Online multi-server convex chasing and optimization

Sébastien Bubeck - MSR AI
Yuval Rabani - HUJI
Mark Sellke - Stanford U.

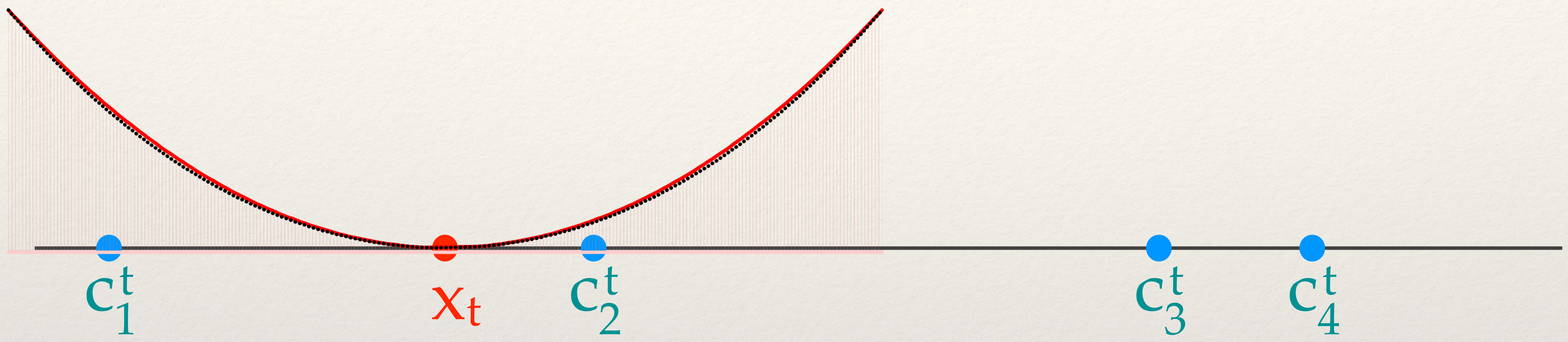
k-means clustering

- (More generally: center-based clustering)
- Input: a finite set $X \subset \mathbb{R}^d$
- Output: a sketch $\{c_1, c_2, \ldots, c_k\}$ of X $\quad (\forall i, c_i \in \mathbb{R}^d)$
- Objective: minimize $\sum_{x \in X} \min_i \|x-c_i\|_2^2$

Online clustering

- The points in X are revealed gradually: $x_1, x_2, x_3, \ldots$
- The cluster centers may shift over time
- $C^t = \{c_1^t, c_2^t, \ldots, c_k^t\}$ predicts x_t
- Measure both quality and stability of prediction:
  - Quality: $\| \square \|_2^2$ distance to closest center
  - Stability: $\| \square \|_2$ measure of center(s) movement

k-chasing convex functions

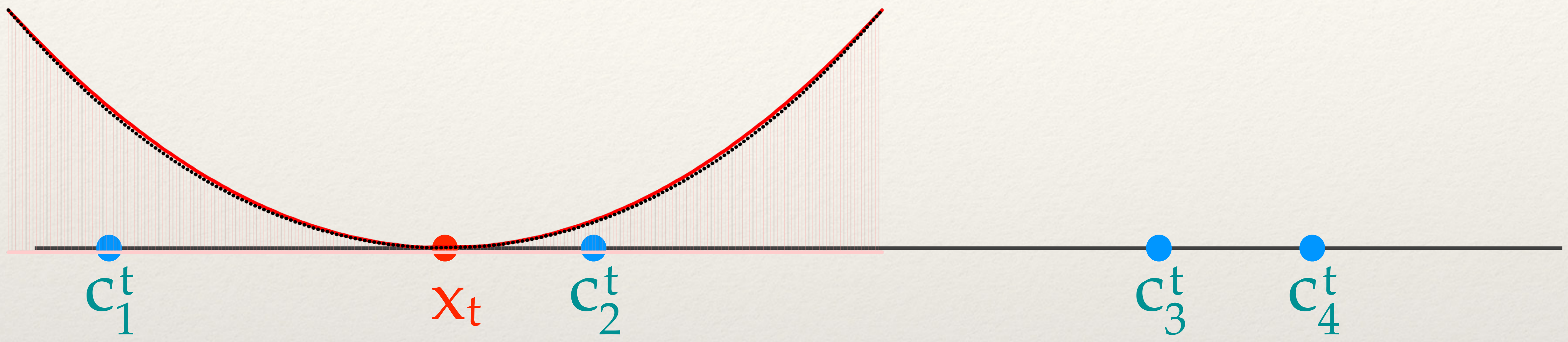


- Every input point x_t defines a convex function

$$f_t(x) = \|x - x_t\|_2^2$$

- We pay centers movement cost + $\min_j f_t(c_j^t)$
- k-server + chasing convex bodies/functions

