Chaînes de Markov, Génération Aléatoire, et Couplage Depuis le Passé (5)

Petite École de Combinatoire

P. Duchon

LaBRI

Quelques exemples

Deux exemples d'évaluation du temps de mélange par les chemins canoniques :

- colorations d'un graphe vide(!)
- couplages d'un graphe quelconque

Colorations d'un graphe

On considère un graphe G, de degré maximal Δ, et un entier $\Delta' > \Delta$.

- Ω (états) : ensemble des Δ'-colorations de G;

Colorations d'un graphe

On considère un graphe G, de degré maximal Δ, et un entier $\Delta' > \Delta$.

- Ω (états) : ensemble des Δ'-colorations de G;
- π : distribution uniforme sur Ω;

Colorations d'un graphe

On considère un graphe G, de degré maximal Δ, et un entier $\Delta' > \Delta$.

- Ω (états) : ensemble des Δ'-colorations de G;
- π : distribution uniforme sur Ω;
- transitions : entre deux colorations c et c' si et seulement si c et c' diffèrent en un seul sommet de G; on pose alors $P_{c,c'} = \frac{1}{2n\Delta'}$ ($P_{c,c} = \ldots \geq 1/2$)

Colorations d’un graphe

On considère un graphe G, de degré maximal Δ, et un entier $\Delta' > \Delta$.

- Ω (états) : ensemble des Δ'-colorations de G;
- π : distribution uniforme sur Ω;
- Algorithme de transition : c l’état courant;
  - avec probabilité $1/2$, poser c' et terminer;
  - choisir un sommet aléatoire uniforme u et une couleur aléatoire uniforme i;
  - si i est une couleur légale pour u (i.e., aucun voisin de u n’est de couleur i), c' est obtenu en recoloriant u en i, sinon $c' = c$

Colorations d'un graphe

- $P_{c,c'} = P_{c',c}$: la chaîne est inversible avec π comme distribution stationnaire.

Colorations d'un graphe

- $P_{c,c'} = P_{c',c}$: la chaîne est inversible avec π comme distribution stationnaire.
- Les probabilités de boucles sont toutes au moins $1/2$.

Colorations d'un graphe

- $P_{c,c'} = P_{c',c}$: la chaîne est inversible avec π comme distribution stationnaire.
- Les probabilités de boucles sont toutes au moins $1/2$.
- Si Δ' est assez grand, le graphe des colorations est connexe; $\Delta' \geq 2\Delta + 1$ suffit (étant données deux colorations c et c', on forme facilement un chemin de longueur au plus $2n$ de c à c')

Colorations d'un graphe

- $P_{c,c'} = P_{c',c}$: la chaîne est inversible avec π comme distribution stationnaire.
- Les probabilités de boucles sont toutes au moins $1/2$.
- Si Δ' est assez grand, le graphe des colorations est connexe; $\Delta' \geq 2\Delta + 1$ suffit (étant données deux colorations c et c', on forme facilement un chemin de longueur au plus $2n$ de c à c')

Donc dans ce cas, la chaîne converge vers la distribution uniforme sur Ω.

Colorations d'un graphe

- $P_{c,c'} = P_{c',c}$: la chaîne est inversible avec π comme distribution stationnaire.
- Les probabilités de boucles sont toutes au moins $1/2$.
- Si Δ' est assez grand, le graphe des colorations est connexe; $\Delta' \geq 2\Delta + 1$ suffit (étant données deux colorations c et c', on forme facilement un chemin de longueur au plus $2n$ de c à c')

Donc dans ce cas, la chaîne converge vers la distribution uniforme sur Ω.
Que peut-on dire de la vitesse de convergence ?

Colorations d'un graphe

- $P_{c,c'} = P_{c',c}$: la chaîne est inversible avec π comme distribution stationnaire.
- Les probabilités de boucles sont toutes au moins $1/2$.
- Si Δ' est assez grand, le graphe des colorations est connexe; $\Delta' \geq 2\Delta + 1$ suffit (étant données deux colorations c et c', on forme facilement un chemin de longueur au plus $2n$ de c à c')

Donc dans ce cas, la chaîne converge vers la distribution uniforme sur Ω.
Que peut-on dire de la vitesse de convergence ?
(dans le cas général: résultats obtenus par couplage)

Colorations d'un graphe

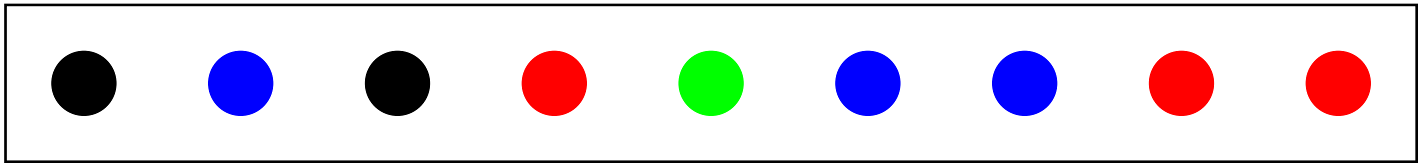
- $P_{c,c'} = P_{c',c}$: la chaîne est inversible avec π comme distribution stationnaire.
- Les probabilités de boucles sont toutes au moins $1/2$.
- Si Δ' est assez grand, le graphe des colorations est connexe; $\Delta' \geq 2\Delta + 1$ suffit (étant données deux colorations c et c', on forme facilement un chemin de longueur au plus $2n$ de c à c')

Donc dans ce cas, la chaîne converge vers la distribution uniforme sur Ω.
Que peut-on dire de la vitesse de convergence ?
(on va utiliser le cas trivial du graphe vide pour illustrer la méthode des chemins canoniques)

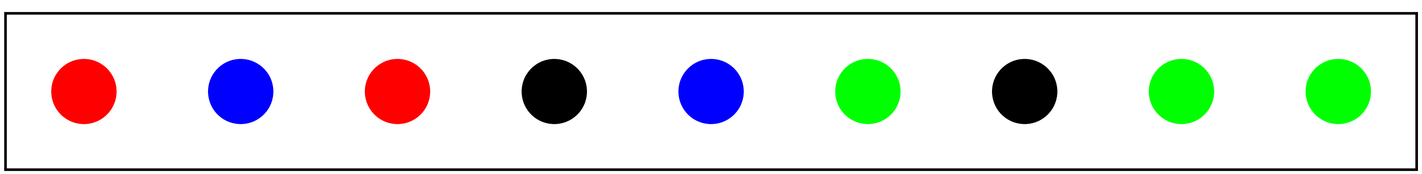
Colorations du graphe vide

Définir des chemins canoniques :

