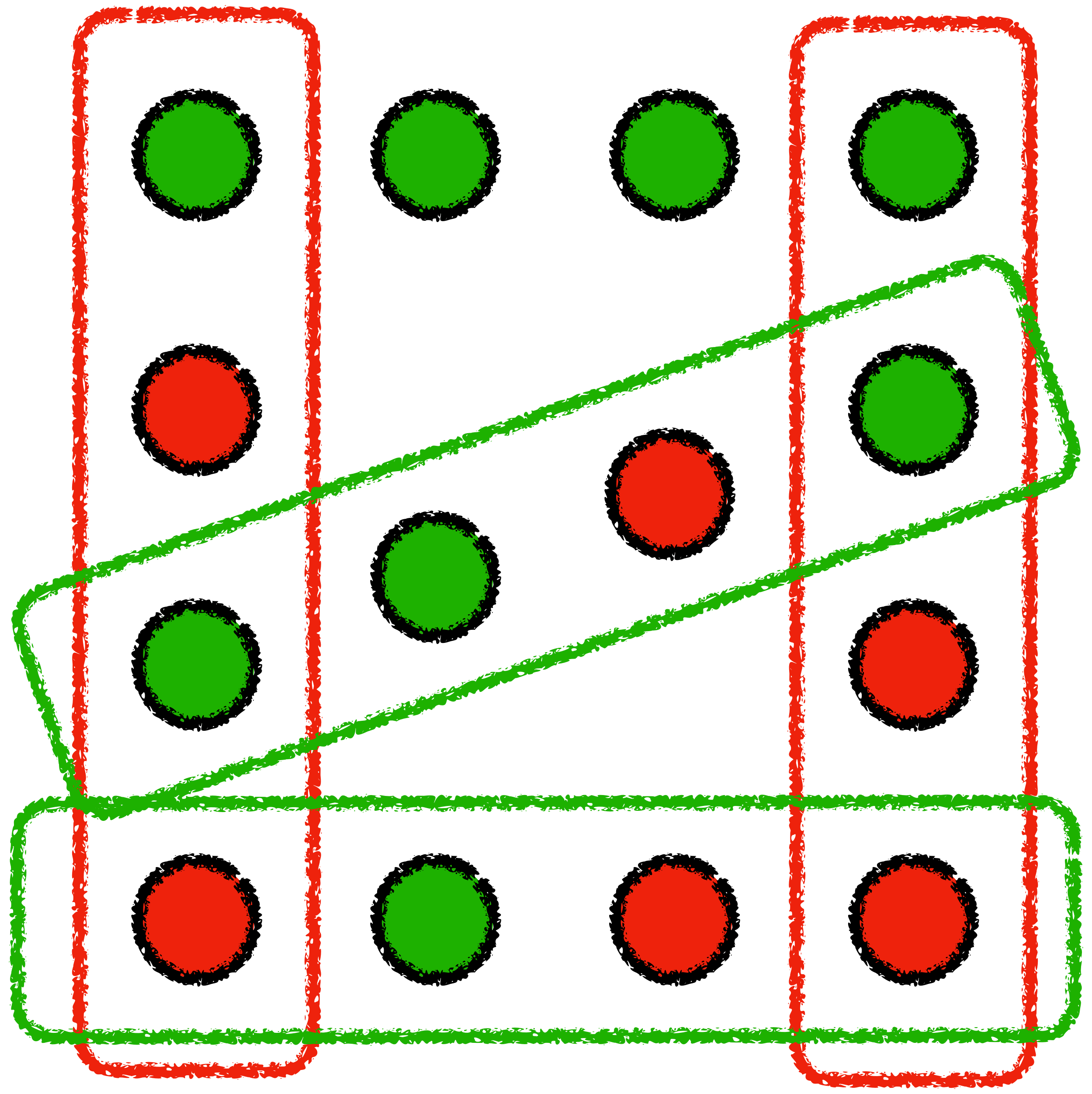
All randomness by the ... t samples:

$\mu_{\mathscr{C}\setminus\{c_0\}}, \quad \mathfrak{Y} \sim \mu_{\mathscr{C}}.$

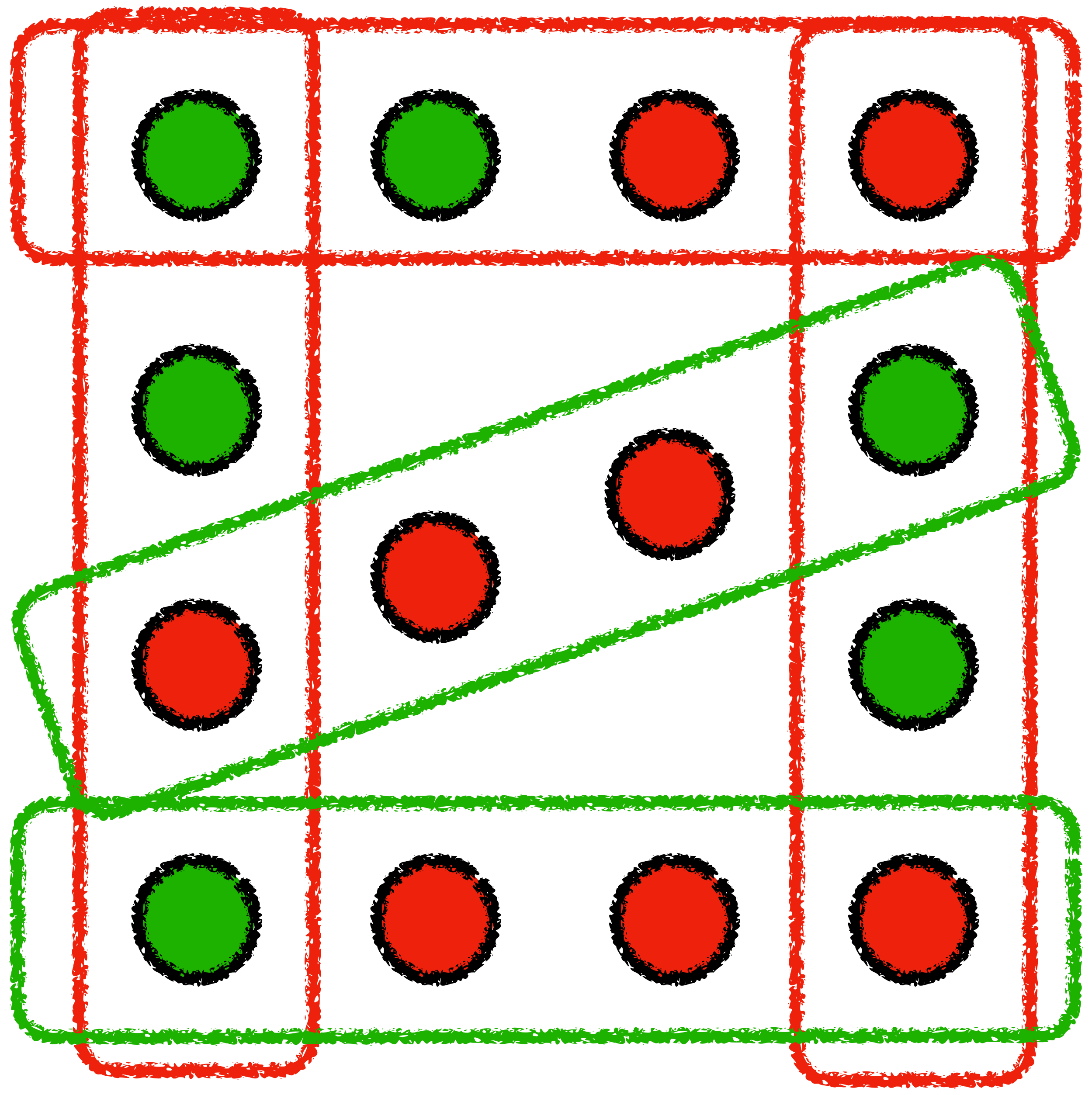


Sampling by ... l distribution = Revealing local information of $\mathfrak{X}$ and $\mathfrak{Y}$

