A framework for Multi-A(rmed)/B(andit) testing with online FDR control

Fanny Yang, UC Berkeley

Joint work with Aaditya Ramdas, Kevin Jamieson and Martin Wainwright

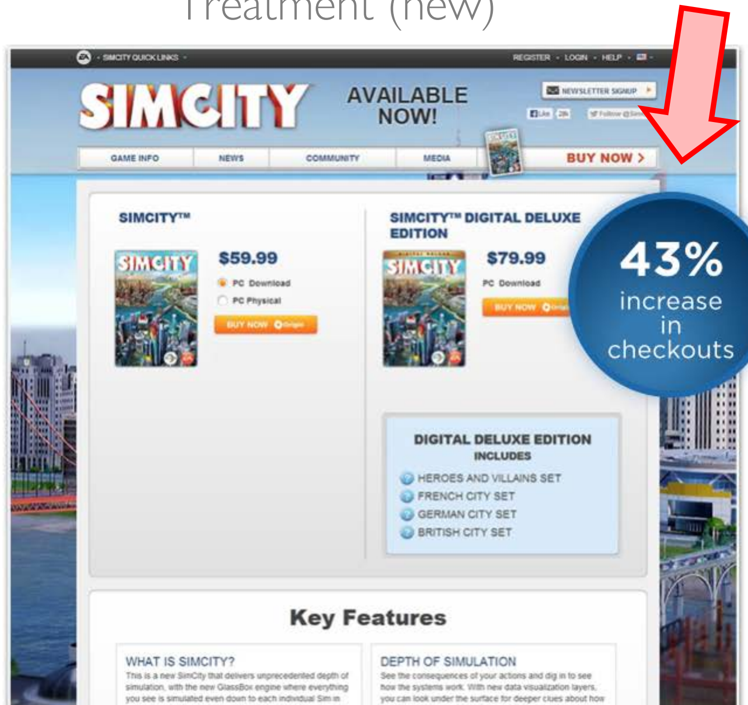
MCP Conference, June 2017

A company has an on-going sale, and it's going alright…

…a new marketing student suggests different layout → much better!

Control (default)

Treatment (new)



Time

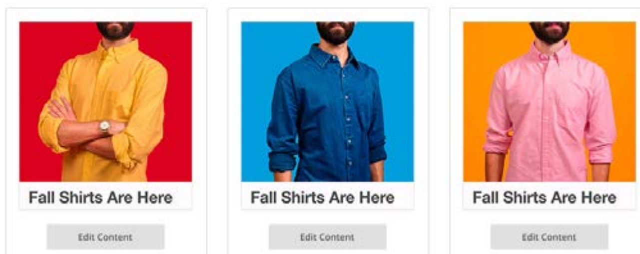
Jan

. . .

April

. . .

August



. . .

.
.
.

What and how to optimize the choices?

Outline

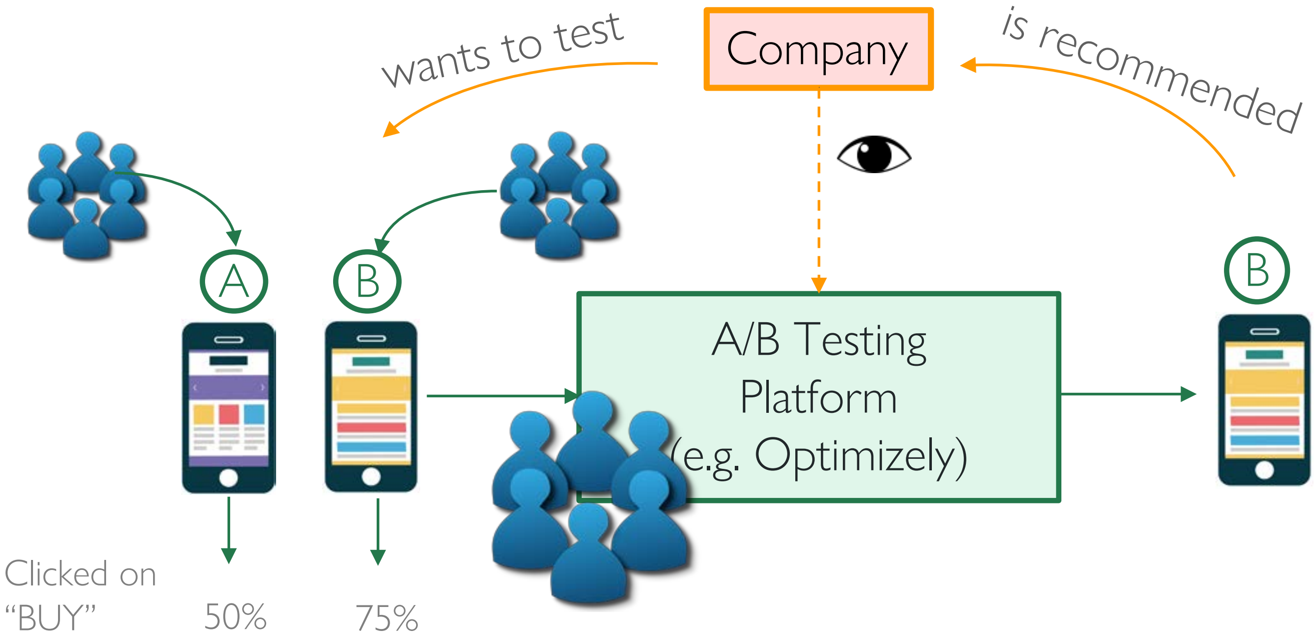
- Status quo (A/B testing)
- Unadded practical requirements
- Review of known sequential procedures