X

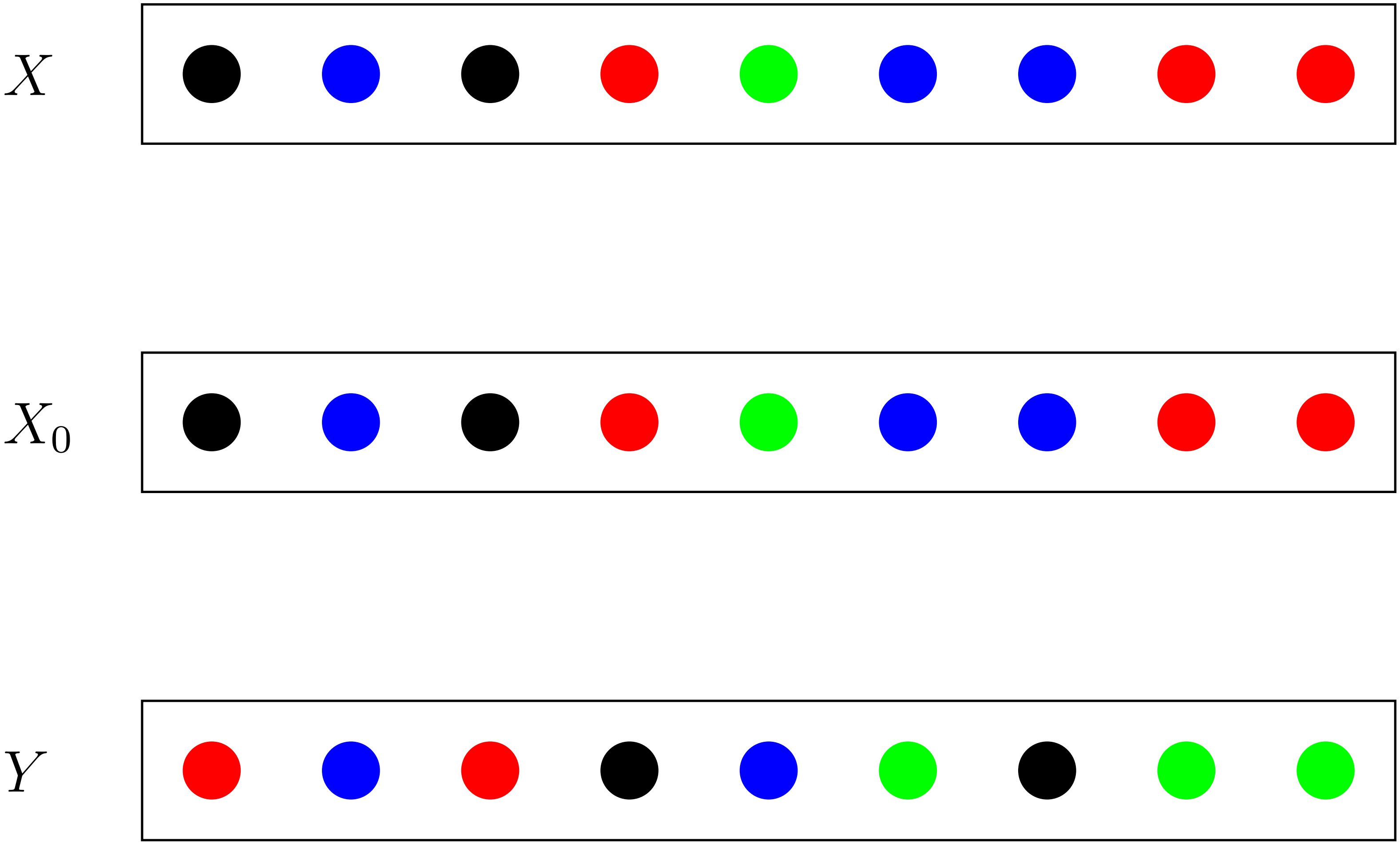


Y



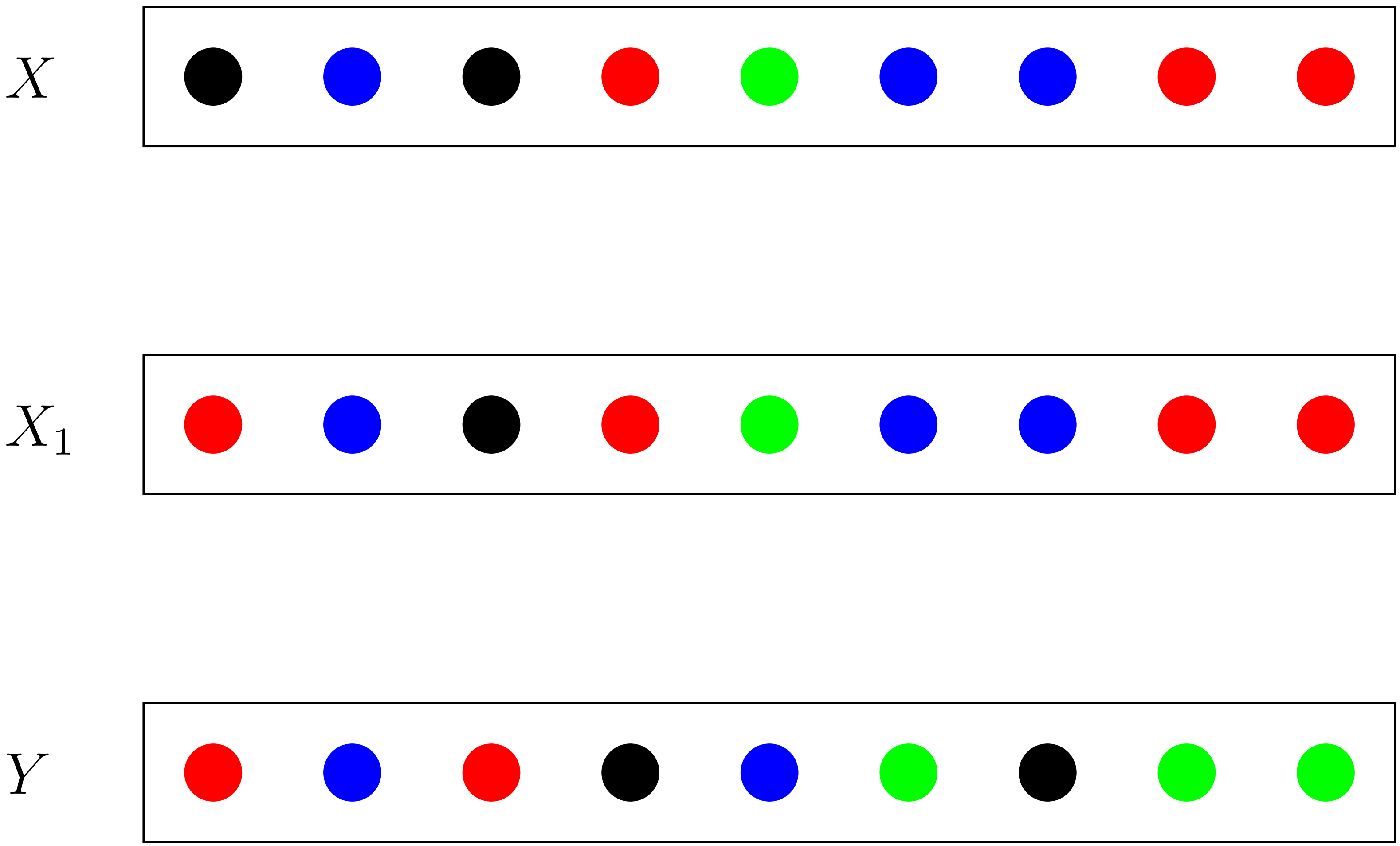
Colorations du graphe vide

Définir des chemins canoniques :