Analysis of the coupling



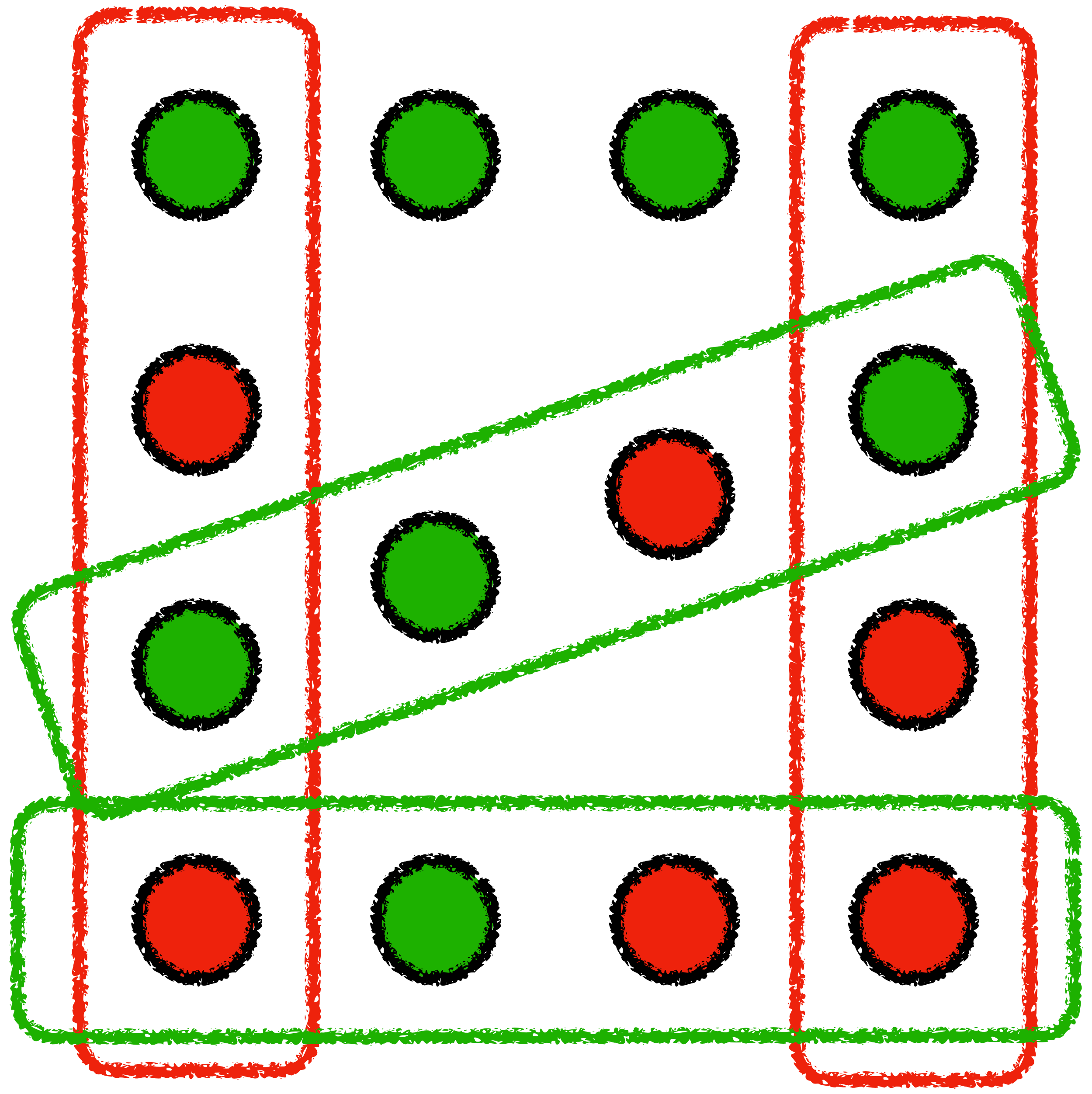
$\mathfrak{X} \sim \mu_{\mathscr{C}\setminus\{c_0\}}$