    Multi-Armed Bandit algorithms

    Online FDR procedures

- Combining all three frameworks and guarantees

Role of A/B Testing platform



Only recommends new version if evidence is "significant"!
Allows continuous monitoring given same number of samples

A/B Testing model

- Distributions P_A, P_B for A (control), B with means μ_{control}, μ_B

  H_0: $\mu_{\text{control}} > \mu_B$ vs. H_1: $\mu_{\text{control}} < \mu_B$

- Data are i.i.d. samples from P_A, P_B
- Compute test statistic T' given data
- Compute probability of T being more extreme: $p = P_{H0}(T > T')$
- Recommend new arm (reject null)
  if p-value $p <$ desired significance α (often 0.05)

A/B testing has been done…

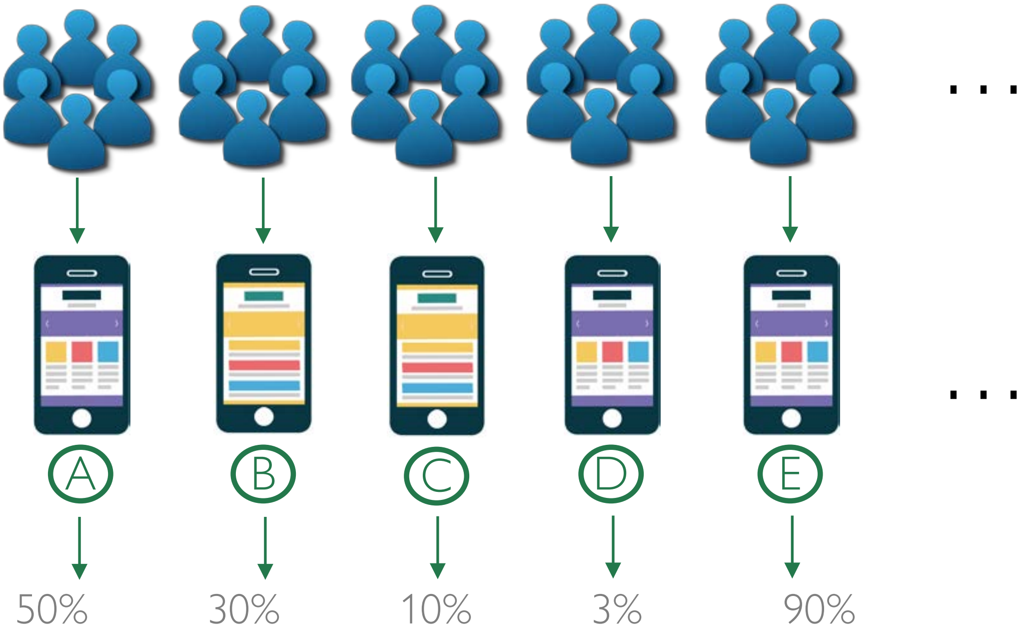
(Commercial Platform: e.g. Optimizely, papers by Johari et al.)

