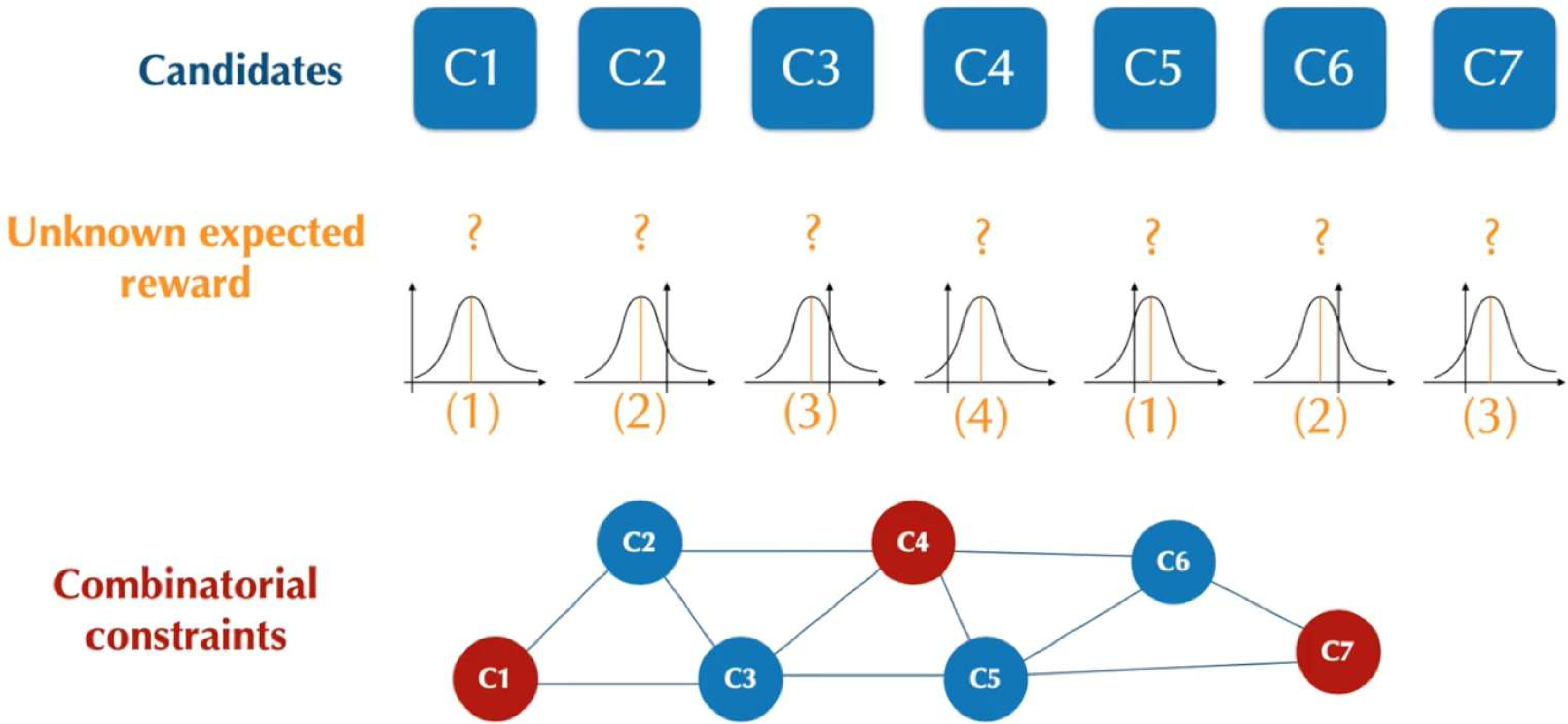
Change Detection for Online Matching Problems

Combinatorial constraints

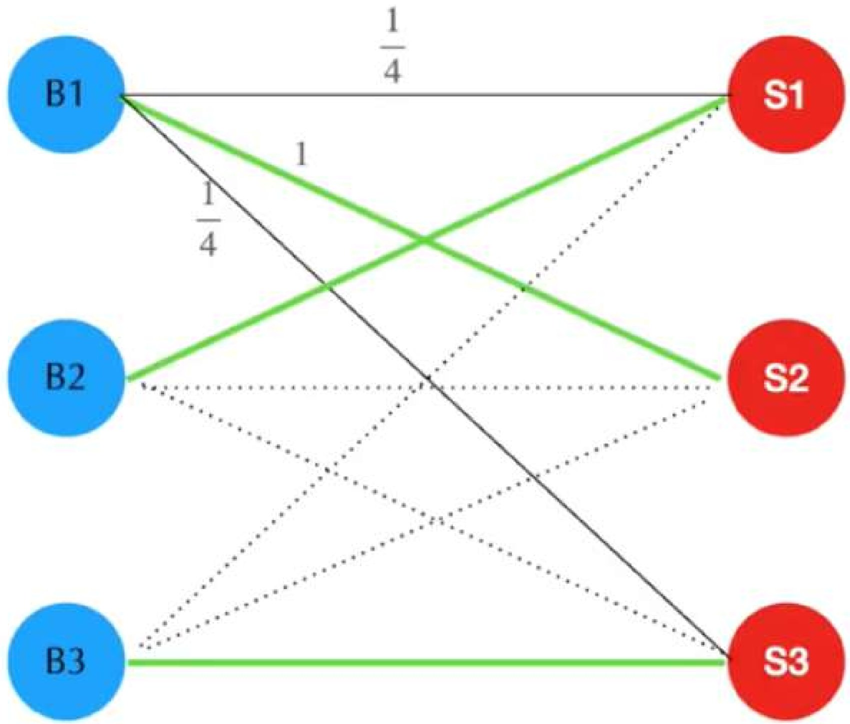


Arms, superarms, objective function

- An ***arm*** is a candidate
- A ***superarm*** is a collection of candidates
- A ***feasible superarm*** is a superarm satisfying the (combinatorial) constraints