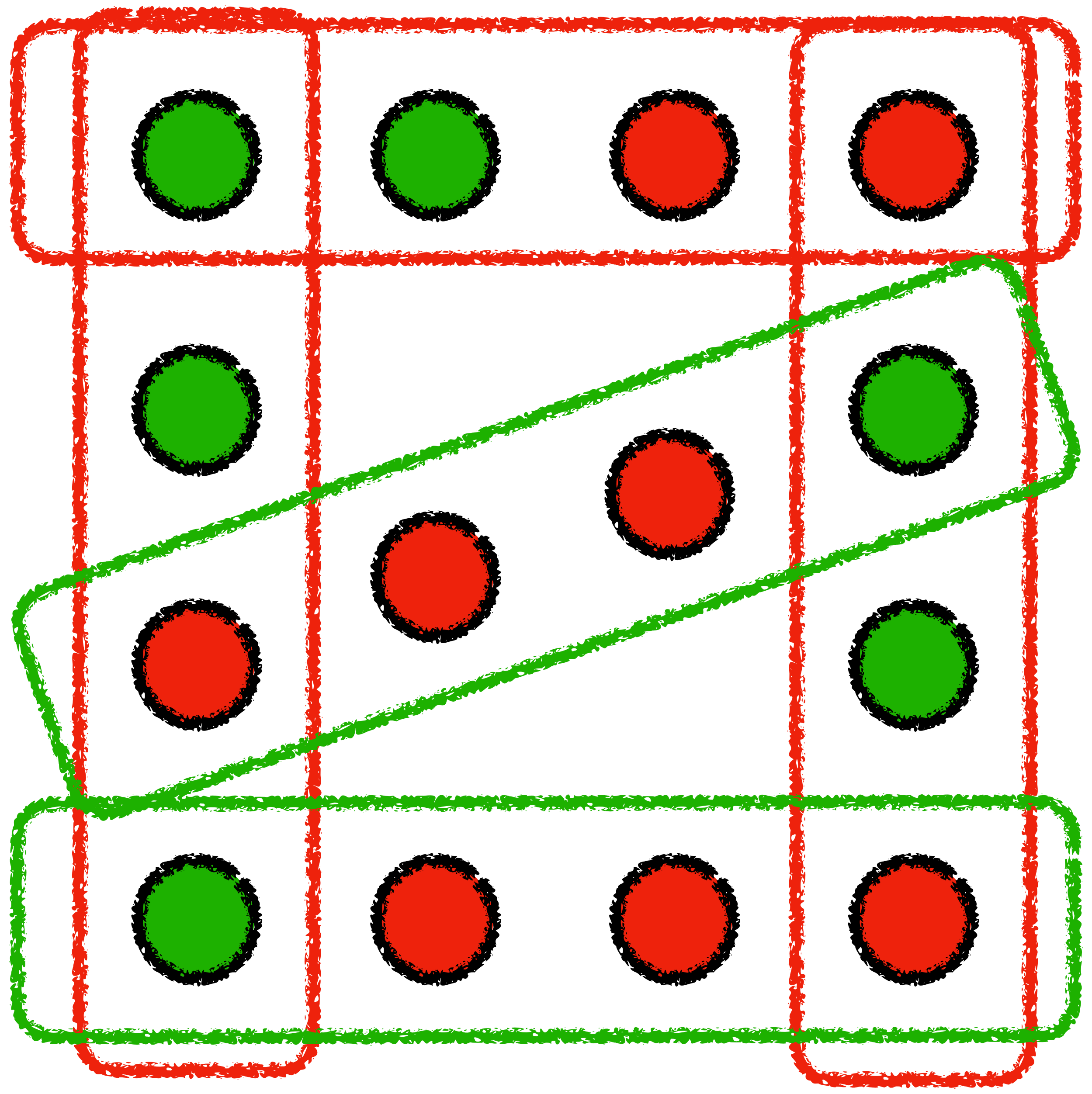
$\mathfrak{Y} \sim \mu_{\mathscr{C}}$

[HSS '14]: under local lemma regimes, $\mathfrak{X}$ and $\mathfrak{Y}$ behave close to uniform