k-chasing convex functions



- Every input point x_t defines a convex function

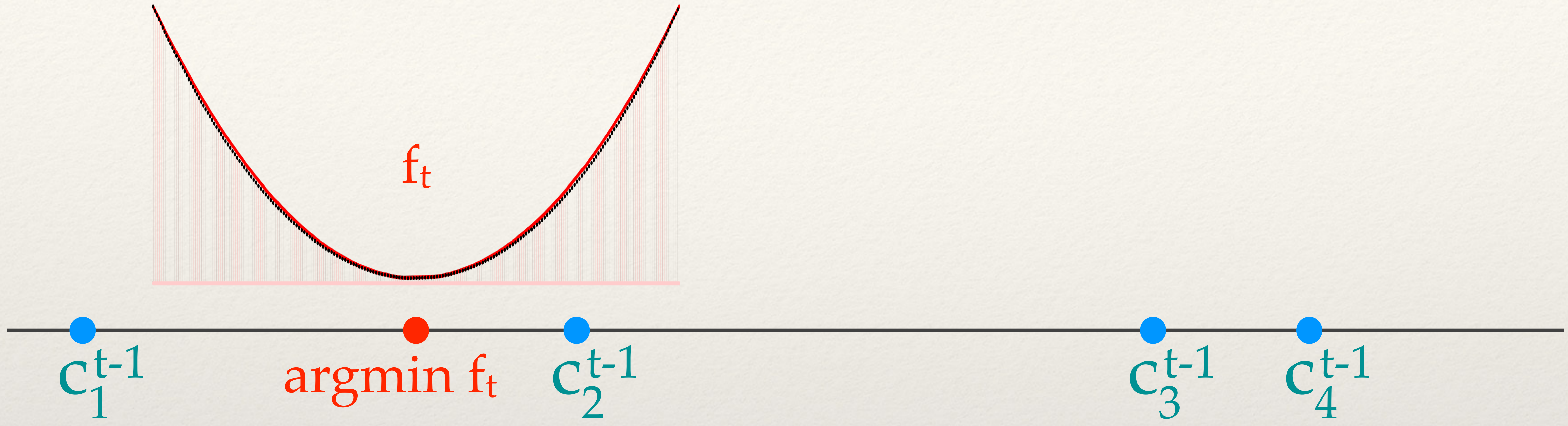
$$f_t(x) = \|x - x_t\|_2^2$$

- We pay centers movement cost + $\min_j f_t(c_j^t)$ (service cost)
- k-server + chasing convex bodies / functions