Practical desiderata still unaddressed

- Many arms, limited budget
- Many tests throughout the year

Still be able to continuously monitor

Desideratum 1: many variants (arms)



Problem: Uniform sampling scales linearly with # arms

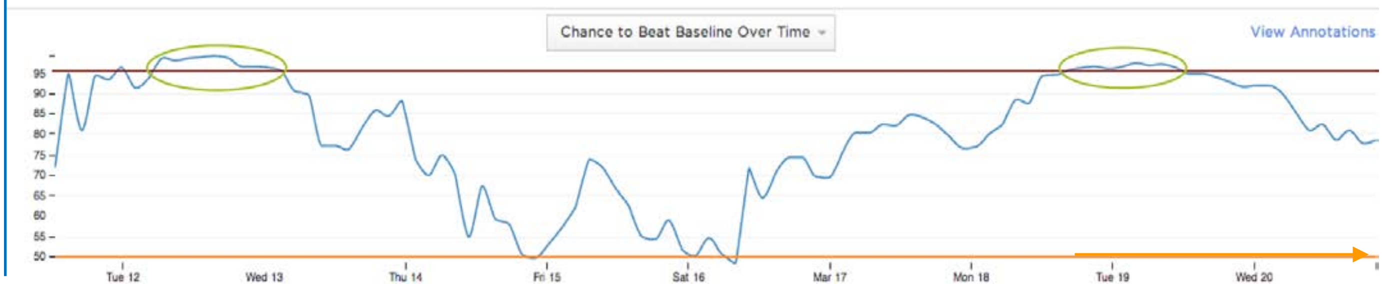
Approach: sample adaptively to need less traffic/time

Desideratum 2: p-value peeking

Confidence

$1\text{-}p_t$

Valid p-value (for each t) satisfies: $P_{H0}(p_t < \alpha) = \alpha$



Source: https://blog.optimizely.com

Number of users (samples)

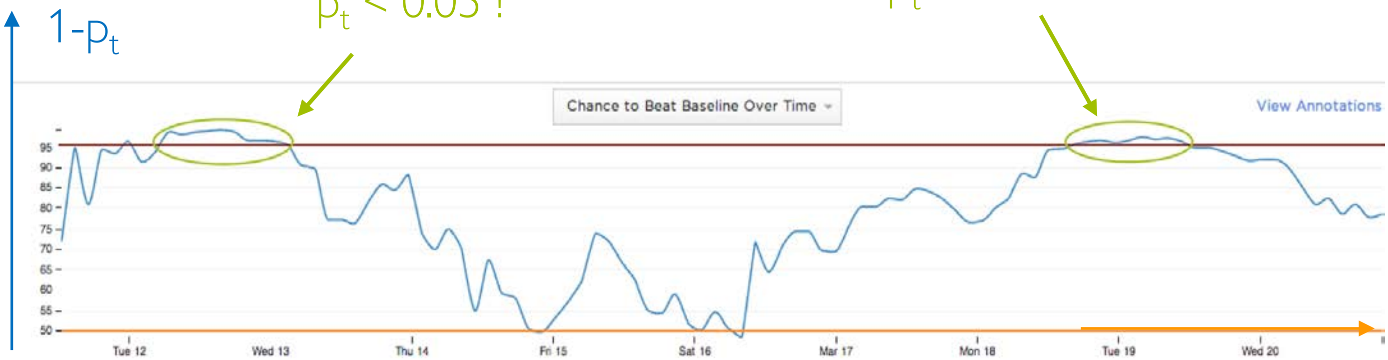
Desideratum 2: p-value peeking

Confidence

$1\text{-}p_t$

$p_t < 0.05$!

$p_t < 0.05$!