$$\mathscr{C}(\mathbf{a}) = 0$$

- The ***reward*** of a superarm is the sum of the reward of arms it contains
- The ***goal*** is to maximize the cumulative, in time, expected reward



Adjacency Matrix

	S1	S2	S3
B1	1/4	1	1/4
B2	1/2	1/4	1/4
B3	1/4	1/4	1

UCB1 pseudocode

1. Play once every arm $a \in A$

2. At every time t play arm a_t such that

$$a_t \leftarrow \arg\max_{a \in A} \left\{ \bar{x}_a + \sqrt{\frac{2\log(t)}{n_a(t-1)}} \right\}$$

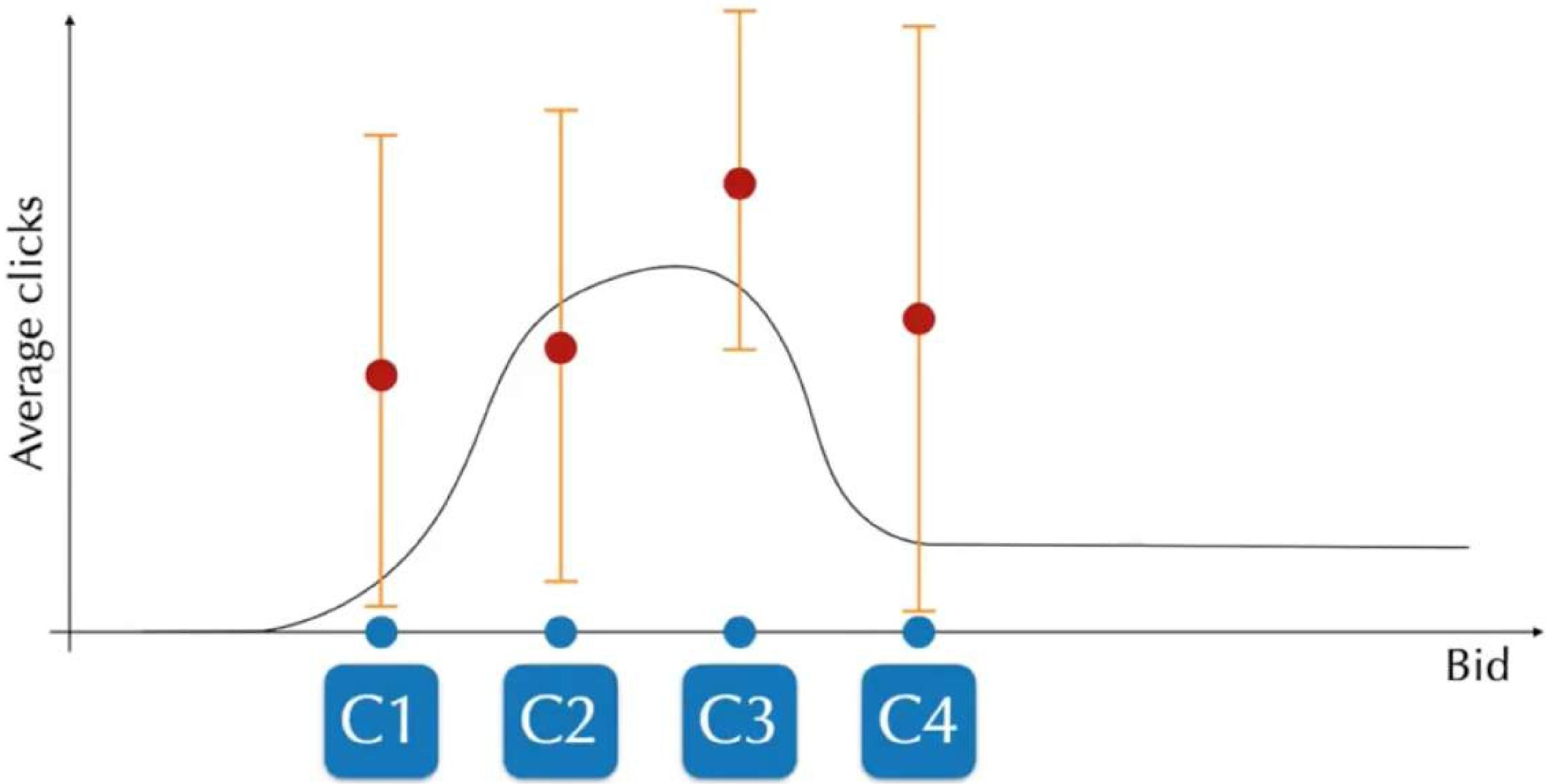
Changes in Online Matching Problems

	S1	S2	S3
B1	1/4	1	1/4