k-server + chasing

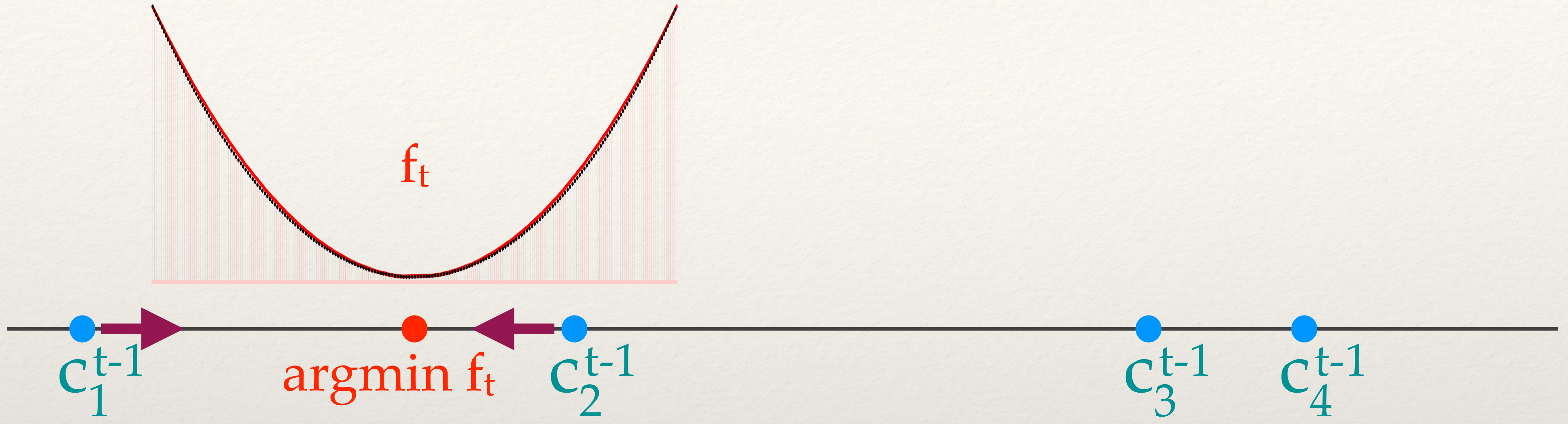
- k-server [MMS, …, BCLLM, Lee]:
  - a request at x_t: $f_t(x_t) = 0$, o.w. $f_t(x) = +\infty$
  - det. l.b. k, u.b. $2k-1$
  - rand. l.b. $\Omega(\frac{\log k}{\log\log k})$, u.b. polylog(k,n)
- chasing convex bodies (1 server) [FL, …, S, AGGT]:
  - requesting B_t: $\forall x \in B_t$, $f_t(x) = 0$, o.w. $f_t(x) = +\infty$
  - l.b. $\sqrt{d}$, u.b. $\min\{d, \sqrt{d \log n}\}$

k-chasing convex functions on the line



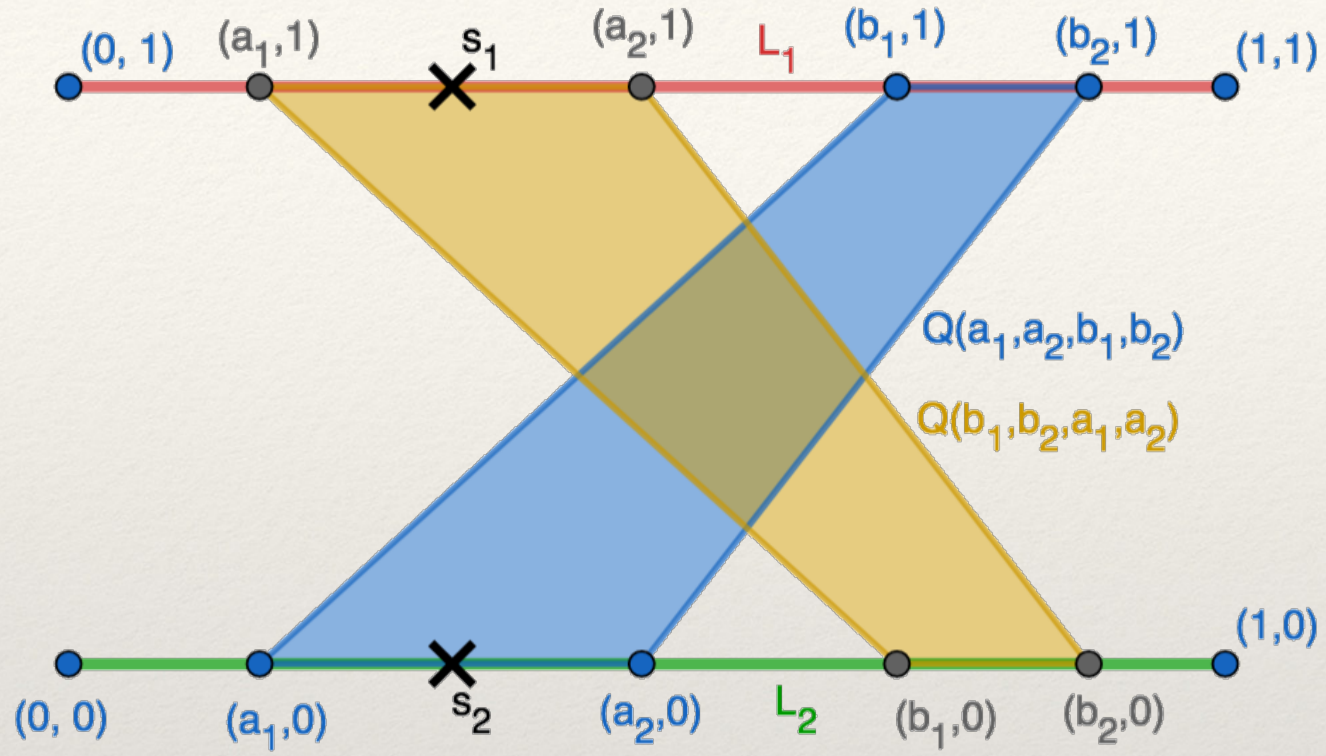
- Double coverage: move adjacent servers towards argmin f_t, until movement = service cost
- competitive ratio: 4k (l.b. k $\Leftarrow$ k-server)