Source: https://blog.optimizely.com

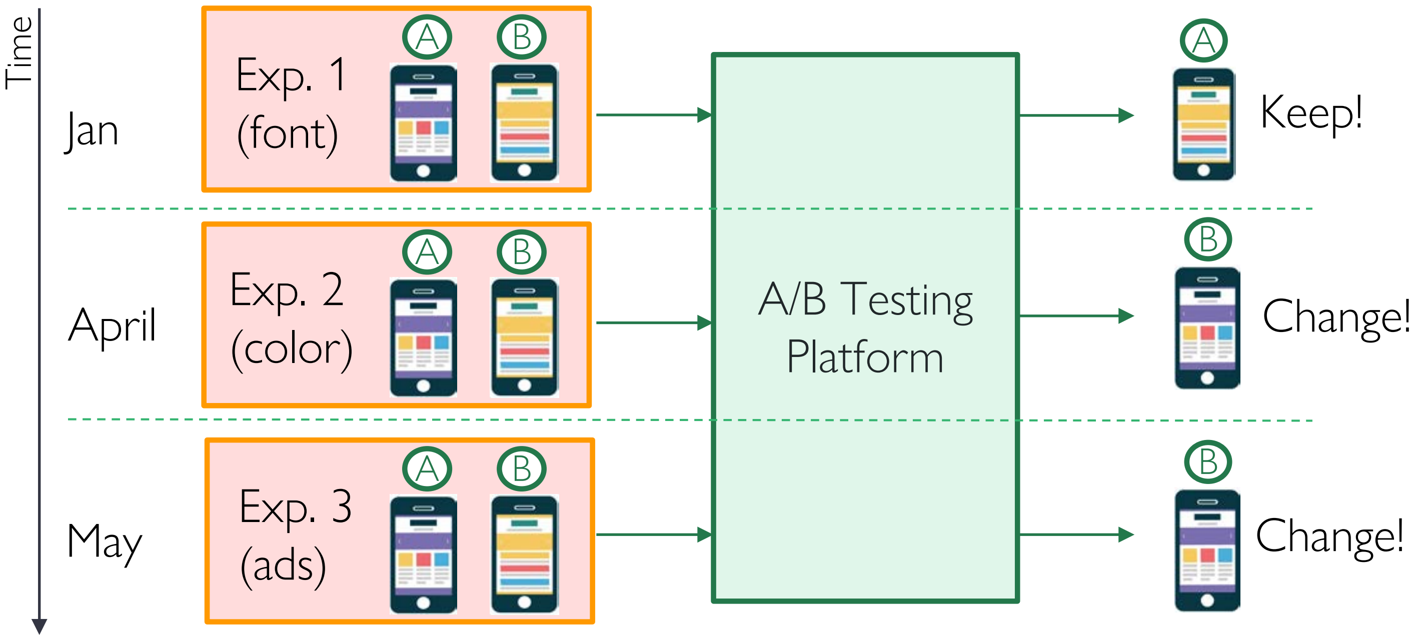
Number of users (samples)

… stop and reject whenever $p_t < 0.05$ …

Problem: $\mathrm{P}(p_t < 0.05$ for some t$) \gg \mathrm{P}(p_t < 0.05) = 0.05$

Approach: construct p-values that are valid when repeatedly queried for non-uniform sampling

Desideratum 3: many tests over time



What kind of error to control and how?

Desideratum 3: many tests over time

Company's interests

- implementing change system-wide has base cost
  → wants to be "sure", i.e. not too many wrong rejections
- detect better treatments if they give higher revenue (high power)
  → FWER & Bonferroni not great 😒

Compromise → Control the expected ratio $\frac{\text{\# false rejections}}{\text{\# rejections}}$ (FDR)

Problem: Neither Vanilla testing nor batch FDR at α is sufficient

Approach: online control of FDR at level α

Summary of desiderata

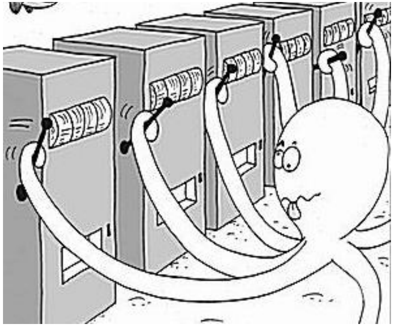
- Many variants (arms)
  → Adaptive Sampling (Best-arm Multi-Armed Bandit)
- Allow p-value peeking
  → Construct always valid p-values
- Many tests over time
  → Online FDR control procedures

➡ Doubly sequential procedure!

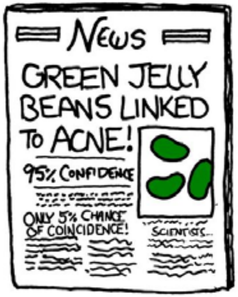
Combining two known adaptive procedures while preserving sample efficiency and statistical guarantees

