Information Filtering for arXiv.org

Bandits,
Exploration vs. Exploitation,
& the Cold Start Problem

Peter Frazier
School of Operations Research & Information Engineering
Cornell University

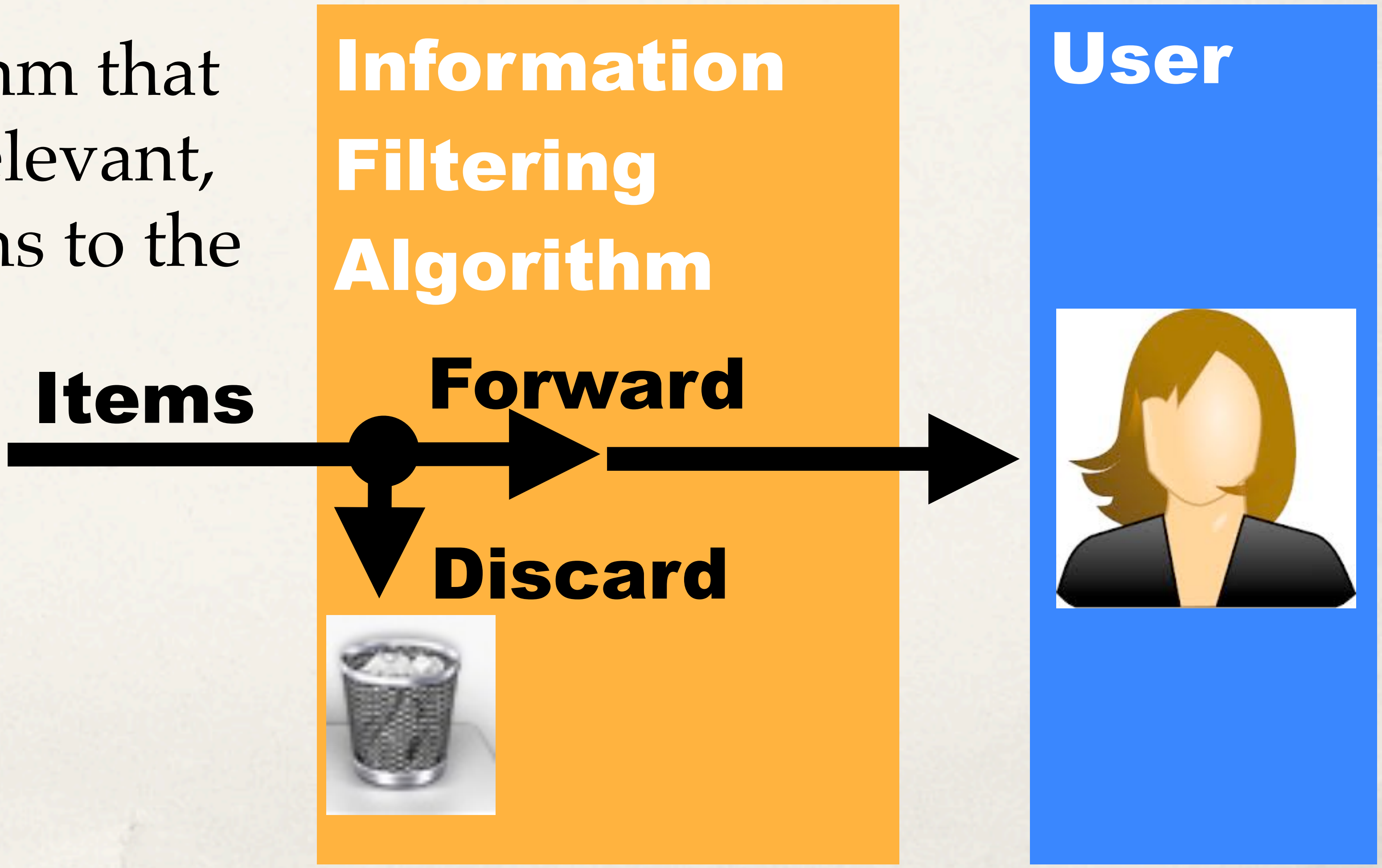
with Xiaoting Zhao, PhD student, School of ORIE, Cornell
Support provided by the National Science Foundation, award IIS-142251, IIS-1247696