Colorations du graphe vide

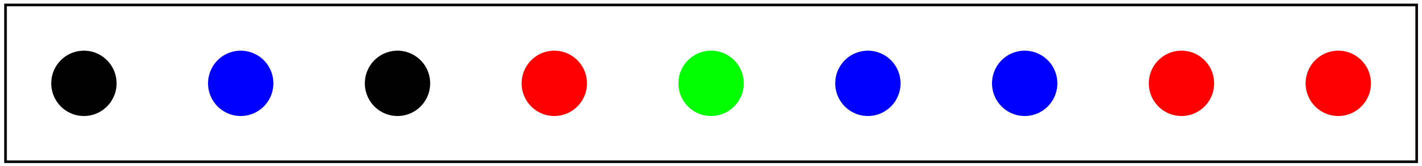
Définir des chemins canoniques :



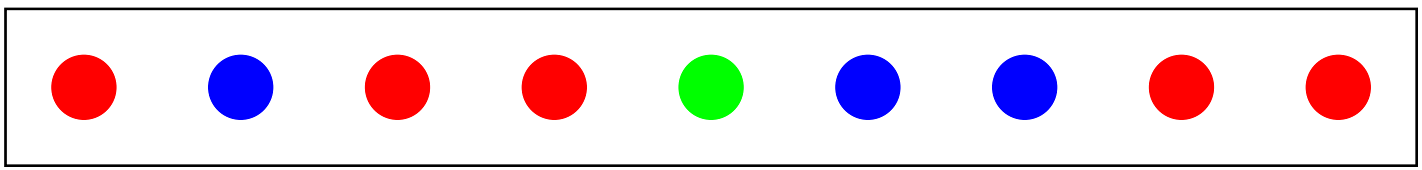
Colorations du graphe vide

Définir des chemins canoniques :

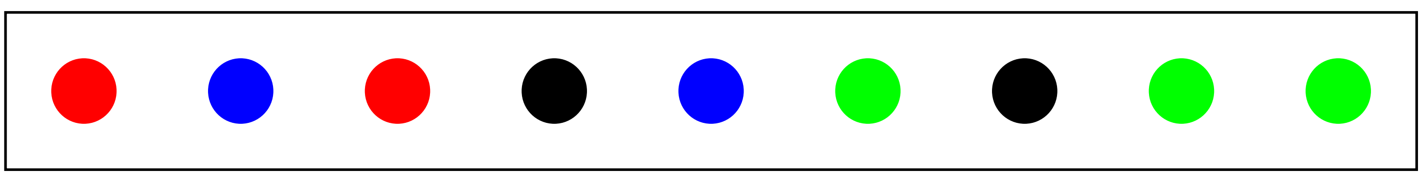
X



X_2

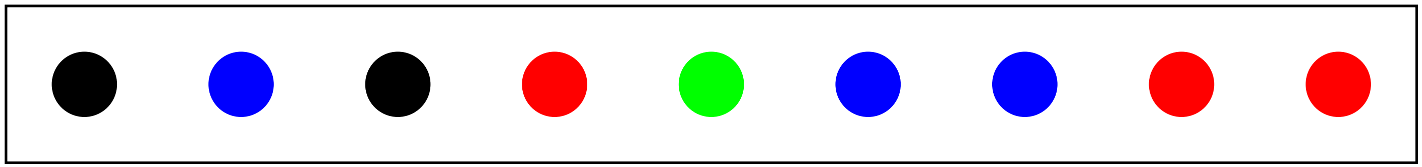


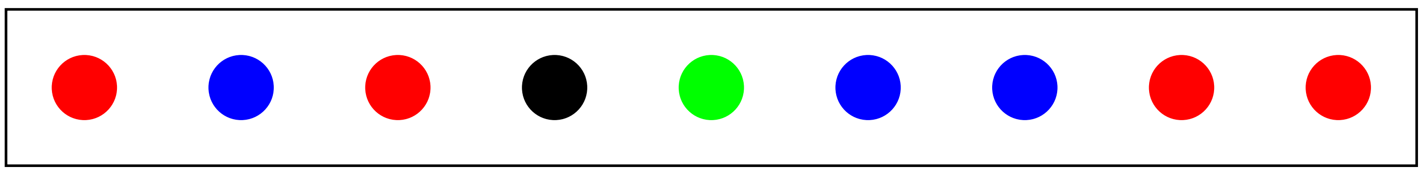
Y

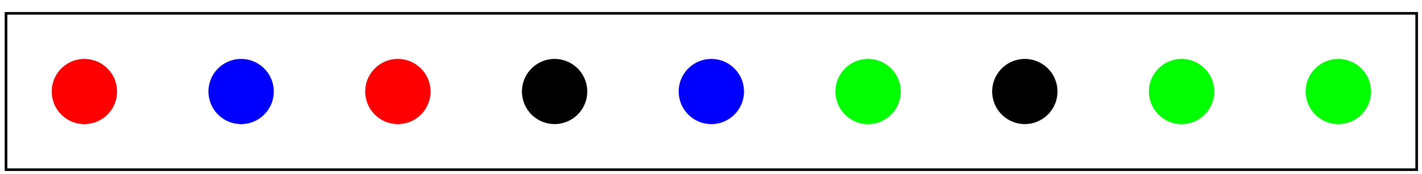


Colorations du graphe vide

Définir des chemins canoniques :