B2	1/2	1/4	1/4
B3	1/4	1/4	1

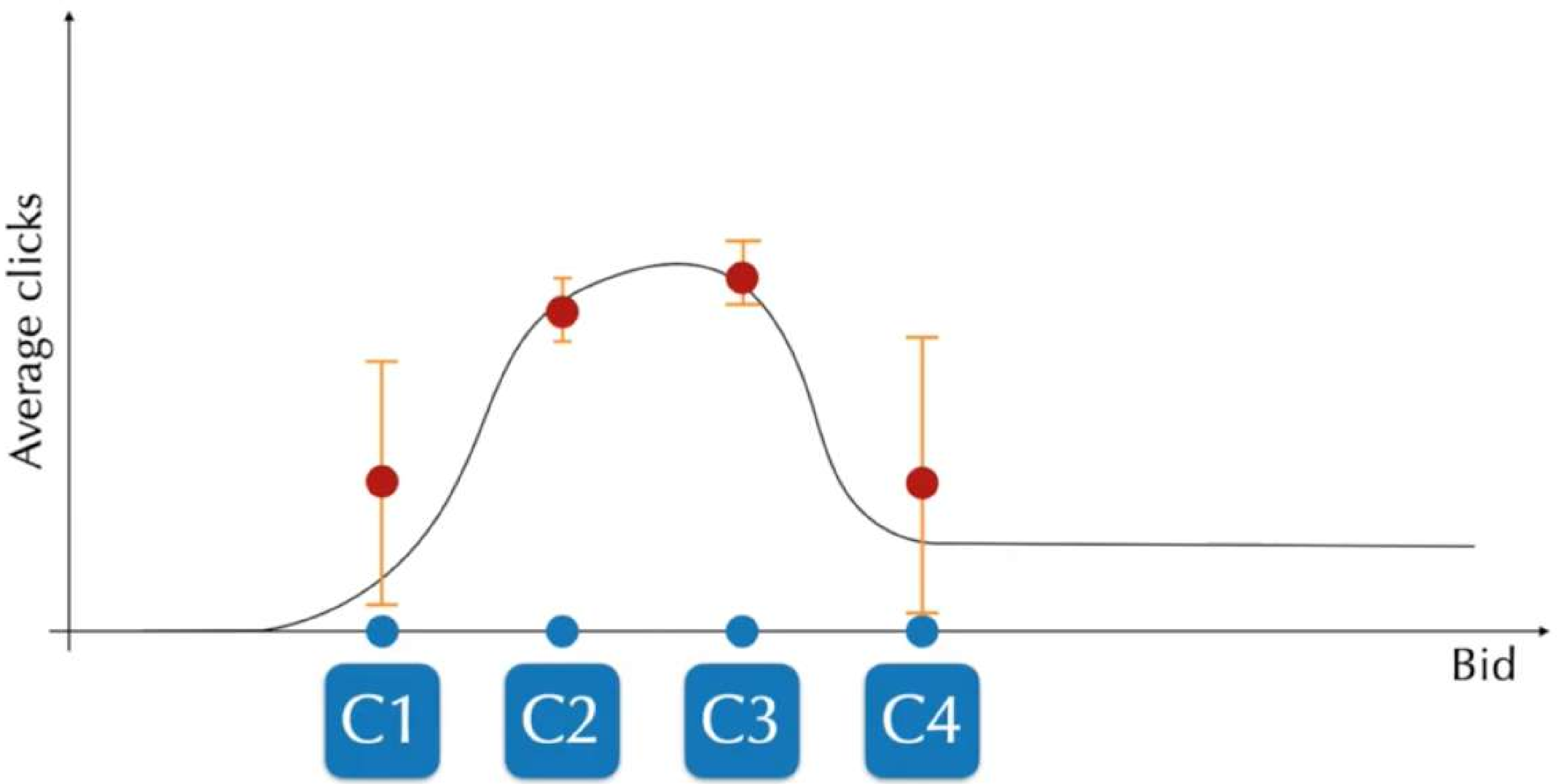
→

	S1	S2	S3
B1	1	1/4	1/4
B2	1/2	1/4	1/4
B3	1/4	1/4	1

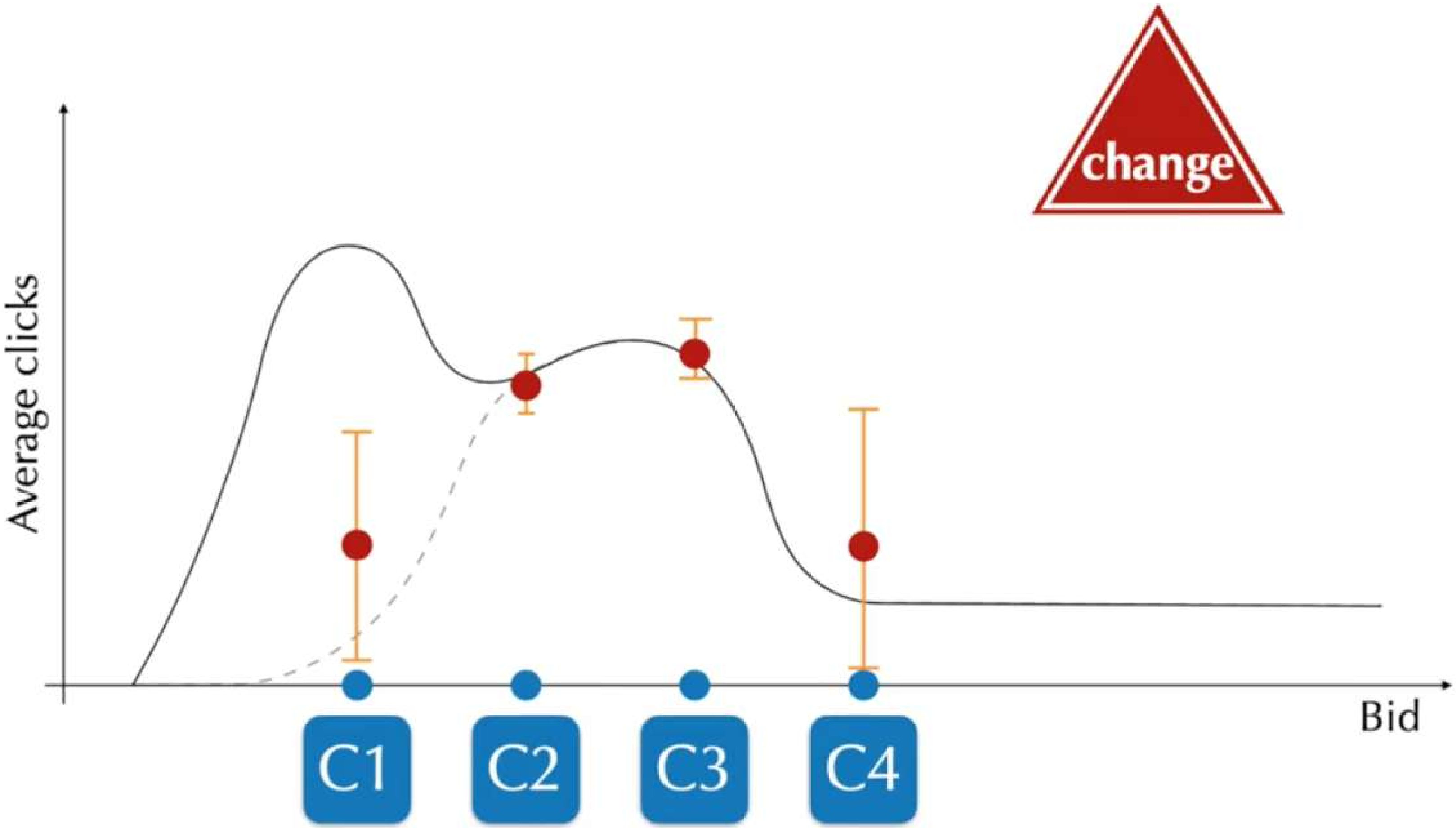
UCB1 and abrupt changes



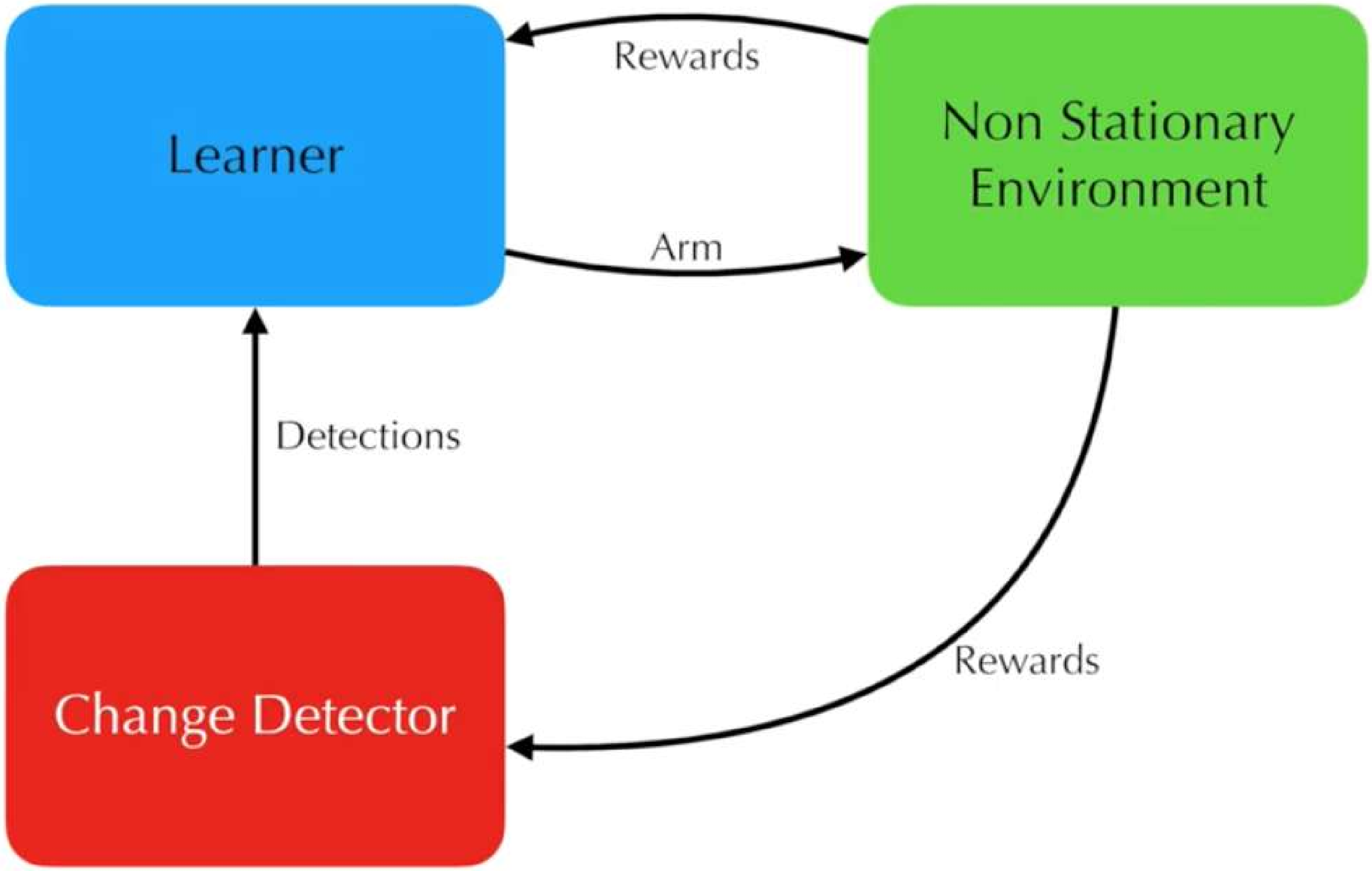
UCB1 and abrupt changes (after a lot of samples)