Yelp, Monday December 16, 2013

We are interested in

information filtering

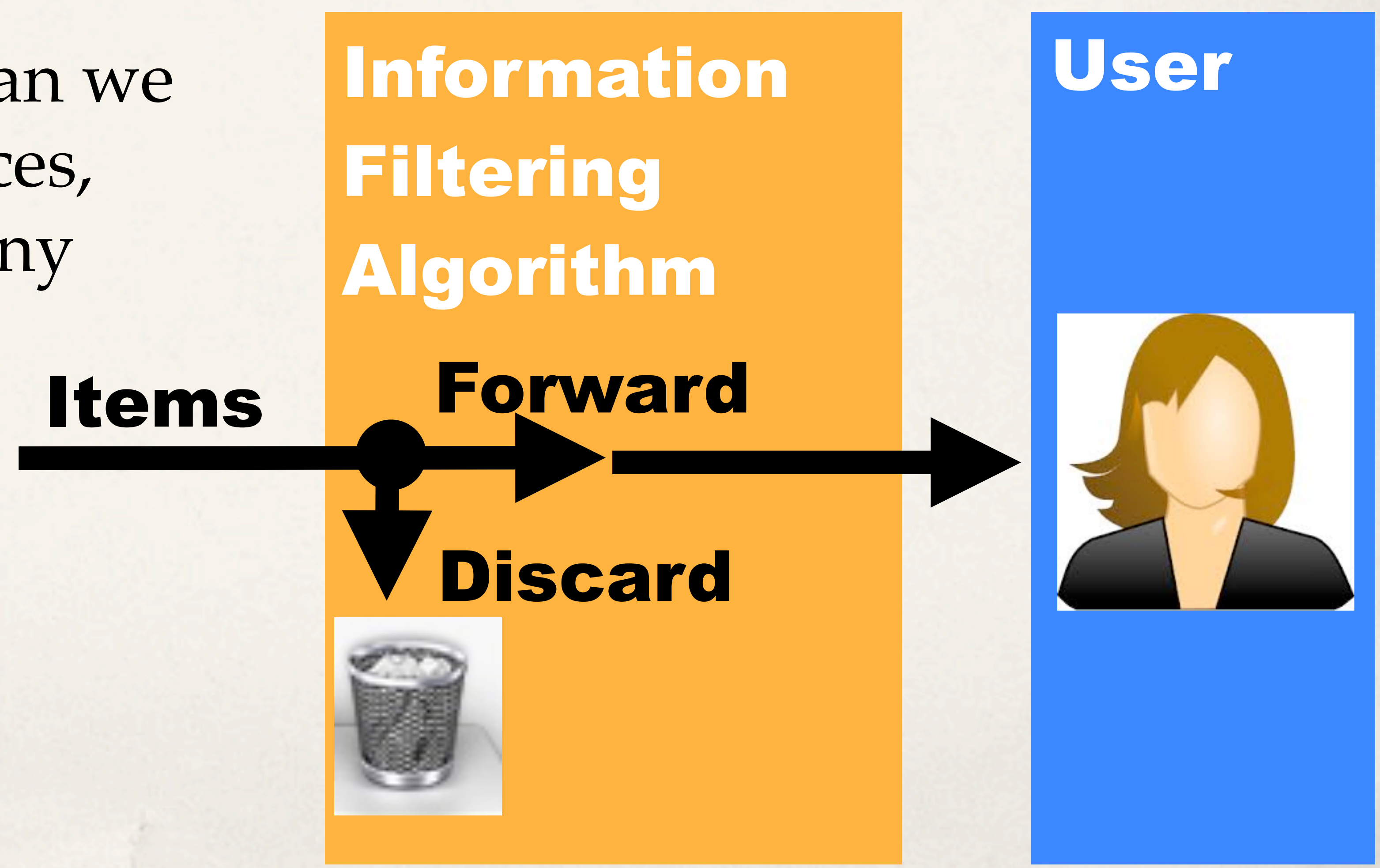
- We face a sequence of time-sensitive items (emails, blog posts, news articles).
- A human is interested in some of these items.
- But, the stream is too voluminous for her to look at all of them.
- **Our goal**: design an algorithm that can learn which items are relevant, and forward only these items to the user.



We are interested in

information filtering

- If we had lots of historical data, we could train a machine learning classifier to predict which items would be relevant to this user.
- But what if we are doing information filtering for a new user?
- **Research Question:** How can we quickly learn user preferences, without forwarding too many irrelevant items?