Analysis of the coupling



$\mathfrak{X} \sim \mu_{\mathscr{C} \setminus \{c_0\}}$



$\mathfrak{Y} \sim \mu_{\mathscr{C}}$

[HSS '14]: under local lemma regimes, $\mathfrak{X}$ and $\mathfrak{Y}$ behave close to uniform

witness of large discrepancy + percolation-style analysis

Algorithmic implications

Algorithmic implications

We cannot really run the coupling, but we can write down linear programs that encode coupling errors to bootstrap the marginal probability.

Algorithmic implications

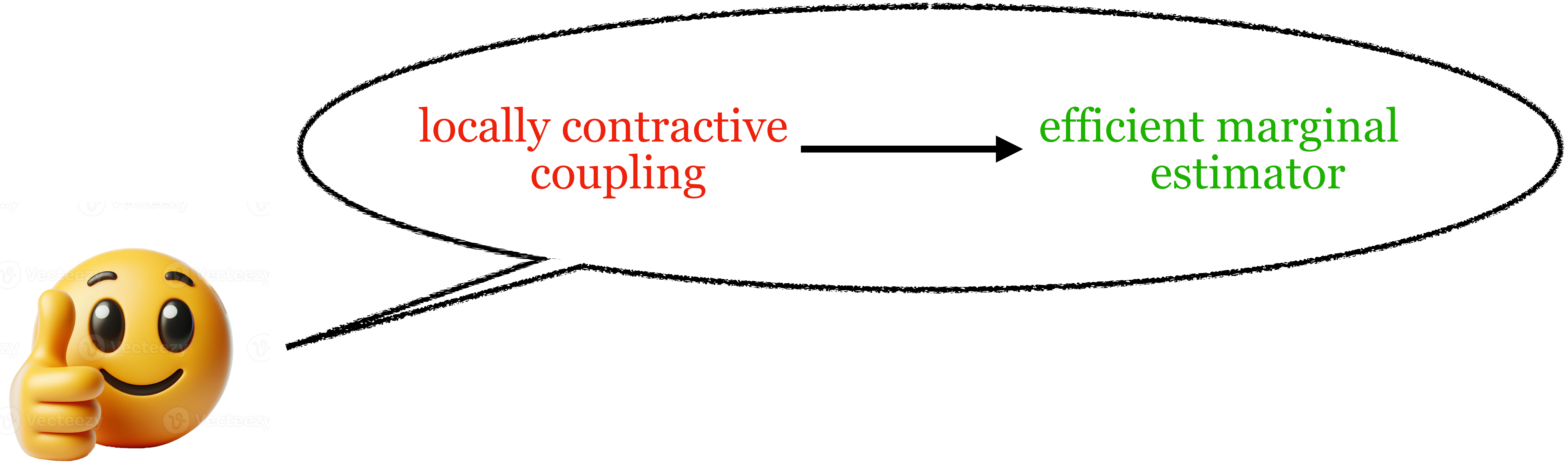
We cannot really run the coupling, but we can write down linear programs that encode coupling errors to bootstrap the marginal probability.