Reviewing known procedures…

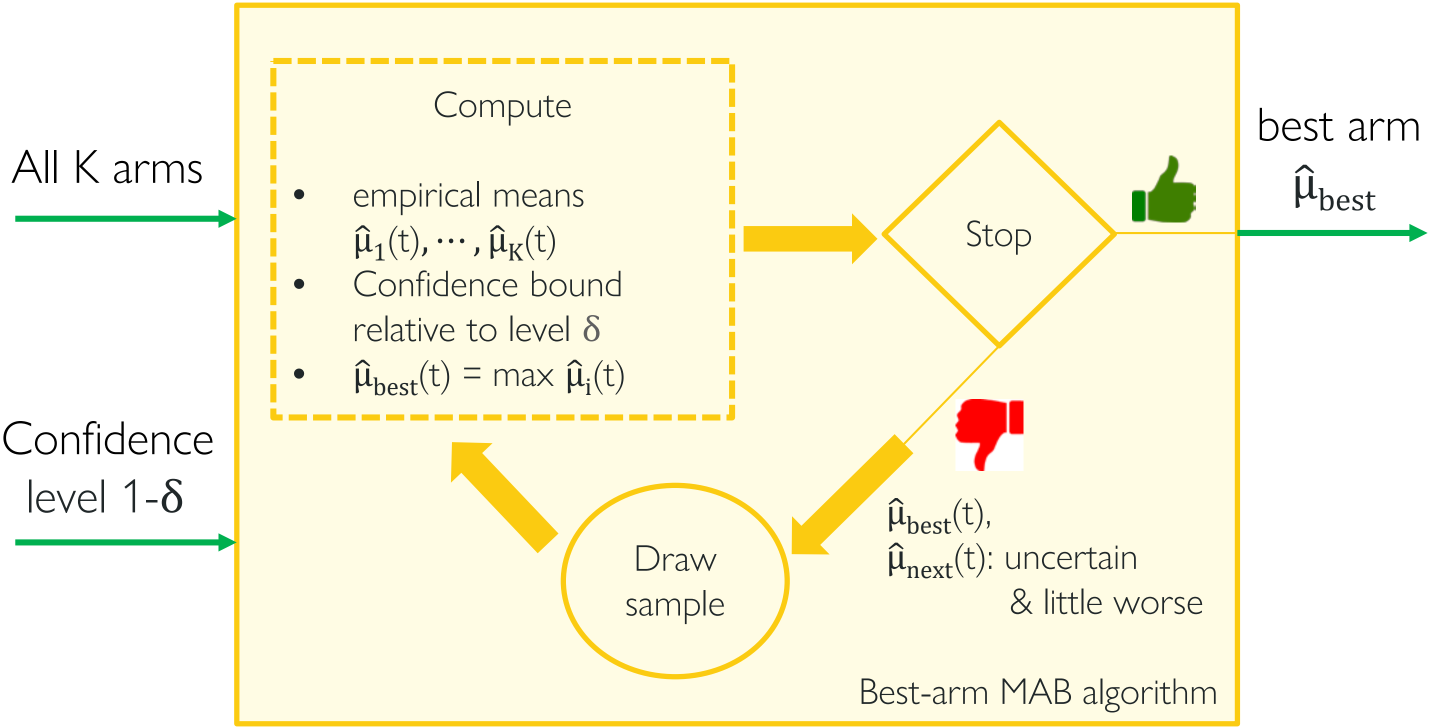
Multi Armed-Bandits



Online FDR procedures



Recap: Best-arm MAB schematic



*Stop criterion: $\hat{\mu}_{best}$ certain, all other $\hat{\mu}_i$ either certain & little worse, uncertain & much worse

Recap: Best-arm MAB guarantees

Known results for K arms, confidence parameter δ,
i.e. P(MAB finds best arm) $\geq 1 - \delta$

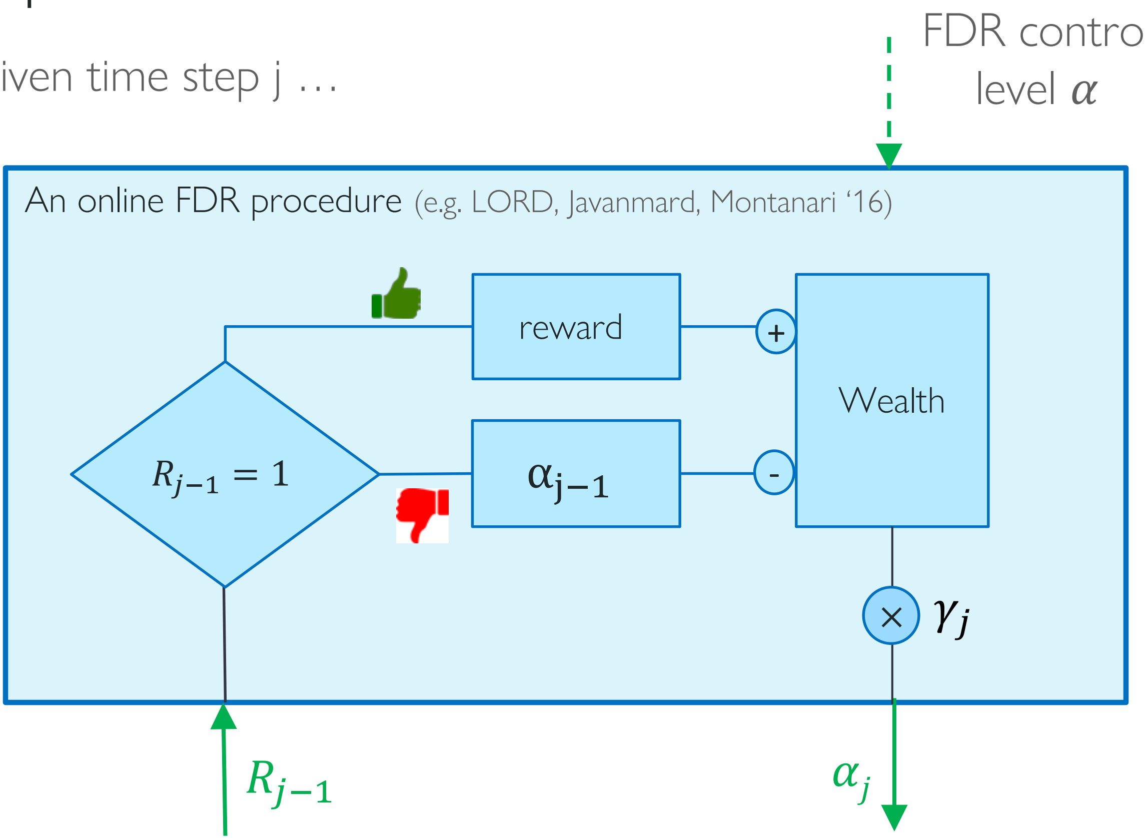
- Sample complexity guarantees e.g. for the LUCB algorithm (Kalyanakrishnan '14) depending on gaps $\Delta_i = \mu_{\mathrm{best}} - \mu_{\mathrm{i}}$