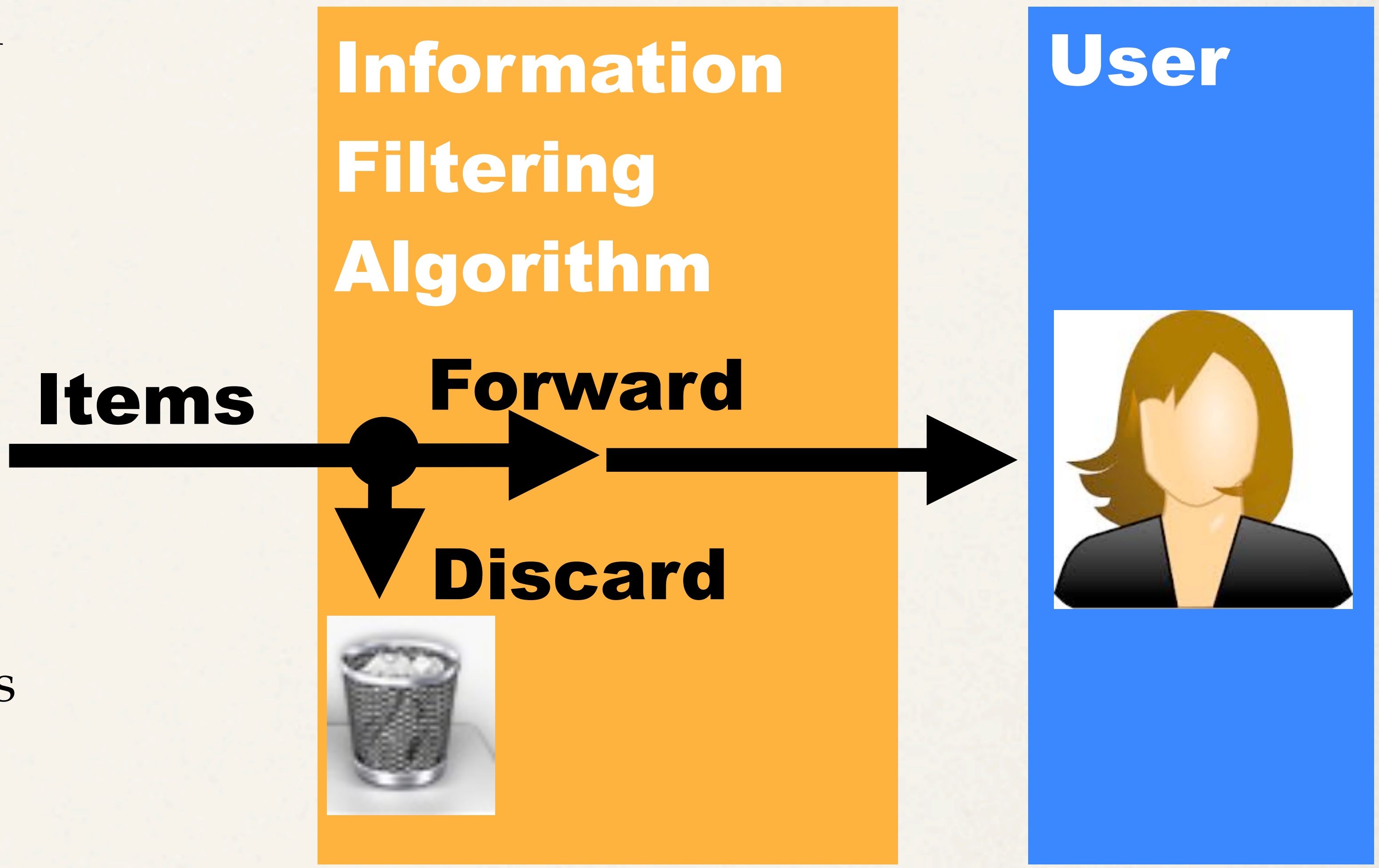


We are interested in

exploration vs. exploitation

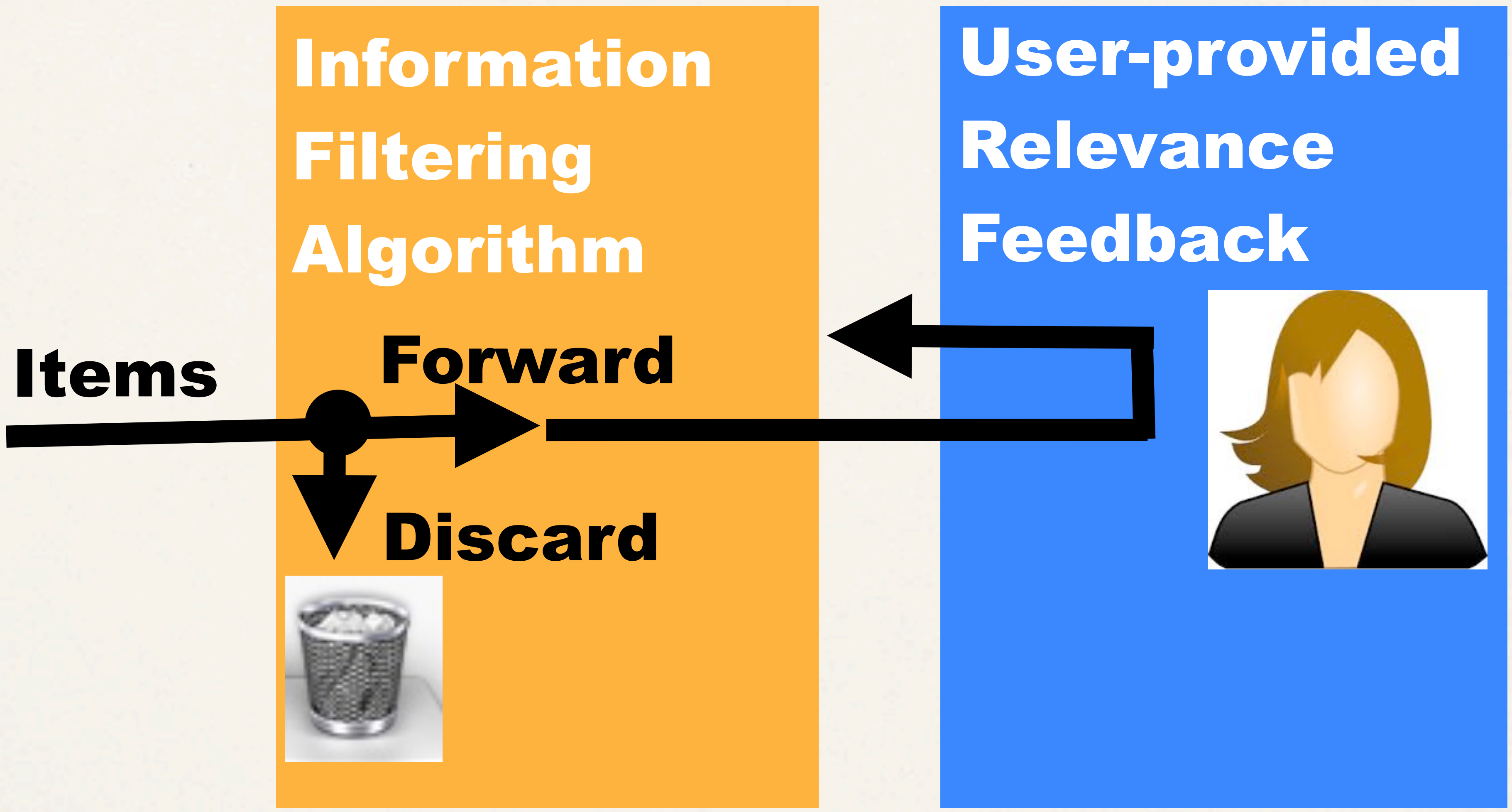
in information filtering

- More generally, suppose there is an item type with little historical data from this user.
- This can arise because:
  - this is a new user;
  - the item mix is changing;
  - the information filtering alg. has not forwarded items of this type.