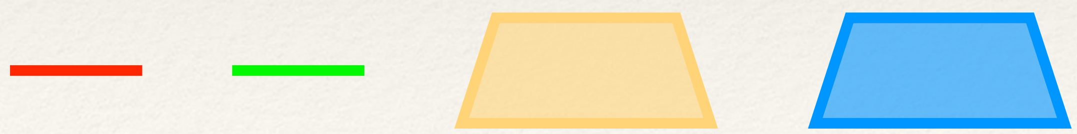
k-chasing convex functions on the line



- Double coverage: move adjacent servers towards argmin f_t, until movement = service cost
- competitive ratio: 4k (l.b. k ⇐ k-server)

d, k = 2



- On the line: chasing unions of 2 segments — unbounded competitiveness
- Emulate in $\mathbb{R}^2$ using convex sets — repeat requesting 

Online k-means

- $\tilde{f}_t(x) = \| c_{min}^{t-1} - x_t \|_2^2 \cdot \frac{\| x - x_t \|_2}{\| c_{min}^{t-1} - x_t \|_2}$
- Use $\tilde{f}_t(\square)$ instead of $\| \square - x_t \|_2^2$
- reduces k-means to k-median: k-chasing functions of the form $\alpha_t \cdot \| \square - x_t \|_2$ with $\alpha_t > 0$
- Move to minimum (MTM) algorithm: either move a center to argmin f_t or stay put.