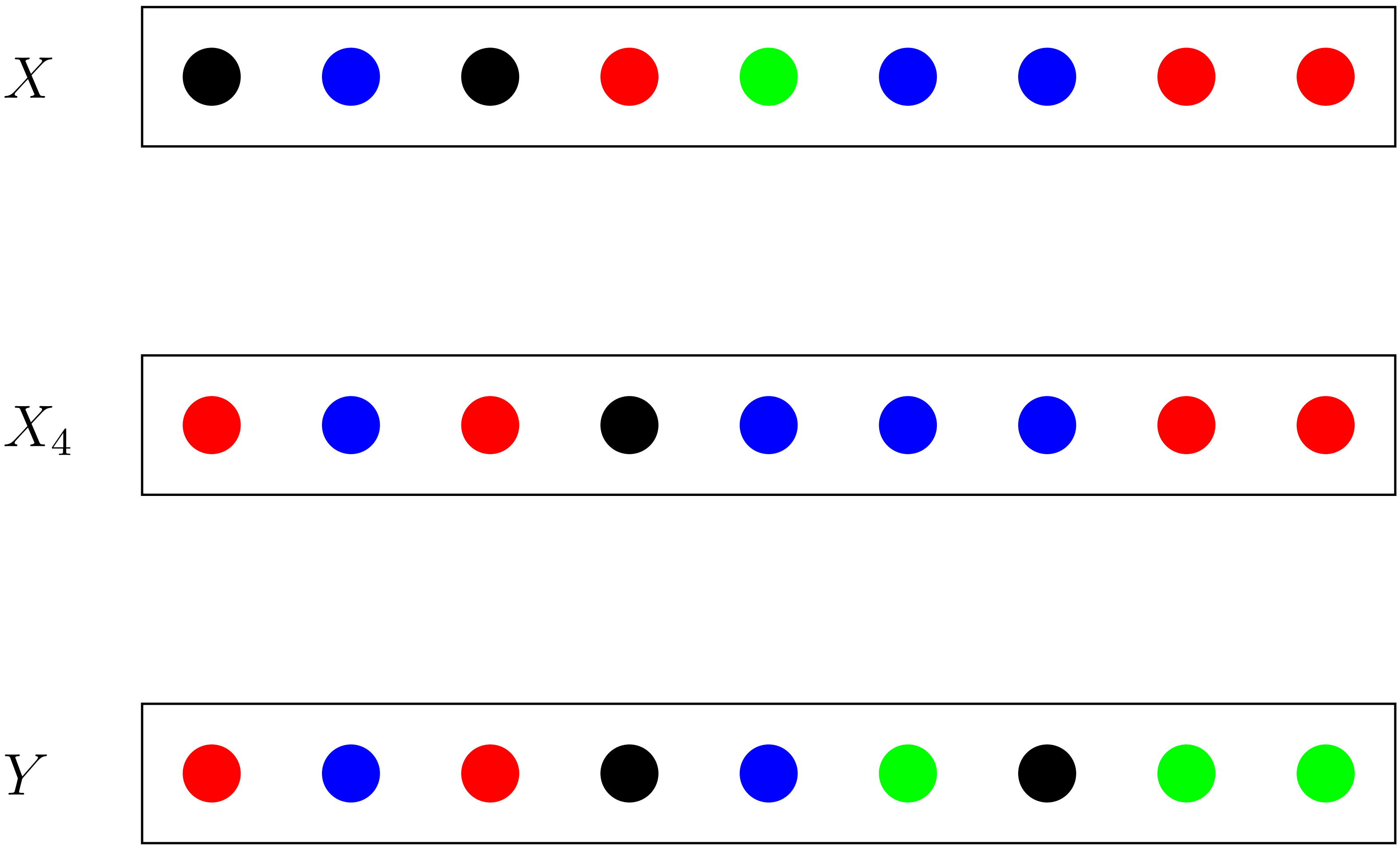
X 

X_3 

Y 

Colorations du graphe vide

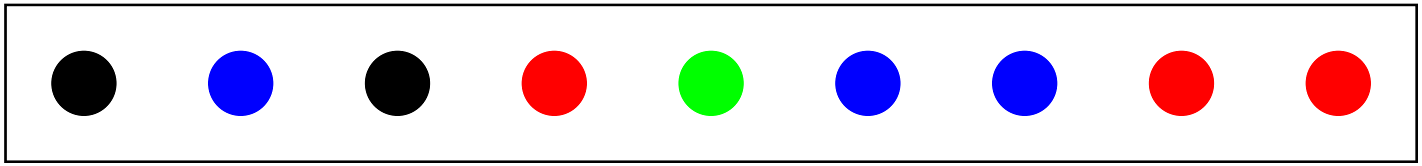
Définir des chemins canoniques :



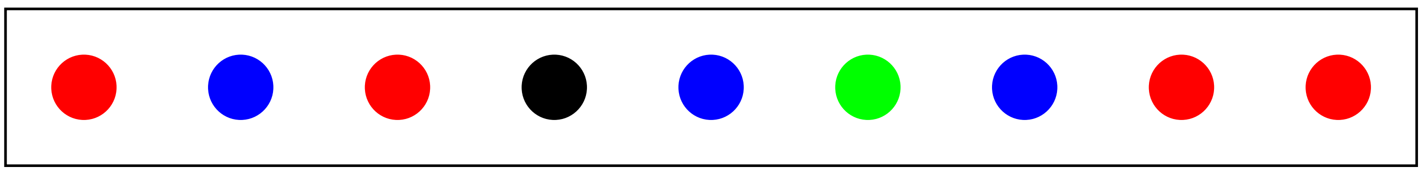
Colorations du graphe vide

Définir des chemins canoniques :

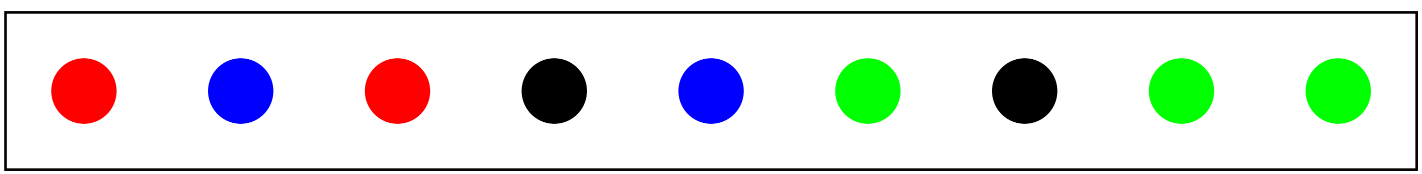
X



X_5



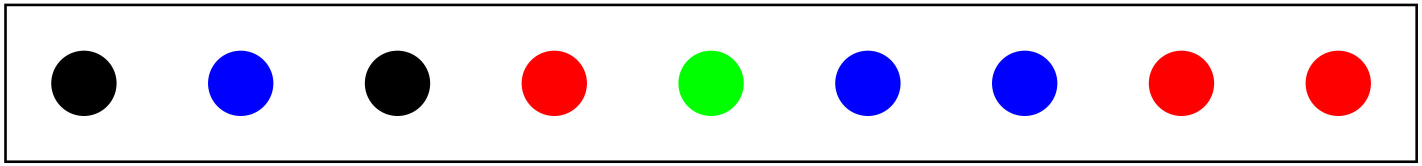
Y



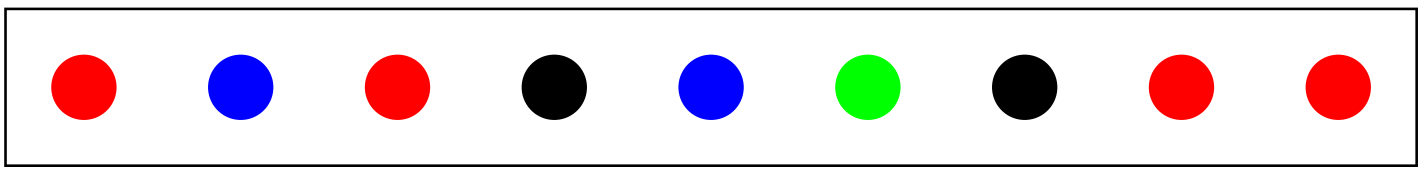
Colorations du graphe vide

Définir des chemins canoniques :