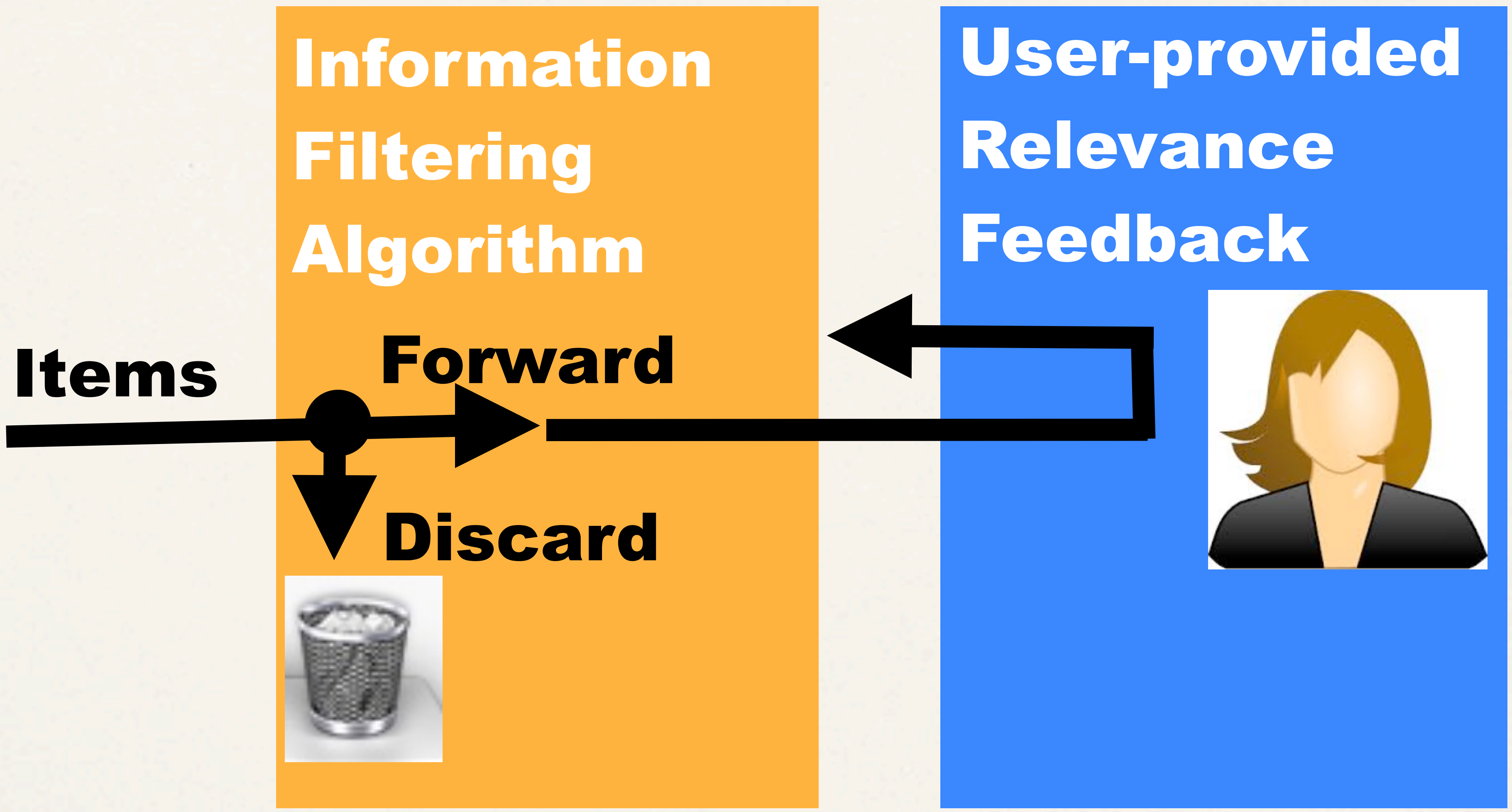
- We may **EXPLORE**, i.e., forward a few items of this type, to better learn this type's relevance.
- But, we may want to **EXPLOIT** what little training data we have, which may suggest this item type is irrelevant.
- What should we do?

We develop an information filtering algorithm that trades exploration vs. exploitation