UCB1 and abrupt changes



Change Detection UCB



CD-UCB Pseudocode

1. inItialize τ_a=0 for each arm $a \in A$
2. for each t

$$a_t \leftarrow \arg\max_{a \in A} \left\{ \bar{x}_{a,\tau_a,t} + \sqrt{\frac{2\log(n(t))}{n_a(\tau_a, t-1)}} \right\} \text{ with probability } 1-\alpha$$

$$a_t \leftarrow \text{random arm with probability } \alpha$$

$n(t)$	is the total number of valid samples
$\bar{x}_{a,\tau_a,t}$	is the empirical mean of arm a over the last valid samples
$n_a(\tau_a, t-1)$	is the number of valid samples for arm a

3. collect reward r_t
4. if $CD_a(r_\tau, \ldots, r_t) = 1$ then $\tau_a = t$ and restart CD_a

Change Detection (CUSUM)

The first *M* valid samples are used to produce the *reference point*

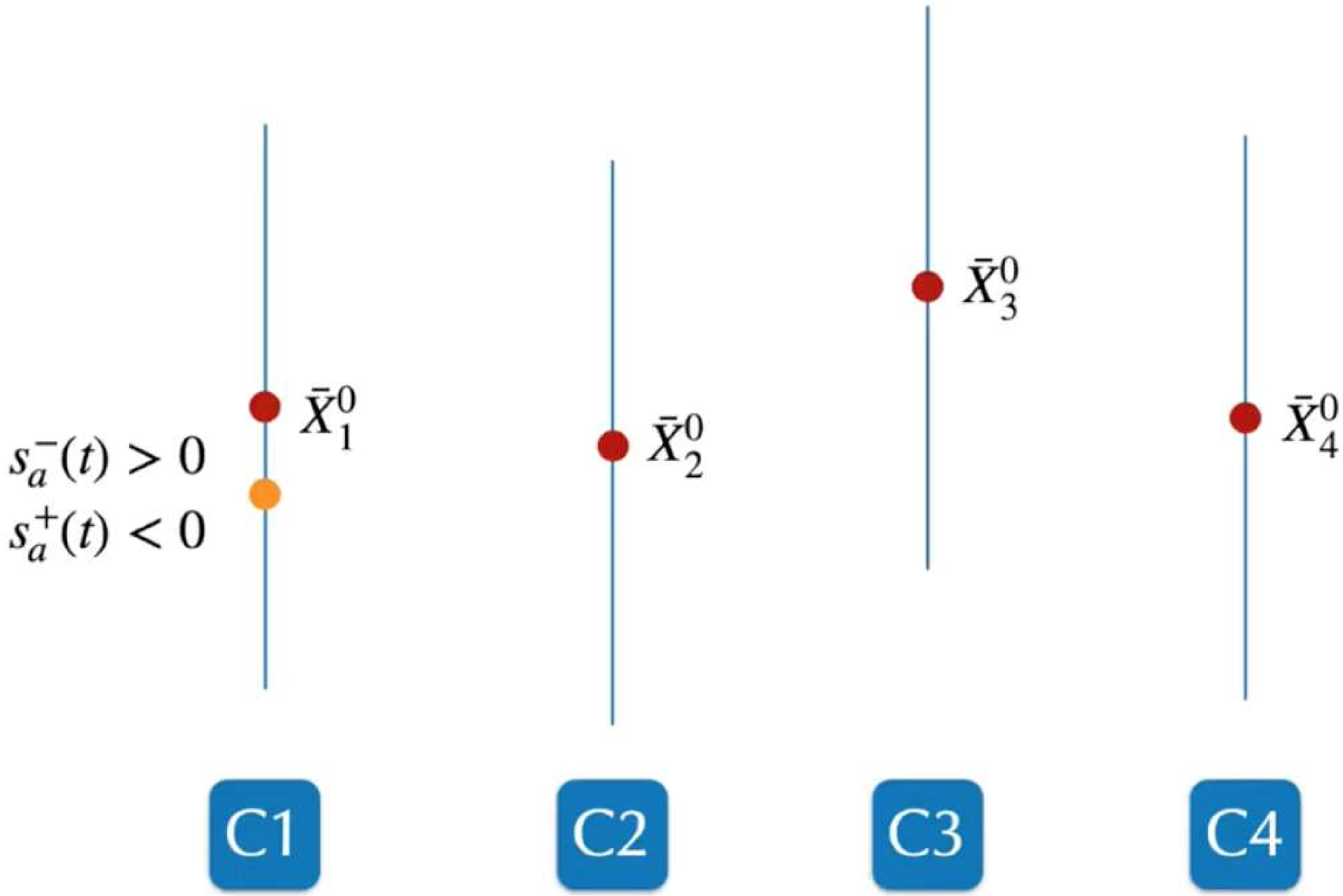
Empirical mean of arm *a* over the first *M* valid samples $\bar{X}_a^0$