  $\sum_{\mathbf{i \neq best}} \Delta_{\mathrm{i}}^{-2} \log(1/\delta)$ (LUCB) vs. $K \cdot \max_{\mathbf{i \neq best}} \Delta_{\mathrm{i}}^{-2} \log(1/\delta)$ (Uniform sampling)

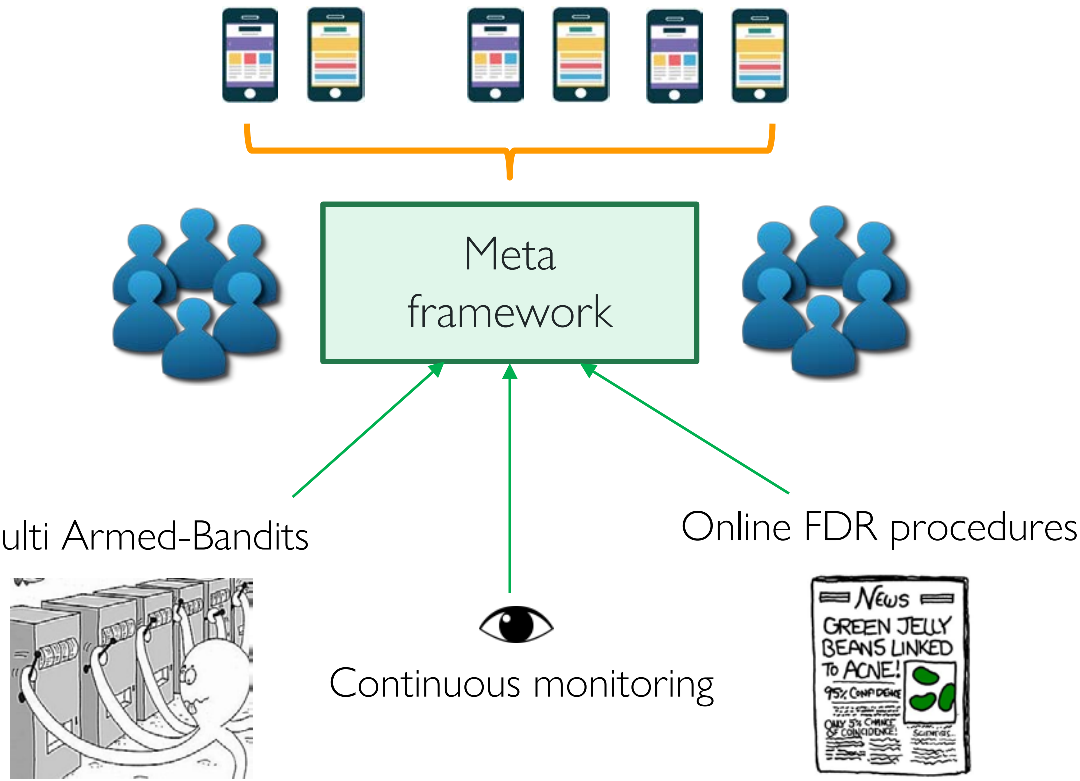
- Matches lower bounds (Garivier et al. '16, Simchowitz et al. '17)

Recap: Online FDR control

At a given time step j …



Open: How to combine everything



Conceptual and theoretical challenges

For embedding an MAB algorithm in a testing setup

- What is the right **null hypothesis**?
  How to incorporate asymmetry in algorithm?
- How to get **always valid** p–values for **non-uniform** and **dependent** samples?