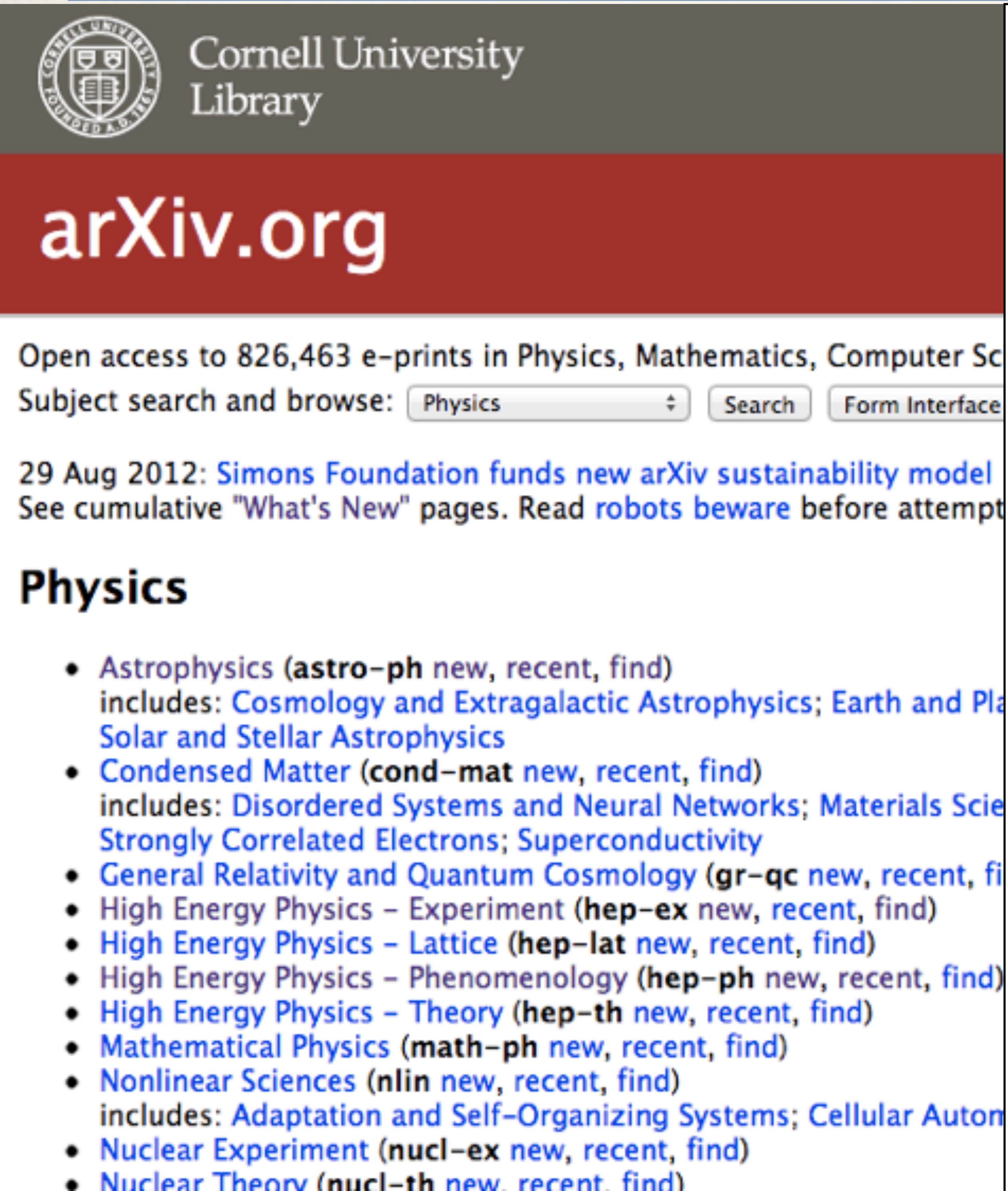
- We use an **optimal learning** approach, which relies on **Bayesian statistics** and **dynamic programming**.

We develop an information filtering algorithm that trades exploration vs. exploitation



- We focus on the **value of the information** in the user's relevance feedback.

We are motivated by an information filtering system we are building for arxiv.org


Cornell University Library

arXiv.org

Open access to 826,463 e-prints in Physics, Mathematics, Computer Sc
Subject search and browse: Physics Search Form Interface

29 Aug 2012: Simons Foundation funds new arXiv sustainability model
See cumulative "What's New" pages. Read robots beware before attempt

Physics

- Astrophysics (**astro-ph** new, recent, find)
  includes: Cosmology and Extragalactic Astrophysics; Earth and Pla
  Solar and Stellar Astrophysics
- Condensed Matter (**cond-mat** new, recent, find)
  includes: Disordered Systems and Neural Networks; Materials Scie
  Strongly Correlated Electrons; Superconductivity
- General Relativity and Quantum Cosmology (**gr-qc** new, recent, fi
- High Energy Physics - Experiment (**hep-ex** new, recent, find)
- High Energy Physics - Lattice (**hep-lat** new, recent, find)
- High Energy Physics - Phenomenology (**hep-ph** new, recent, find)
- High Energy Physics - Theory (**hep-th** new, recent, find)
- Mathematical Physics (**math-ph** new, recent, find)
- Nonlinear Sciences (**nlin** new, recent, find)
  includes: Adaptation and Self-Organizing Systems; Cellular Auton
- Nuclear Experiment (**nucl-ex** new, recent, find)
- Nuclear Theory (**nucl-th** new, recent, find)

- arXiv.org is an electronic repository of scientific papers hosted by Cornell.
- Papers are in physics, math, CS, statistics, finance, and biology.
- arXiv currently has ≈800,000 articles, and 16 million unique users accessing the site each month.

The arXiv is an important repository of scientific articles

Google Scholar

physics

Top 20 publications matching *physics*

	Publication	h5-index	h5-median
1.	arXiv High Energy Physics - Theory (hep-th)	137	180
2.	arXiv High Energy Physics - Phenomenology (hep-ph)	135	182
3.	arXiv Mesoscale and Nanoscale Physics (cond-mat.mes-hall)	132	193
4.	arXiv Quantum Physics (quant-ph)	126	181
5.	Journal of High Energy Physics	124	167
6.	Applied Physics Letters	121	147
7.	Nature Physics	117	160
8.	Reviews of Modern Physics	94	210
9.	Physics Letters B	89	130
10.	The Journal of Chemical Physics	80	112
11.	arXiv High Energy Physics - Experiment (hep-ex)	78	113

- In several research areas in physics, the arXiv's impact factor is higher than that of any journal.

Our goal is to improve daily & weekly new-article feeds

Cornell University Library

arXiv.org > astro-ph

Astrophysics

New submissions

Submissions received from Mon 4 Mar 13 to Tue 5 Mar 13, announced

- New submissions
- Cross-lists
- Replacements

[total of 79 entries: 1-79]
[showing up to 2000 entries per page: fewer | more]