From the *M+1-th* valid sample on, we check whether there is a change

Positive deviation from the reference point at *t* $s_a^+(t) = \left(x_a(t) - \bar{X}_a^0(t)\right) - \epsilon$

Negative deviation from the reference point at *t* $s_a^-(t) = -\left(x_a(t) - \bar{X}_a^0(t)\right) - \epsilon$

Example



Change Detection (CUSUM)

The first M valid samples are used to produce the *reference point*

Empirical mean of arm a over the first M valid samples $\bar{X}_a^0$

From the *M+1-th* valid sample on, we check whether there is a change

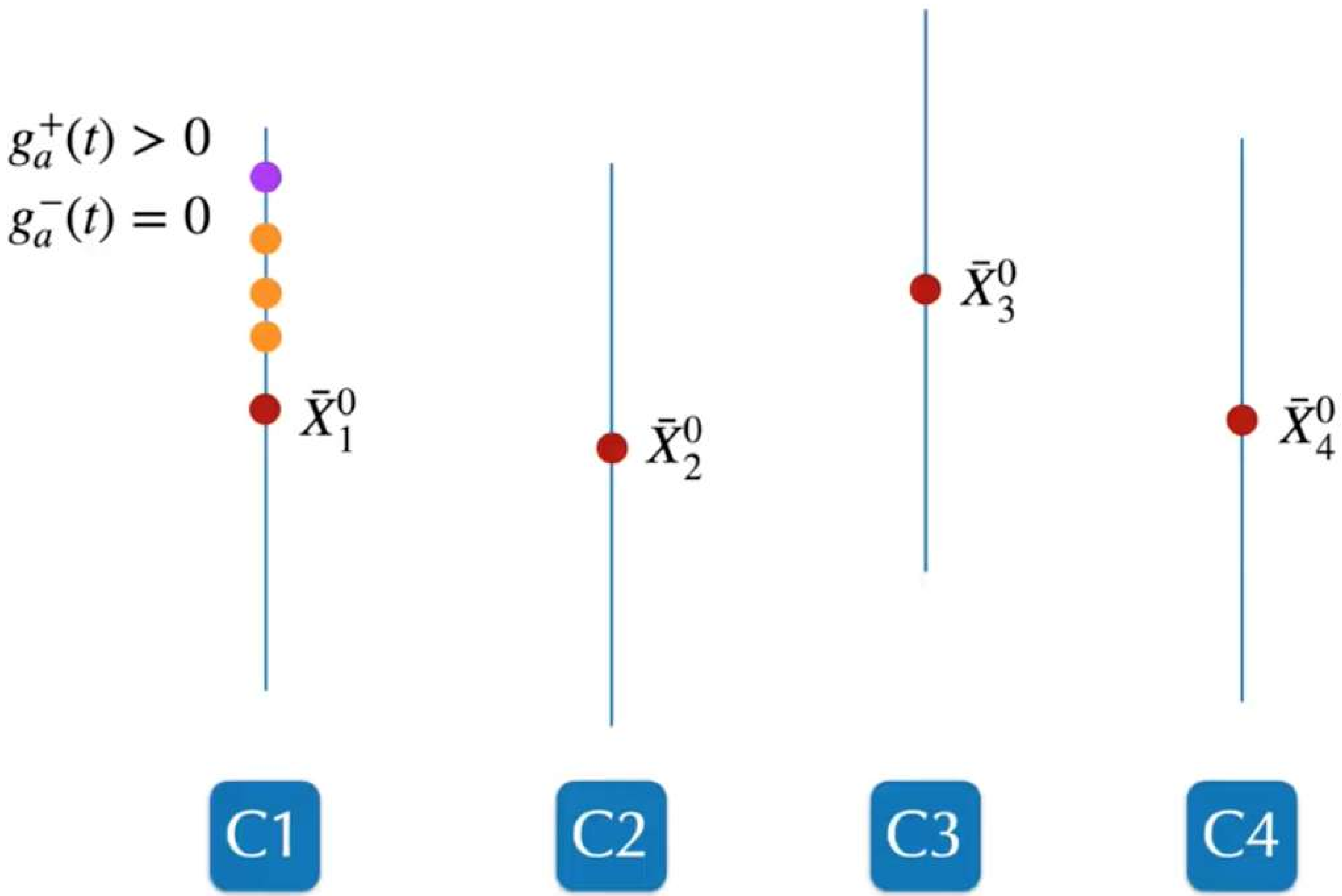
Positive deviation from the reference point at t $s_a^+(t) = \left(x_a(t) - \bar{X}_a^0(t)\right) - \epsilon$