For using MAB in online FDR framework

- What **interaction** between MAB and FDR **preserves**
  best of both worlds (FDR control, low sample complexity)?

Our contribution 1: Embedding MAB

Null hypothesis:

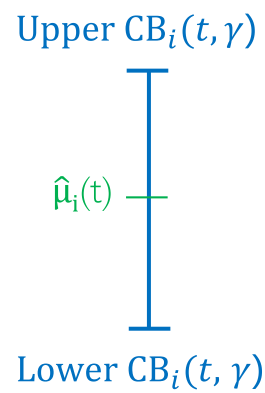
$$H_0: \mu_{\text{control}} > \mu_i - \varepsilon \quad \forall\, i=1\cdots K \qquad \text{vs.} \qquad H_1: \mu_{\text{control}} + \varepsilon < \mu_i \quad \exists\, i$$

Prop. Modified MAB finds ε –better arm with confidence

Always valid p-value p_t , i.e. $P(p_t < \alpha \text{ for some } t) \leq \alpha$:

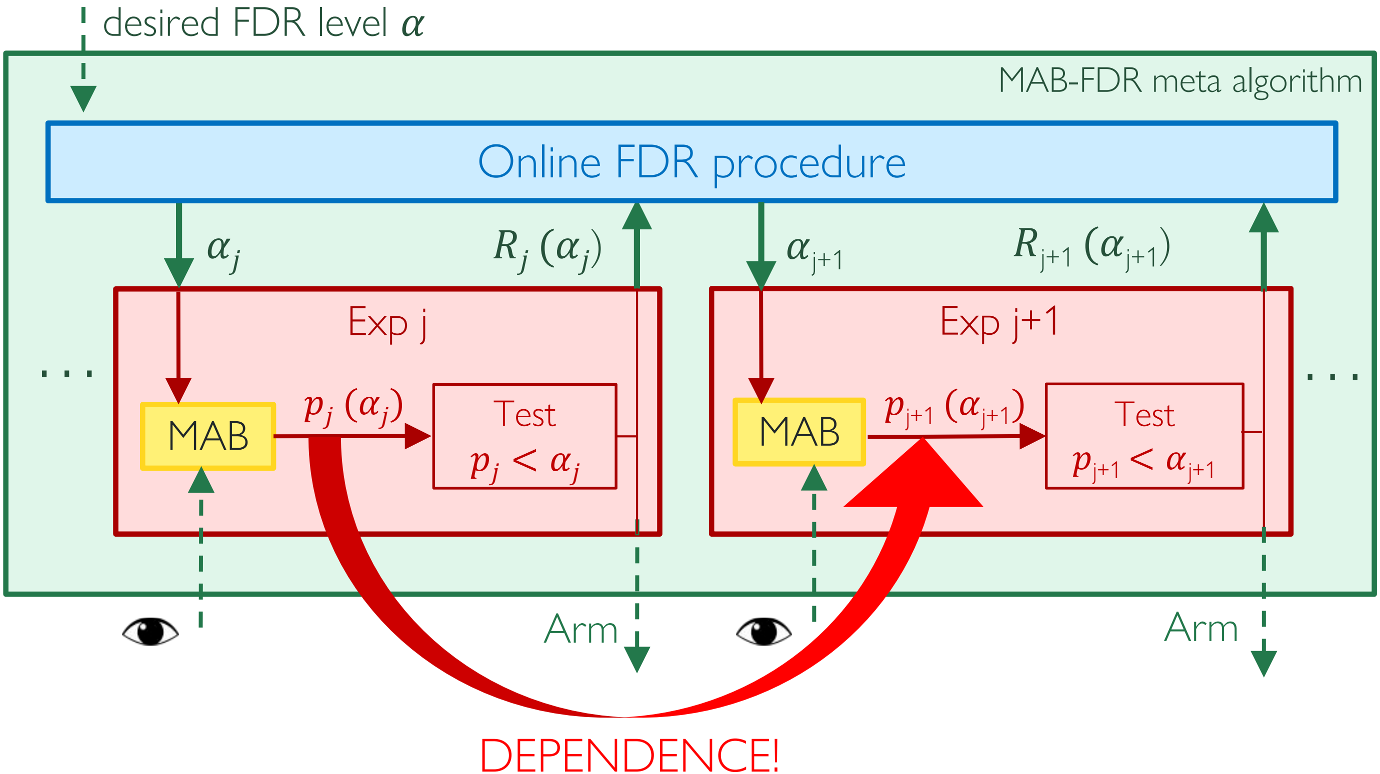
- Law of Iterated Logarithm (LIL)⟶ Always validity
  $P(\exists\, t : \mu_i \in [\mathrm{LCB}_i(t,\gamma), \mathrm{UCB}_i(t,\gamma)]) \leq \gamma$