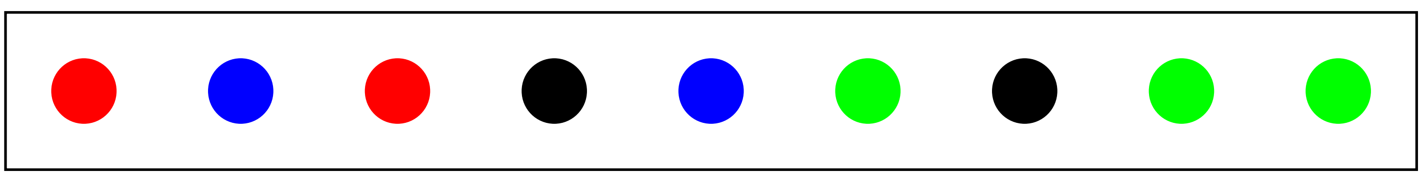
X



X_6

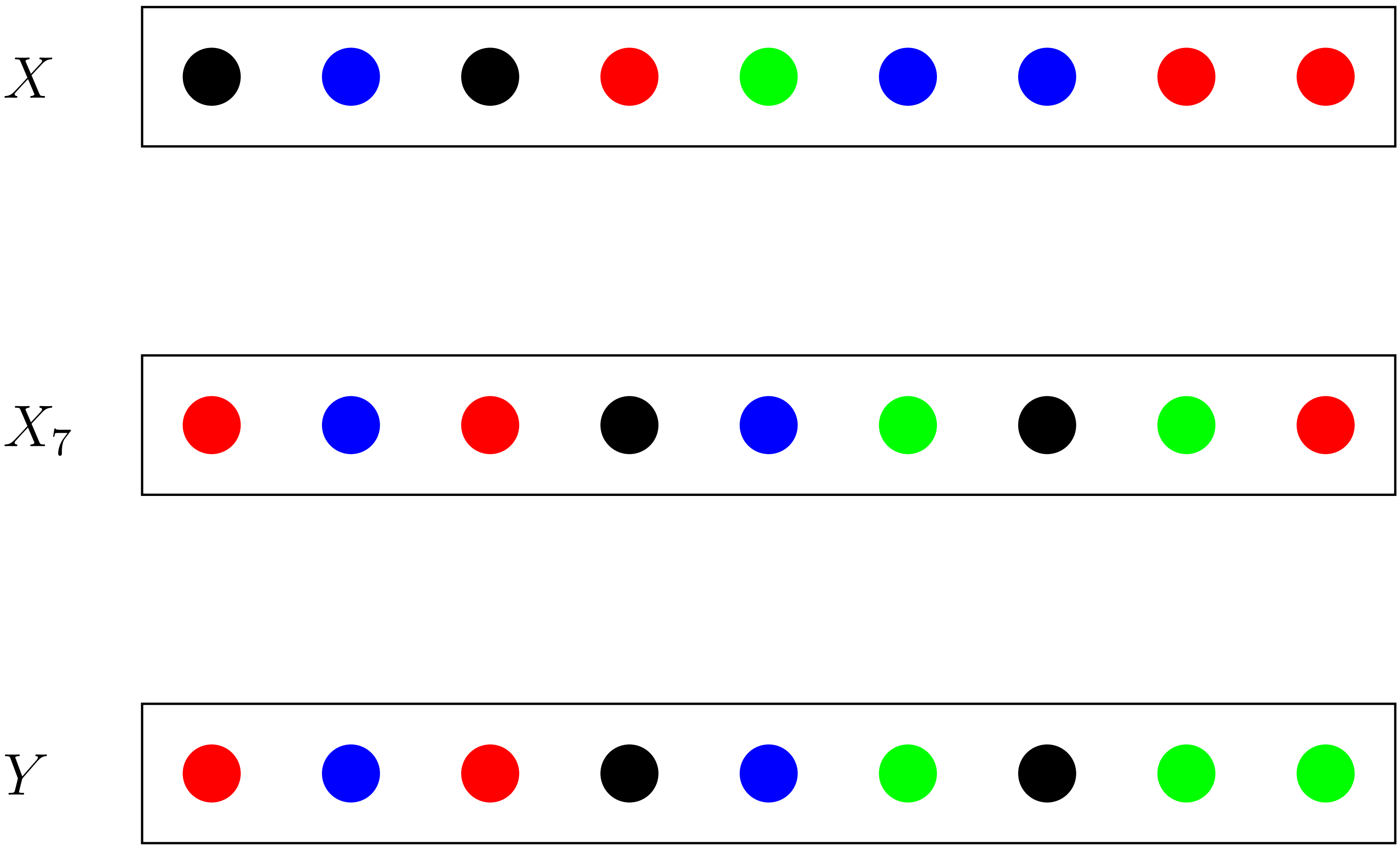


Y



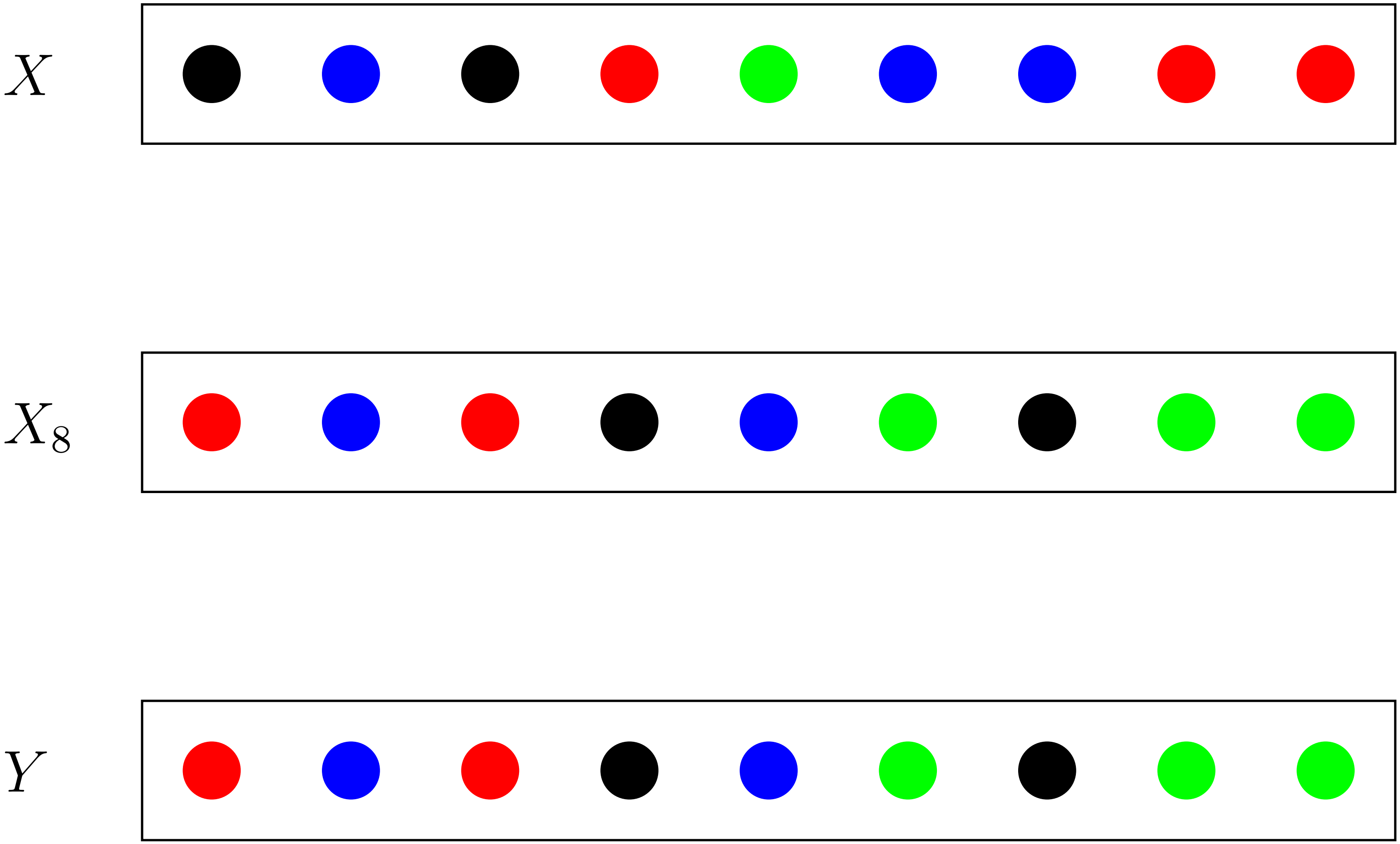
Colorations du graphe vide