New submissions for Wed, 6 Mar 13

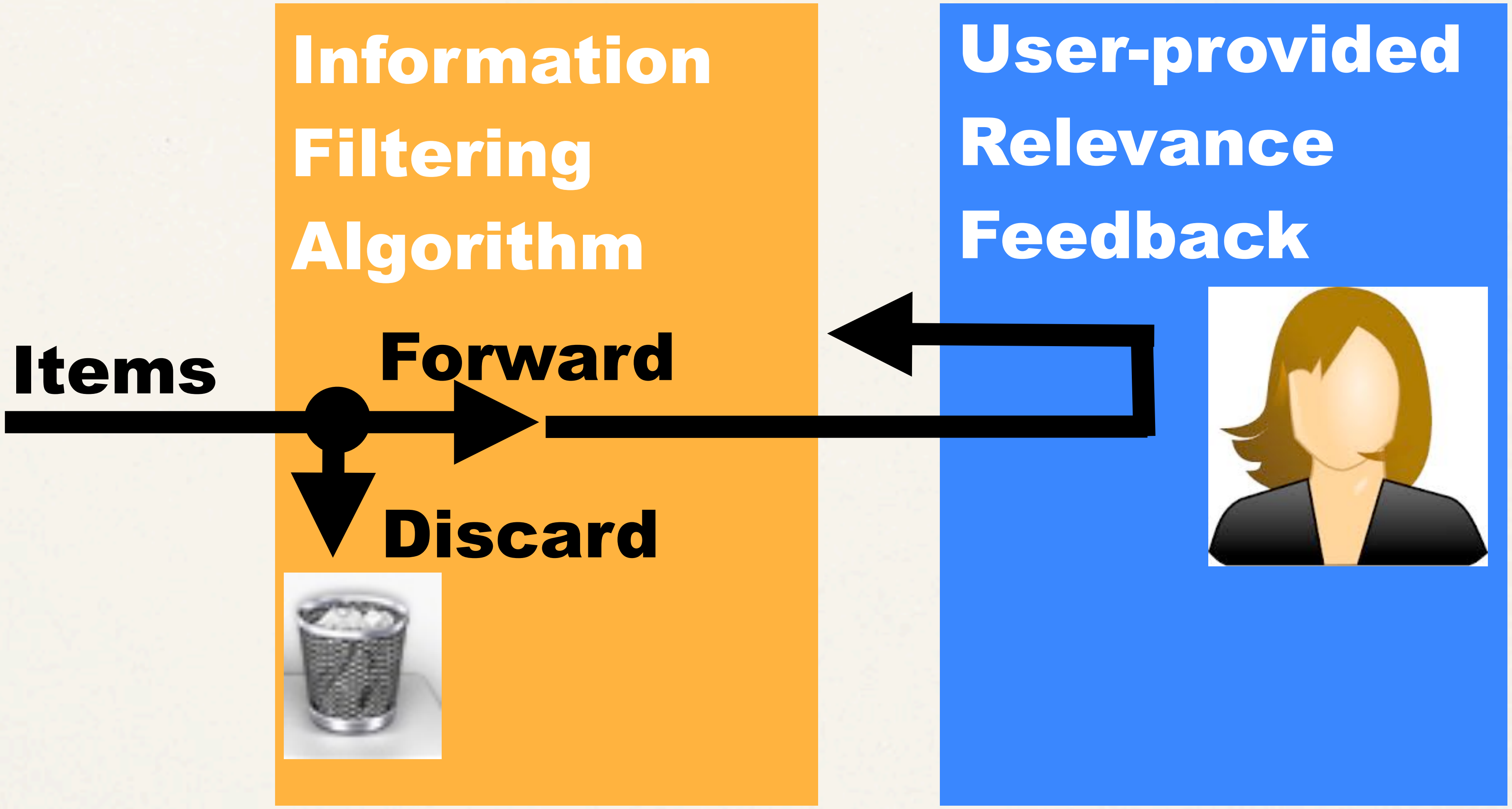
[1] arXiv:1303.0833 [pdf, ps, other]

Transverse oscillations in solar spicules induce
H. Ebadi, M. Hosseinpour, Z. Fazel
Comments: Accepted for publication in Astrophysics and Space Scien
Subjects: **Solar and Stellar Astrophysics (astro-ph.SR)**

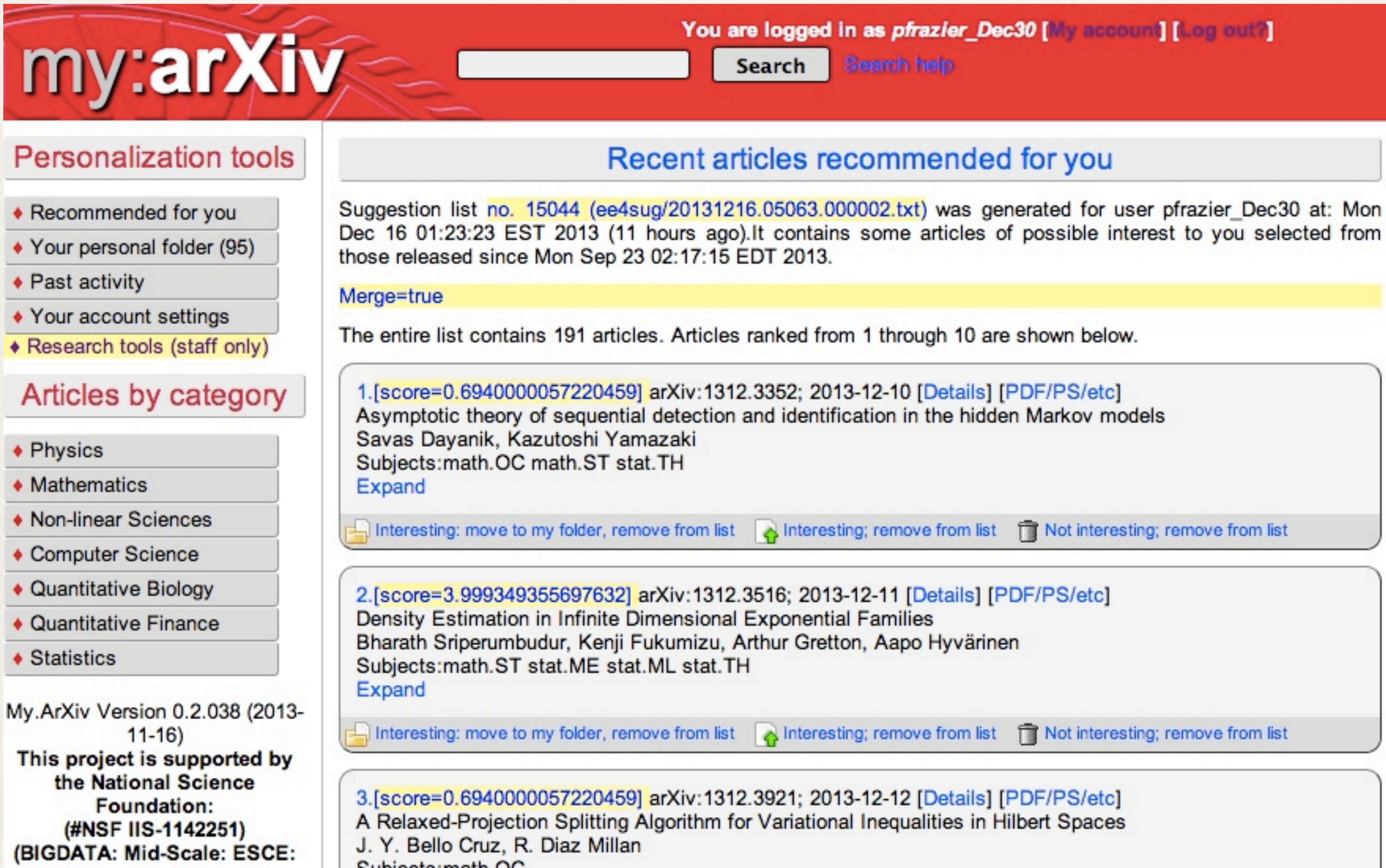
The excitation of Alfvenic waves in the solar spicules due to the
sheared magnetic fields is solved. Stratification due to gravity ar
transition region can penetrate from transition region into the c