Negative deviation from the reference point at t $s_a^-(t) = -\left(x_a(t) - \bar{X}_a^0(t)\right) - \epsilon$

Cumulative positive deviation from the reference point at t $g_a^+(t) = \max\left\{0, g_a^+(t-1) + s_a^+(t)\right\}$

Cumulative negative deviation from the reference point at t $g_a^-(t) = \max\left\{0, g_a^-(t-1) + s_a^-(t)\right\}$

Example



Change Detection (CUSUM)

The first M valid samples are used to produce the *reference point*

Empirical mean of arm a over the first M valid samples $\bar{X}_a^0$

From the *M+1-th* valid sample on, we check whether there is a change

Positive deviation from the reference point at t $s_a^+(t) = \left(x_a(t) - \bar{X}_a^0(t)\right) - \epsilon$

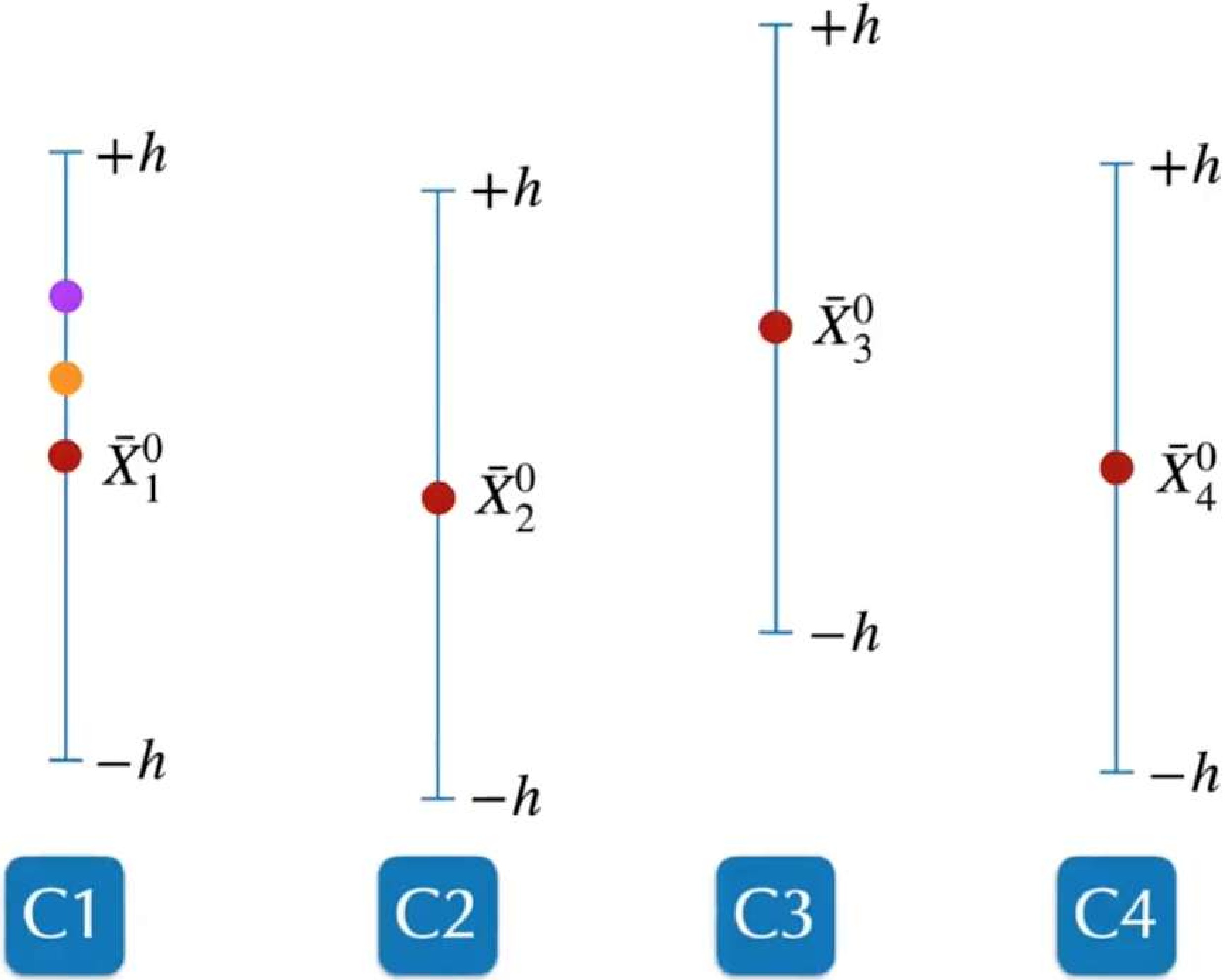
Negative deviation from the reference point at t $s_a^-(t) = -\left(x_a(t) - \bar{X}_a^0(t)\right) - \epsilon$