Définir des chemins canoniques :



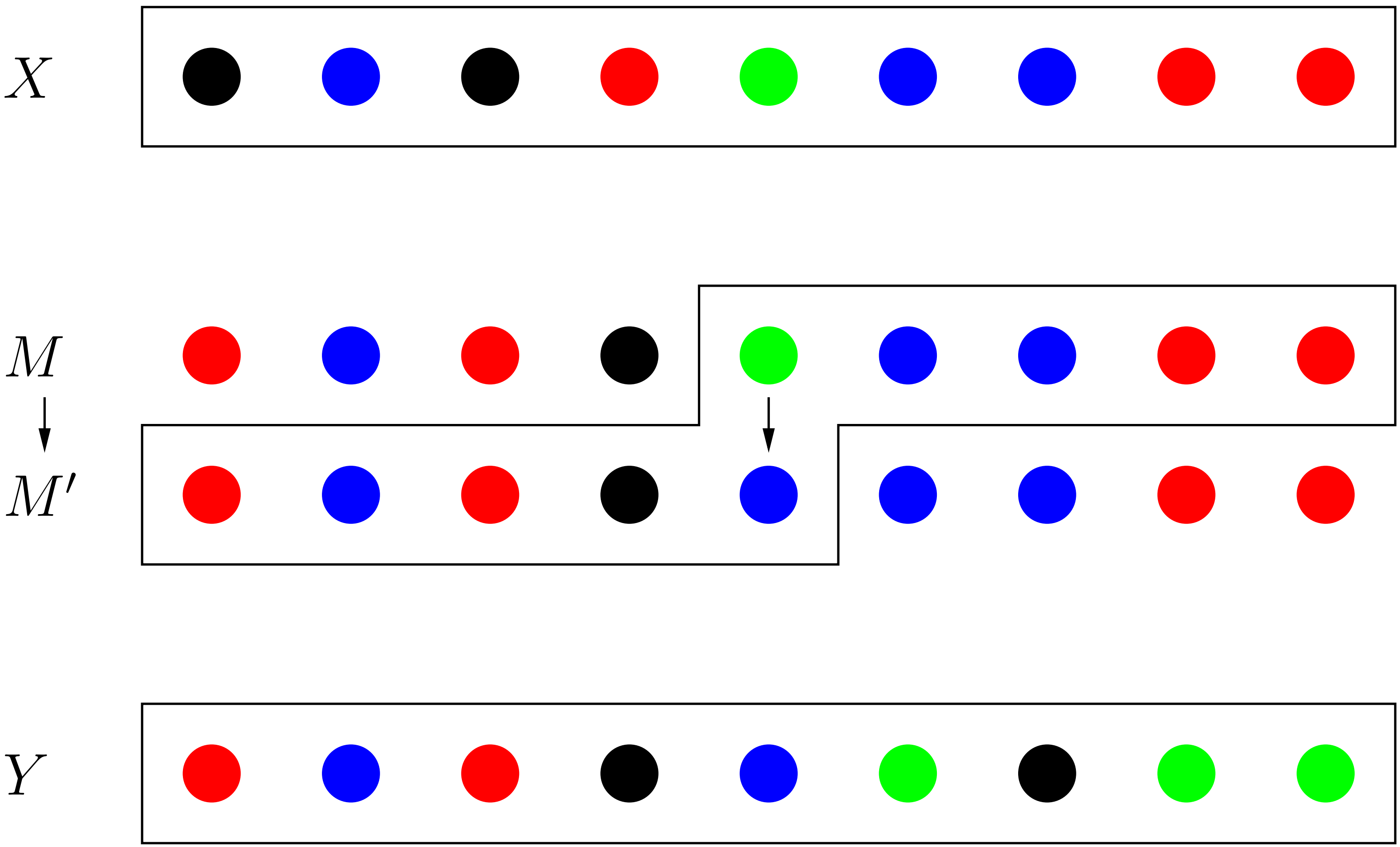
Colorations du graphe vide

Définir des chemins canoniques :