This method was invented by Moitra [Moi '19], applied in other works for sampling/counting LLL, [GLLZ '19, JPV '21b], and has recently been applied to other sampling/counting settings. [HLQZ '24]

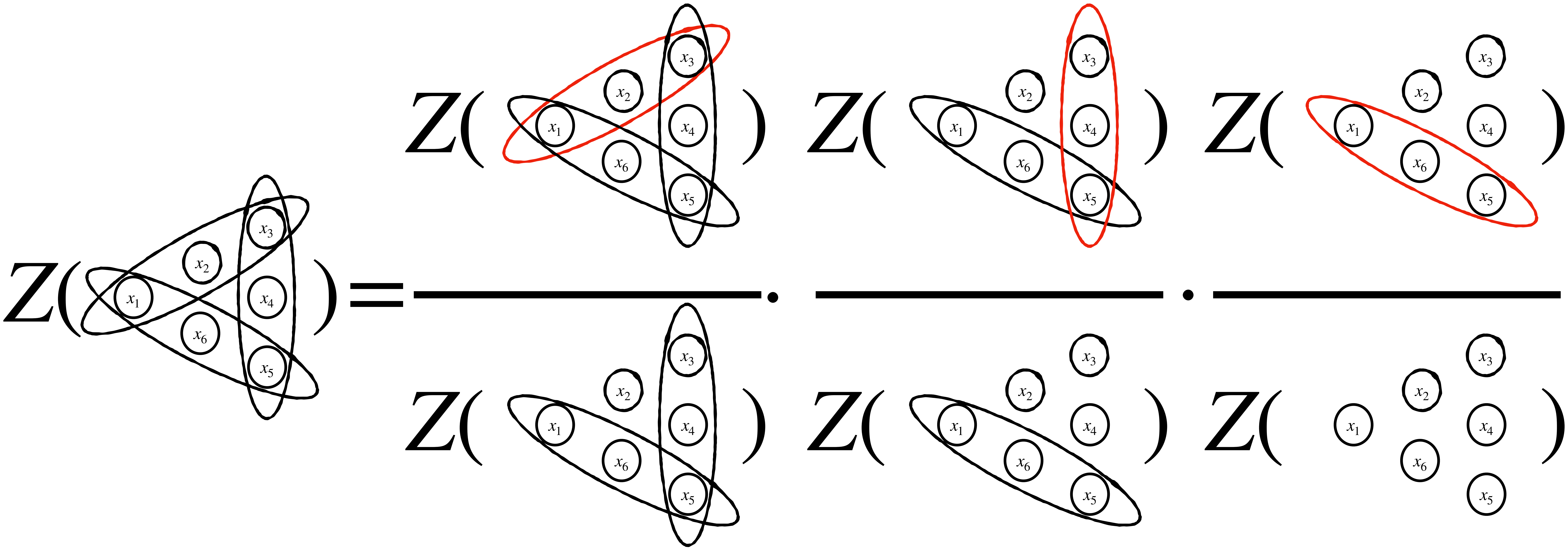
Algorithmic implications

We cannot really run the coupling, but we can write down linear programs that encode coupling errors to bootstrap the marginal probability.