Cumulative positive deviation from the reference point at t $g_a^+(t) = \max\left\{0, g_a^+(t-1) + s_a^+(t)\right\}$

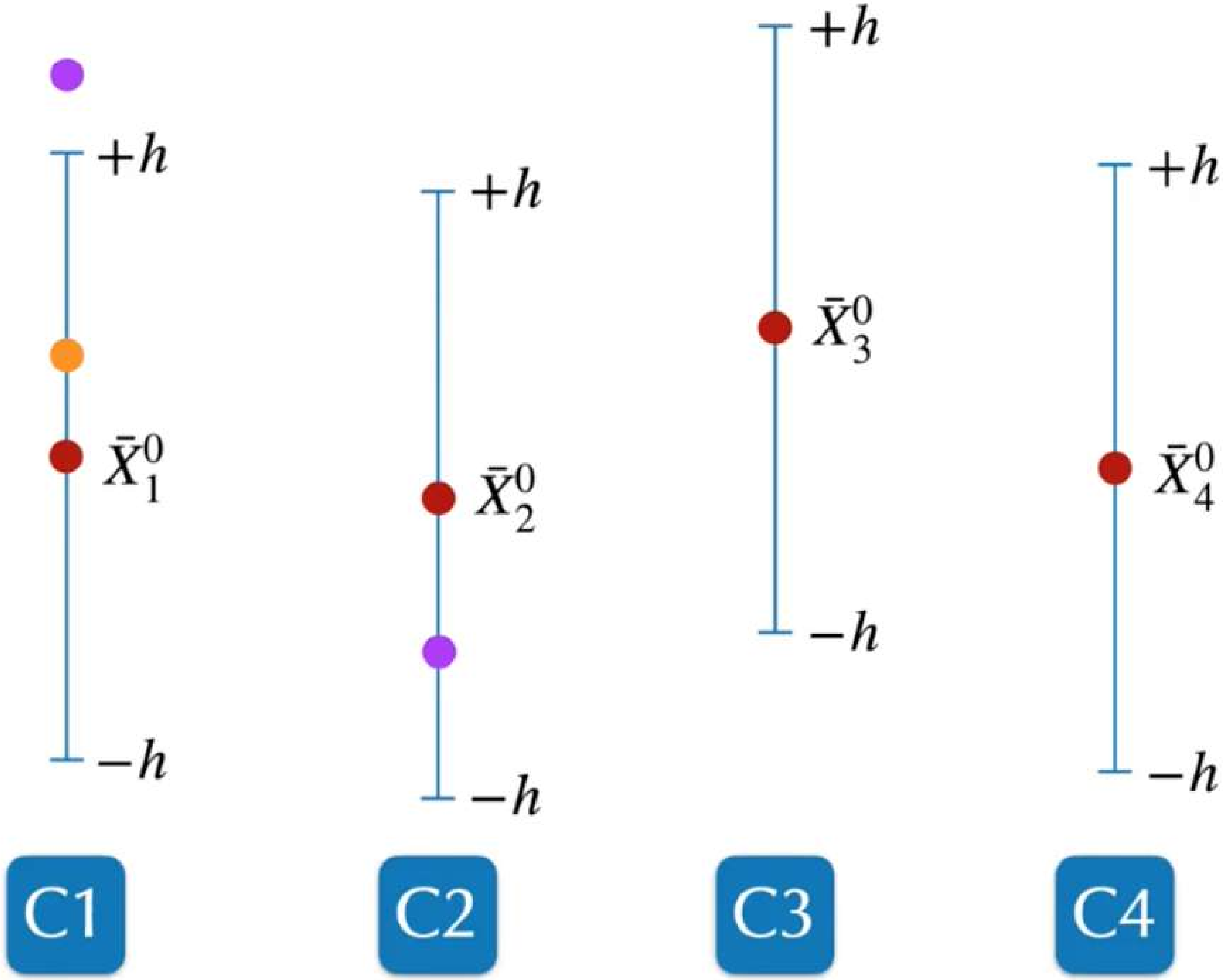
Cumulative negative deviation from the reference point at t $g_a^-(t) = \max\left\{0, g_a^-(t-1) + s_a^-(t)\right\}$

We have a change if $g_a^-(t) > h$ or $g_a^+(t) > h$

Example



Example



CD-UCB for Online Matching Pseudocode

1. inltialize τ_a=0 for each arm $a \in A$
2. for each t

$$\boldsymbol{a}_t \leftarrow \arg\max_{\boldsymbol{a} \in \mathcal{M}} \left\{ \sum_{a \in \boldsymbol{a}} \bar{x}_{a,\tau_a,t} + \sqrt{\frac{2\log(n(t))}{n_a(\tau_a, t-1)}} \right\} \text{ with probability } 1-\alpha$$

$\boldsymbol{a}_t \leftarrow$ random **match** with probability α

$n(t)$	is the total number of valid samples
$\bar{x}_{a,\tau_a,t}$	is the empirical mean of arm a over the last valid samples
$n_a(\tau_a, t-1)$	is the number of valid samples for arm a

3. collect reward r_t
4. if $CD_a(r_\tau, \ldots, r_t) = 1$ then $\tau_a = t$ and restart CD_a