- Many physicists visit the arXiv every day to browse the list of new papers, to stay aware of the latest research.
- There are lots of new papers (roughly 80 new papers / day in astrophysics.)
- Problem 1: Browsing this many papers is a lot of work for researchers.
- Problem 2: Researchers still miss important developments.

Our goal is to improve daily & weekly new-article feeds



Our goal is to improve daily & weekly new-article feeds



my:arXiv

You are logged in as *pfrazier_Dec30* [My account] [Log out?]

Search | Search help

Personalization tools

- Recommended for you
- Your personal folder (95)
- Past activity
- Your account settings
- Research tools (staff only)

Articles by category

- Physics
- Mathematics
- Non-linear Sciences
- Computer Science
- Quantitative Biology
- Quantitative Finance
- Statistics

My.ArXiv Version 0.2.038 (2013-11-16)

This project is supported by the National Science Foundation: (#NSF IIS-1142251) (BIGDATA: Mid-Scale: ESCE:

Recent articles recommended for you

Suggestion list no. 15044 (ee4sug/20131216.05063.000002.txt) was generated for user pfrazier_Dec30 at: Mon Dec 16 01:23:23 EST 2013 (11 hours ago).It contains some articles of possible interest to you selected from those released since Mon Sep 23 02:17:15 EDT 2013.

Merge=true

The entire list contains 191 articles. Articles ranked from 1 through 10 are shown below.

1.[score=0.6940000057220459] arXiv:1312.3352; 2013-12-10 [Details] [PDF/PS/etc]
Asymptotic theory of sequential detection and identification in the hidden Markov models
Savas Dayanik, Kazutoshi Yamazaki
Subjects:math.OC math.ST stat.TH
Expand

Interesting: move to my folder, remove from list | Interesting; remove from list | Not interesting; remove from list

2.[score=3.999349355697632] arXiv:1312.3516; 2013-12-11 [Details] [PDF/PS/etc]
Density Estimation in Infinite Dimensional Exponential Families
Bharath Sriperumbudur, Kenji Fukumizu, Arthur Gretton, Aapo Hyvärinen
Subjects:math.ST stat.ME stat.ML stat.TH
Expand

Interesting: move to my folder, remove from list | Interesting; remove from list | Not interesting; remove from list

3.[score=0.6940000057220459] arXiv:1312.3921; 2013-12-12 [Details] [PDF/PS/etc]
A Relaxed-Projection Splitting Algorithm for Variational Inequalities in Hilbert Spaces
J. Y. Bello Cruz, R. Diaz Millan

We also want to **understand** exploration vs. exploitation in information retrieval

- In this talk, we focus on the simplest of several models we have developed.
- The simplicity of the model makes clear the essential insights of our analysis into the exploration vs. exploitation tradeoff.
- However, building a system that provides value to users requires a number of tweaks to this simple model.
- We will discuss these tweaks briefly at the end of the talk.

Literature Review

- Exploration vs. exploitation has been studied extensively in the context of the multi-armed bandit problem:
  - Bayesian treatments: [Gittins & Jones, 1974; Whittle 1980] …
  - non-Bayesian treatments: [Auer, Cesa-Bianchi, Freund, Schapire, 1995; Auer, Cesa-Bianchi & Fischer, 2002] …
- Exploration vs. exploitation has also been studied in reinforcement learning [Kaelbling et al., 1998, Sutton and Barto, 1998].
- Exploration vs. exploitation has also been studied in information retrieval: [Zhang, Xu & Callan 2003; Agarwal, Chen & Elango 2009; Yue, Broder, Kleinberg & Joachims 2009; Hofmann, Whitestone & Rijke 2012]