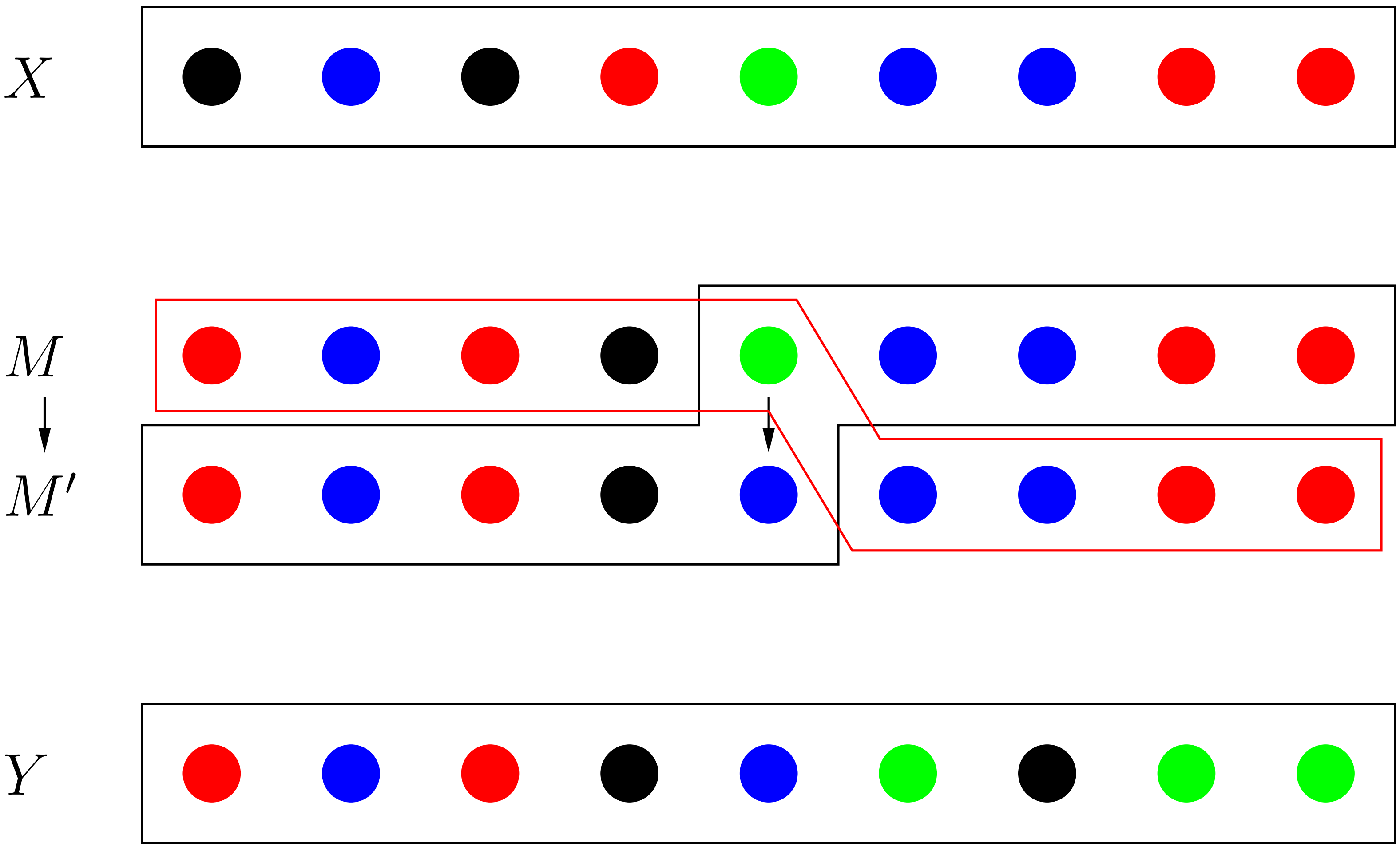
Colorations du graphe vide

Définir des chemins canoniques :



Colorations du graphe vide

Définir des chemins canoniques :



Calcul de ρ et ρ'

$\mathsf{CC}(e) = \{(x,y) : e \in \gamma_{x,y}\}$

$$\rho(e) = \frac{1}{Q(e)} \sum_{(x,y)\in \mathsf{CC}(e)} \pi(x)\pi(y)$$

Calcul de ρ et ρ'

$\mathsf{CC}(e) = \{(x,y) : e \in \gamma_{x,y}\}$

$$\rho'(e) = \frac{1}{Q(e)} \sum_{(x,y)\in\mathsf{CC}(e)} \pi(x)\pi(y)|\gamma_{x,y}|$$

Calcul de ρ et ρ'

$\textsf{CC}(e) = \{(x, y) : e \in \gamma_{x,y}\}$

$$\rho'(e) = \frac{1}{Q(e)} \sum_{(x,y)\in \textsf{CC}(e)} \pi(x)\pi(y)|\gamma_{x,y}|$$

- $|\Omega| = \Delta^n$, $|\gamma_{x,y}| \leq n$

Calcul de ρ et ρ'

$\mathsf{CC}(e) = \{(x,y) : e \in \gamma_{x,y}\}$

$$\rho'(e) = \frac{1}{Q(e)} \sum_{(x,y)\in \mathsf{CC}(e)} \pi(x)\pi(y)|\gamma_{x,y}|$$

- $|\Omega| = \Delta^n$, $|\gamma_{x,y}| \leq n$
- $|\mathsf{CC}(e)| = \Delta^{n-1}$

Calcul de ρ et ρ'

$\textsf{CC}(e) = \{(x,y) : e \in \gamma_{x,y}\}$

$$\rho'(e) = \frac{1}{Q(e)} \sum_{(x,y)\in \textsf{CC}(e)} \pi(x)\pi(y)|\gamma_{x,y}|$$

- $|\Omega| = \Delta^n$, $|\gamma_{x,y}| \leq n$
- $|\textsf{CC}(e)| = \Delta^{n-1}$
- $\rho(e) = 2n$, $\rho'(e) \leq 2n^2$

Calcul de ρ et ρ'

$\textsf{CC}(e) = \{(x,y) : e \in \gamma_{x,y}\}$

$$\rho'(e) = \frac{1}{Q(e)} \sum_{(x,y)\in \textsf{CC}(e)} \pi(x)\pi(y)|\gamma_{x,y}|$$

- $|\Omega| = \Delta^n$, $|\gamma_{x,y}| \leq n$
- $|\textsf{CC}(e)| = \Delta^{n-1}$
- $\rho(e) = 2n$, $\rho'(e) \leq 2n^2$

Donne comme bornes sur le temps de mélange :

$$(\rho) \qquad \tau(\epsilon) \leq 8\rho^2 \left(\ln(1/\pi_{\min}) + \ln(1/\epsilon)\right) = 32n^2 \left(n\ln(\Delta) + \ln(1/\epsilon)\right)$$

$$(\rho') \qquad \tau(\epsilon) \leq \rho' \left(\ln(1/\pi_{\min}) + \ln(1/\epsilon)\right) = 2n^2 \left((n\ln(\Delta) + \ln(1/\epsilon)\right)$$

Couplages d'un graphe

G un graphe (simple non orienté sans boucles) dont les arêtes ont des poids $w(e)$, Ω ensemble des couplages de G.

$$\begin{aligned} w(M) &= \prod_{e \in M} w(e) \\ \pi(M) &= \frac{w(M)}{\sum_{M' \in \Omega} w(M')} \end{aligned}$$

On fabrique facilement une chaîne de Markov ergodique, inversible, dont π soit la distibution stationnaire/limite.

Couplages: Chaîne de Markov

Algorithme de la chaîne de Markov : dans l'état M:

- Avec probabilité $1/2$, ne rien faire ($M' = M$);
- Choisir une arête $e = (u, v)$ uniforme;
- Déterminer un nouveau couplage M'' :
  - si $e \in M$, $M'' = M - e$
  - si u et v sont libres dans M, $M'' = M + e$
  - si un seul de u et v est libre dans M, et si l'autre est incident à une arête e', $M'' = M - e' + e$
  - sinon, $M'' = M$
- $M' = M''$ avec probabilité $\min(1, w(M')/w(M))$, sinon $M' = M$.