- For each arm compute
  $P_{i,t} = \sup\{\gamma \in [0,1]: \mathrm{LCB}_i(t,\gamma) < \mathrm{UCB}_0(t,\gamma)\}$
  Final p-value: $p_t = \min_{i=1,\cdots,K} P_{i,t}$

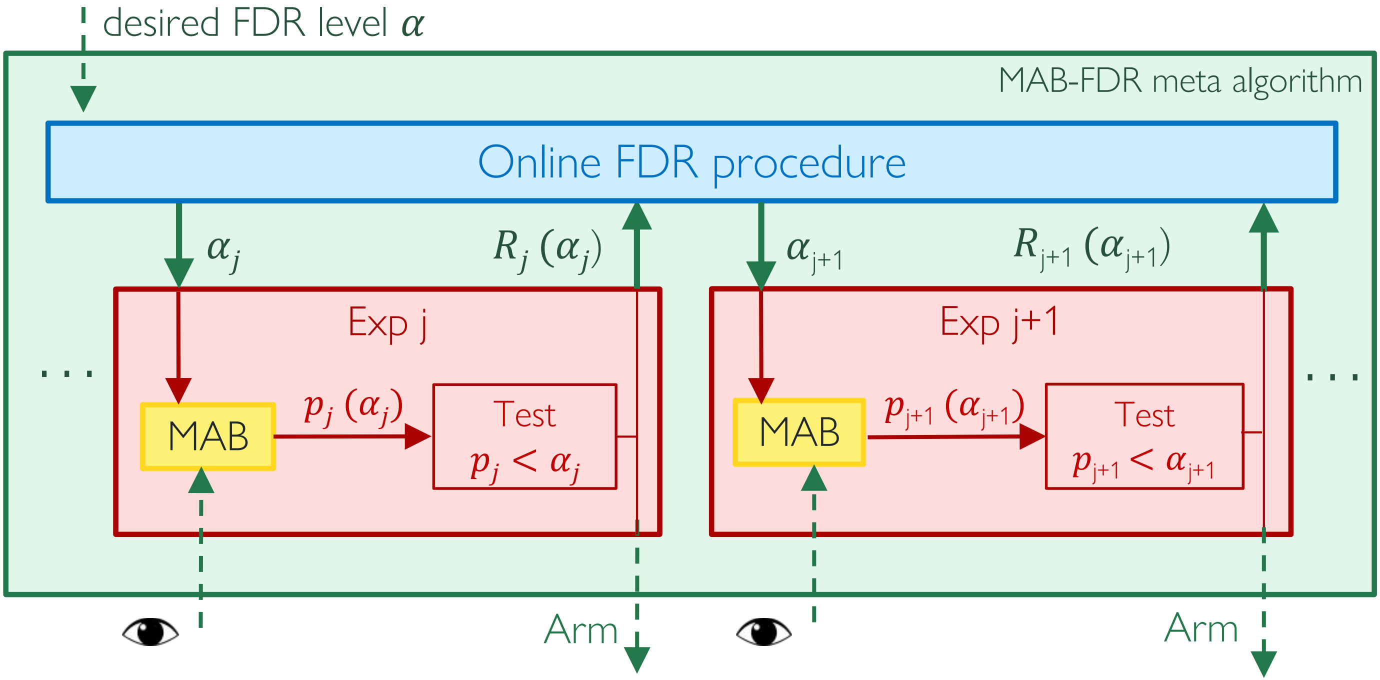


Prop. Can compute always valid p-values by sampling with MAB!

Contribution 2: MAB-FDR and guarantees



Contribution 2: MAB-FDR and guarantees



Theorem "FDR" is at most α at any time.

If the algorithm is not terminated early, then "power" is at least (1- α).

Summary

Introduced a doubly sequential procedure, which simultaneously

- Yields good sample complexity
- Allows continuous monitoring
- Controls FDR in an online fashion

Fanny
Yang

Aaditya
Ramdas

Kevin
Jamieson

Martin
Wainwright

"A framework for Multi-A(rmed)/B(andit) testing with online FDR control"
Available as arXiv preprint arXiv:1706.05378 (2017)