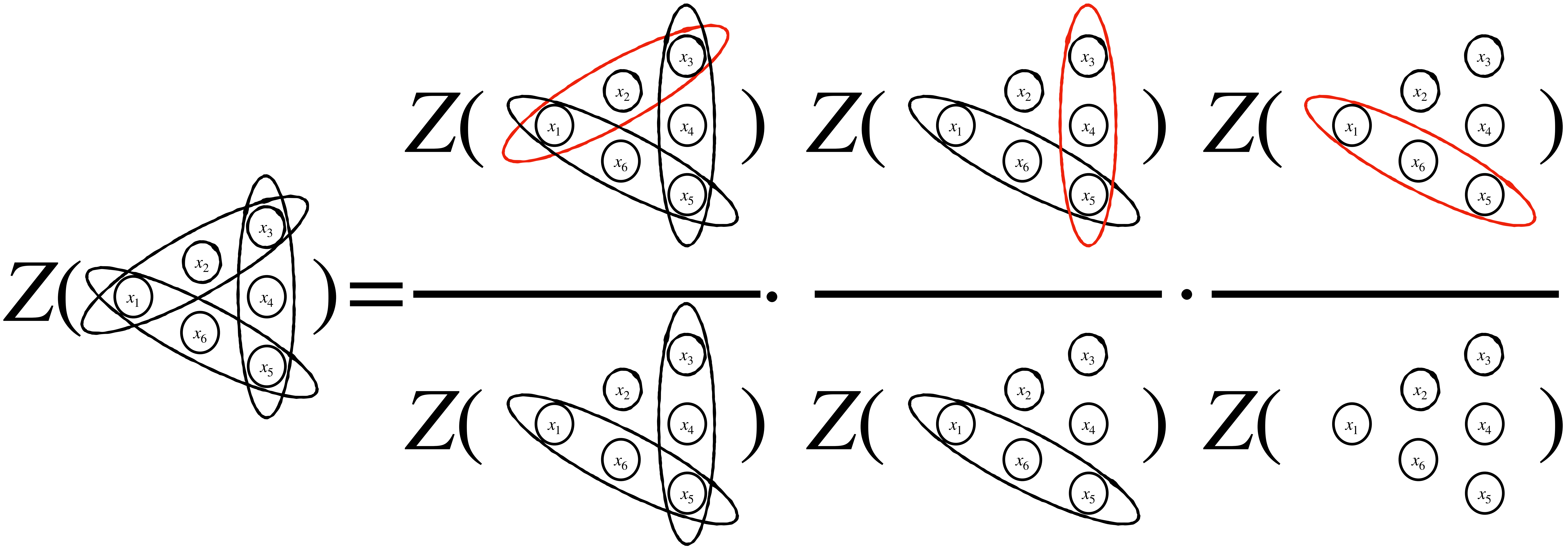
This method was invented by Moitra [Moi '19], applied in other works for sampling/counting LLL, [GLLZ '19, JPV '21b], and has recently been applied to other sampling/counting settings. [HLQZ '24]



Constraint-wise self-reducibility



Constraint-wise self-reducibility



Marginal estimator for $\mu_{\mathscr{C}\setminus\{c_0\}}(c_0)$ ⟶ Efficient counting

Dynamic sampler that updates $X \sim \mu_{\mathscr{C}\setminus\{c_0\}}$ to $Y \sim \mu_{\mathscr{C}}$ ⟶ Efficient sampling

Summary

We present polynomial-time algorithms for approximate counting/almost uniform sampling atomic constraint satisfaction solutions in the regime of $pD^{2+o_q(1)} \lesssim 1$.

This regime almost matches the lower bound $pD^2 \lesssim 1$, and still improves over the previous best regime in the worst case of Boolean domains.

At the heart of our approach is a novel constraint-wise coupling for CSPs, which may be of independent interest.

Summary

We present polynomial-time algorithms for approximate counting/almost uniform sampling atomic constraint satisfaction solutions in the regime of $pD^{2+o_q(1)} \lesssim 1$.

This regime almost matches the lower bound $pD^2 \lesssim 1$, and still improves over the previous best regime in the worst case of Boolean domains.

At the heart of our approach is a novel constraint-wise coupling for CSPs, which may be of independent interest.

Open Problems

sampling Lovász local lemma …

- … for small domain sizes? (especially, k-CNF)
- … for general CSPs?
- … with a faster running time? (our result works in $n^{\mathrm{poly}(k,D,\log q)}$ time)

Summary

We present polynomial-time algorithms for approximate counting/almost uniform sampling atomic constraint satisfaction solutions in the regime of $pD^{2+o_q(1)} \lesssim 1$.
This regime almost matches the lower bound $pD^2 \lesssim 1$, and still improves over the previous best regime in the worst case of Boolean domains.
At the heart of our approach is a novel constraint-wise coupling for CSPs, which may be of independent interest.

Thank you!

Open Problems

sampling Lovász local lemma …

- … for small domain sizes? (especially, k-CNF)
- … for general CSPs?
- … with a faster running time? (our result works in $n^{\mathrm{poly}(k,D,\log q)}$ time)