Vitesse de convergence ?

Couplages: Chaîne de Markov

Algorithme de la chaîne de Markov : dans l'état M:

- Avec probabilité $1/2$, ne rien faire ($M' = M$);
- Choisir une arête $e = (u, v)$ uniforme;
- Déterminer un nouveau couplage M'' :
  - si $e \in M$, $M'' = M - e$ ($\downarrow$-transition)
  - si u et v sont libres dans M, $M'' = M + e$
  - si un seul de u et v est libre dans M, et si l'autre est incident à une arête e', $M'' = M - e' + e$
  - sinon, $M'' = M$
- $M' = M''$ avec probabilité $\min(1, w(M')/w(M))$, sinon $M' = M$.

Vitesse de convergence ?

Couplages: Chaîne de Markov

Algorithme de la chaîne de Markov : dans l'état M:

- Avec probabilité $1/2$, ne rien faire ($M' = M$);
- Choisir une arête $e = (u, v)$ uniforme;
- Déterminer un nouveau couplage M'' :
  - si $e \in M$, $M'' = M - e$ (↓-transition)
  - si u et v sont libres dans M, $M'' = M + e$ (↑-transition)
  - si un seul de u et v est libre dans M, et si l'autre est incident à une arête e', $M'' = M - e' + e$
  - sinon, $M'' = M$
- $M' = M''$ avec probabilité $\min(1, w(M')/w(M))$, sinon $M' = M$.

Vitesse de convergence ?

Couplages: Chaîne de Markov

Algorithme de la chaîne de Markov : dans l'état M:

- Avec probabilité $1/2$, ne rien faire ($M' = M$);
- Choisir une arête $e = (u, v)$ uniforme;
- Déterminer un nouveau couplage M'' :
  - si $e \in M$, $M'' = M - e$ ($\downarrow$-transition)
  - si u et v sont libres dans M, $M'' = M + e$ ($\uparrow$-transition)
  - si un seul de u et v est libre dans M, et si l'autre est incident à une arête e', $M'' = M - e' + e$ ($\leftrightarrow$-transition)
  - sinon, $M'' = M$
- $M' = M''$ avec probabilité $\min(1, w(M')/w(M))$, sinon $M' = M$.

Vitesse de convergence ?

Chemins canoniques pour les couplages

On doit trouver, pour deux couplages X et Y quelconques, un chemin γ_{XY} : une suite de transitions qui fasse passer de X à Y, en essayant de faire en sorte que

Chemins canoniques pour les couplages

On doit trouver, pour deux couplages X et Y quelconques, un chemin γ_{XY} : une suite de transitions qui fasse passer de X à Y, en essayant de faire en sorte que

- les chemins ne soient pas trop longs

Chemins canoniques pour les couplages

On doit trouver, pour deux couplages X et Y quelconques, un chemin γ_{XY} : une suite de transitions qui fasse passer de X à Y, en essayant de faire en sorte que

- les chemins ne soient pas trop longs
- ils évitent de trop utiliser certaines transitions

Chemins canoniques pour les couplages

On doit trouver, pour deux couplages X et Y quelconques, un chemin γ_{XY} : une suite de transitions qui fasse passer de X à Y, en essayant de faire en sorte que

- les chemins ne soient pas trop longs
- ils évitent de trop utiliser certaines transitions

Par exemple, le choix consistant, pour passer de X à Y, à passer systématiquement par le couplage vide est probablement mauvais : les arêtes incidentes au couplage vide seront très chargées.

Chemins canoniques pour couplages

X et Y étant donnés : $Z = X \oplus Y$
Dans Z, chaque sommet est de degré 0, 1 ou 2 : ensemble de chemins et de cycles.

Chemins canoniques pour couplages

X et Y étant donnés : $Z = X \oplus Y$
Dans Z, chaque sommet est de degré 0, 1 ou 2 : ensemble de chemins et de cycles.
Ingrédients :

- définir un ordre total entre ces chemins et cycles;

Chemins canoniques pour couplages

X et Y étant donnés : $Z = X \oplus Y$
Dans Z, chaque sommet est de degré 0, 1 ou 2 : ensemble de chemins et de cycles.
Ingrédients :

- définir un ordre total entre ces chemins et cycles;
- "détricoter" chacun des chemins et cycles, pour passer de X à Y :

Chemins canoniques pour couplages

X et Y étant donnés : $Z = X \oplus Y$
Dans Z, chaque sommet est de degré 0, 1 ou 2 : ensemble de chemins et de cycles.
Ingrédients :

- définir un ordre total entre ces chemins et cycles;
- “détricoter” chacun des chemins et cycles, pour passer de X à Y :
  - pour un cycle, commencer par une $\downarrow$-transition (on obtient un chemin)

Chemins canoniques pour couplages

X et Y étant donnés : $Z = X \oplus Y$
Dans Z, chaque sommet est de degré 0, 1 ou 2 : ensemble de chemins et de cycles.
Ingrédients :

- définir un ordre total entre ces chemins et cycles;
- “détricoter” chacun des chemins et cycles, pour passer de X à Y :
  - pour un cycle, commencer par une $\downarrow$-transition (on obtient un chemin)
  - pour un chemin, une suite de $\leftrightarrow$-transitions, éventuellement suivie par une $\uparrow$- ou $\downarrow$-transition

Évaluation de ρ

On a (presque?) la possibilité, pour chaque transition $t = (M, M')$, de définir une injection $\eta_t : \mathsf{cc}(t) \rightarrow \Omega$, avec la propriété que

$$\forall X, Y \in \mathsf{cc}(t), \pi(X)\pi(Y) = \pi(M)\pi(\eta_t(X, Y))$$

Évaluation de ρ

On a (presque?) la possibilité, pour chaque transition $t = (M, M')$, de définir une injection η_t : cc$(t) \to \Omega$, avec la propriété que

$$\forall X, Y \in \textsf{cc}(t), \pi(X)\pi(Y) = \pi(M)\pi(\eta_t(X, Y))$$

On aurait alors

$$\rho(t) = \frac{1}{\pi(M)P_{M,M'}} \sum_{(X,Y)\in\textsf{CC}(t)} \pi(M)\pi(\eta_t(X, Y)) \leq \frac{1}{P_{M,M'}}$$

Évaluation de ρ

On a (presque?) la possibilité, pour chaque transition $t = (M, M')$, de définir une injection η_t : cc$(t) \to \Omega$, avec la propriété que

$$\forall X, Y \in \text{cc}(t), \pi(X)\pi(Y) = \pi(M)\pi(\eta_t(X, Y))$$

On aurait alors

$$\rho(t) = \frac{1}{\pi(M)P_{M,M'}} \sum_{(X,Y)\in \text{CC}(t)} \pi(M)\pi(\eta_t(X, Y)) \leq \frac{1}{P_{M,M'}}$$

(Pour le cas $w(e) = \lambda$, Jerrum et Sinclair obtiennent : $\rho' \leq 4m.n. \max(1, \lambda)$)