Online k-means

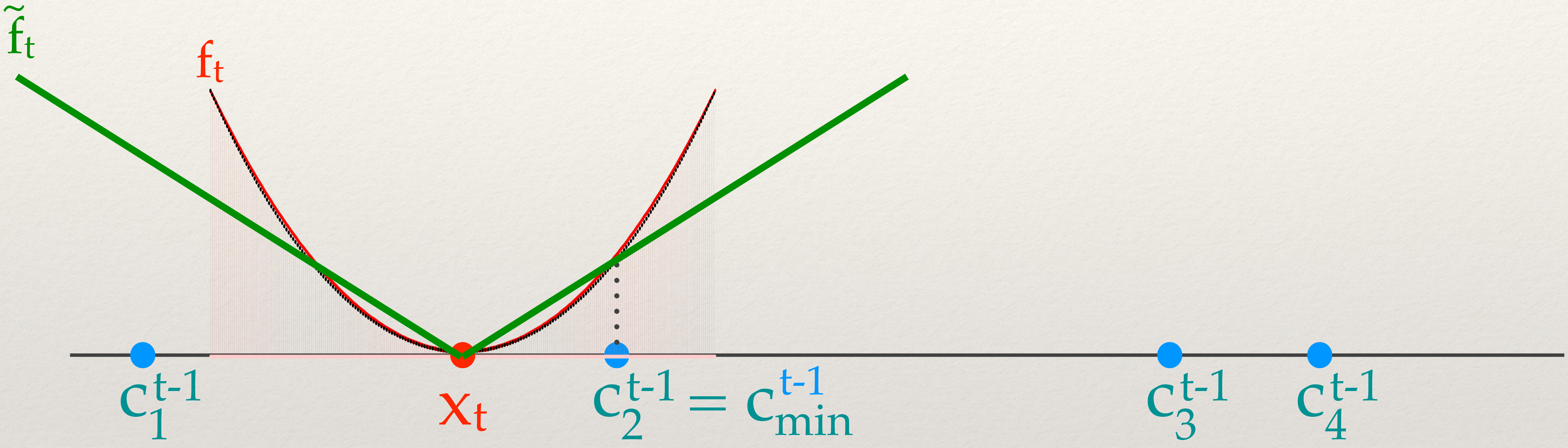
closest center to x_t

- $\tilde{f}_t(x) = \| c_{min}^{t-1} - x_t \|_2^2 \cdot \frac{\| x - x_t \|_2}{\| c_{min}^{t-1} - x_t \|_2}$
- Use $\tilde{f}_t(\square)$ instead of $\| \square - x_t \|_2^2$
- reduces k-means to k-median: k-chasing functions of the form $\alpha_t \cdot \| \square - x_t \|_2$ with $\alpha_t > 0$
- Move to minimum (MTM) algorithm: either move a center to argmin f_t or stay put.

Online k-means

- $\tilde{f}_t(x) = \| c^{t-1}_{min} - x_t \|_2^2 \cdot \frac{\| x - x_t \|_2}{\| c^{t-1}_{min} - x_t \|_2}$
- Use $\tilde{f}_t(\square)$ instead of $\| \square - x_t \|_2^2$
- reduces k-means to k-median: k-chasing functions of the form $\alpha_t \cdot \| \square - x_t \|_2$ with $\alpha_t > 0$
- Move to minimum (MTM) algorithm: either move a center to argmin f_t or stay put.

Illustration



A sequence of reductions

- Online k-means
- Repeat M_t times $f_t = \| \square - x_t \|_2^2 \ / \ M_t$
- Blind MTM k-chasing of convex functions a la f_t
- Blind MTM online k-median with $\alpha_t < 1$ (use $\tilde{f}_t$)
- k-server (pass request with prob. α_t)

A sequence of reductions

choose M_t s.t. $\forall j,\ f_t(c_j^{t-1}) < \| c_j^{t-1} - x_t \|_2$

- Online k-means
- Repeat M_t times $f_t = \| \square - x_t \|_2^2 \,/\, M_t$
- Blind MTM k-chasing of convex functions a la f_t
- Blind MTM online k-median with $\alpha_t < 1$ (use $\tilde{f}_t$)
- k-server (pass request with prob. α_t)

A sequence of reductions

- Online k-means
- Repeat M_t times $f_t = \| \square - x_t \|_2^2 \ / \ M_t$
- Blind MTM k-chasing of convex functions a la f_t
- Blind MTM online k-median with $\alpha_t < 1$ (use $\tilde{f}_t$)
- k-server (pass request with prob. α_t)

A sequence of reductions

Blind: service cost paid before move

- Online k-means
- Repeat M_t times $f_t = \| \square - x_t \|_2^2 \; / \; M_t$
- Blind MTM k-chasing of convex functions a la f_t
- Blind MTM online k-median with $\alpha_t < 1$ (use $\tilde{f}_t$)
- k-server (pass request with prob. α_t)

A sequence of reductions

- Online k-means
- Repeat M_t times $f_t = \| \square - x_t \|_2^2 \; / \; M_t$
- Blind MTM k-chasing of convex functions a la f_t
- Blind MTM online k-median with $\alpha_t < 1$ (use $\tilde{f}_t$)
- k-server (pass request with prob. α_t)

A sequence of reductions

- Online k-means
- Repeat M_t times $f_t = \| \square - x_t \|_2^2 \;/\; M_t$
- Blind MTM k-chasing of convex functions a la f_t
- Blind MTM online k-median with $\alpha_t < 1$ (use $\tilde{f}_t$)
- k-server (pass request with prob. α_t)

squares the competitive ratio, so $O(k^2)$

More results

- The reduction applies to a class of *well-sharpened* functions.
- $O(k)$-competitive online k-median algorithm.
- Regret minimization: 1-Lipschitz convex functions on the unit ball, no movement cost —
  - T-step regret $O(\sqrt{kdT \log T})$ (inefficient alg.)
  - With linear loss functions $O(\sqrt{kT \log k \log^3 T})$