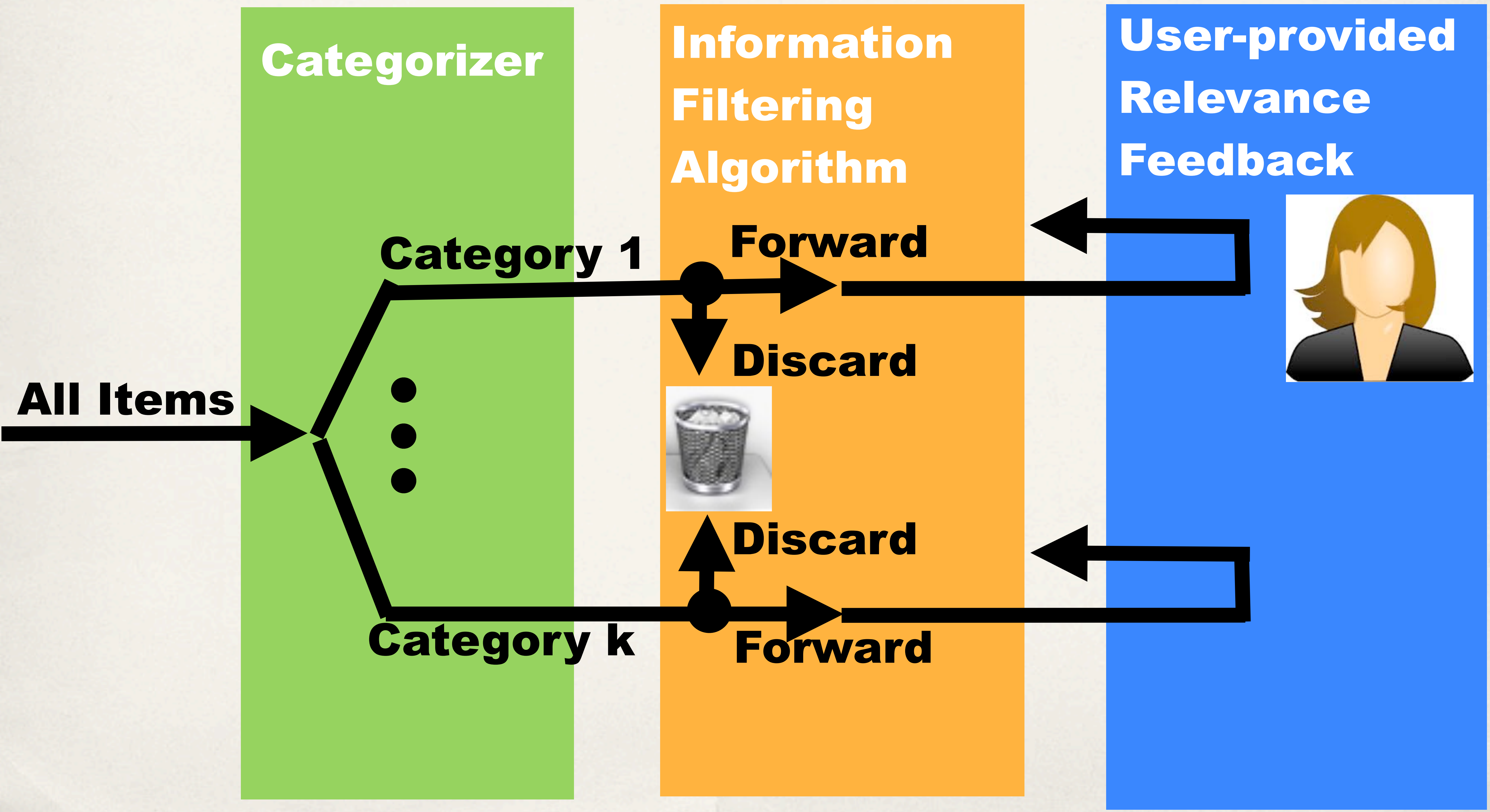
Outline

- Categorizing items
- Mathematical Model
- Extensions & Tweaks

Outline

- Categorizing items
- Mathematical Model
- Extensions & Tweaks

We use a pre-processing step that divides items into categories



We use a pre-processing step that divides items into categories

- **Step 1**: We use historical data to create a ratings matrix with older items and users with lots of history.
- **Step 2:** We use a singular value decomposition to represent older items as points in a low-dimensional space. Dimensions correspond roughly to “topics”.
- **Step 3:** We use kmeans clustering on the low-dimensional space to cluster older items.
- **Step 4**: We train a multi-class SVM to predict the cluster from item features, e.g., the words in a paper, or the authors.

We use a pre-processing step that divides items into categories

- Arxiv papers are also pre-labeled with categories: e.g., Artificial Intelligence; Computation and Language; Computational Complexity; Computational Engineering, Finance, and Science; Computational Geometry; Computer Science and Game Theory; Computer Vision and Pattern Recognition; …
- We are also experimenting with a Bayesian methods for categorizing documents into groups, designed to optimally support filtering.
- The specific method used to divide documents into groups is not important for understanding the main ideas in this talk.

Outline

- Categorizing items
- Mathematical Model
- Extensions & Tweaks

Mathematical Model

- An item from category x is relevant to the user with probability θ_x.
- We begin with a Bayesian prior distribution on θ_x, which is independent across x.

$$\theta_x \sim \mathrm{Beta}(\alpha_{0x}, \beta_{0x})$$

- Items arrive according to a Poisson process with rate λ.
- An item falls into category x with probability p_x. An item's category is observable. Thus, items from category x arrive according to a Poisson process with rate $\lambda_x = \lambda p_x$.
- When each paper arrives, we decide whether to forward or discard. For the n^{th} item from category x, let $U_{nx}=1$ if we forward it, and 0 if not.

Mathematical Model

- When each item arrives, we decide whether to forward or discard. For the n^{th} item from category x, let $U_{nx}=1$ if we forward it, and 0 if not.

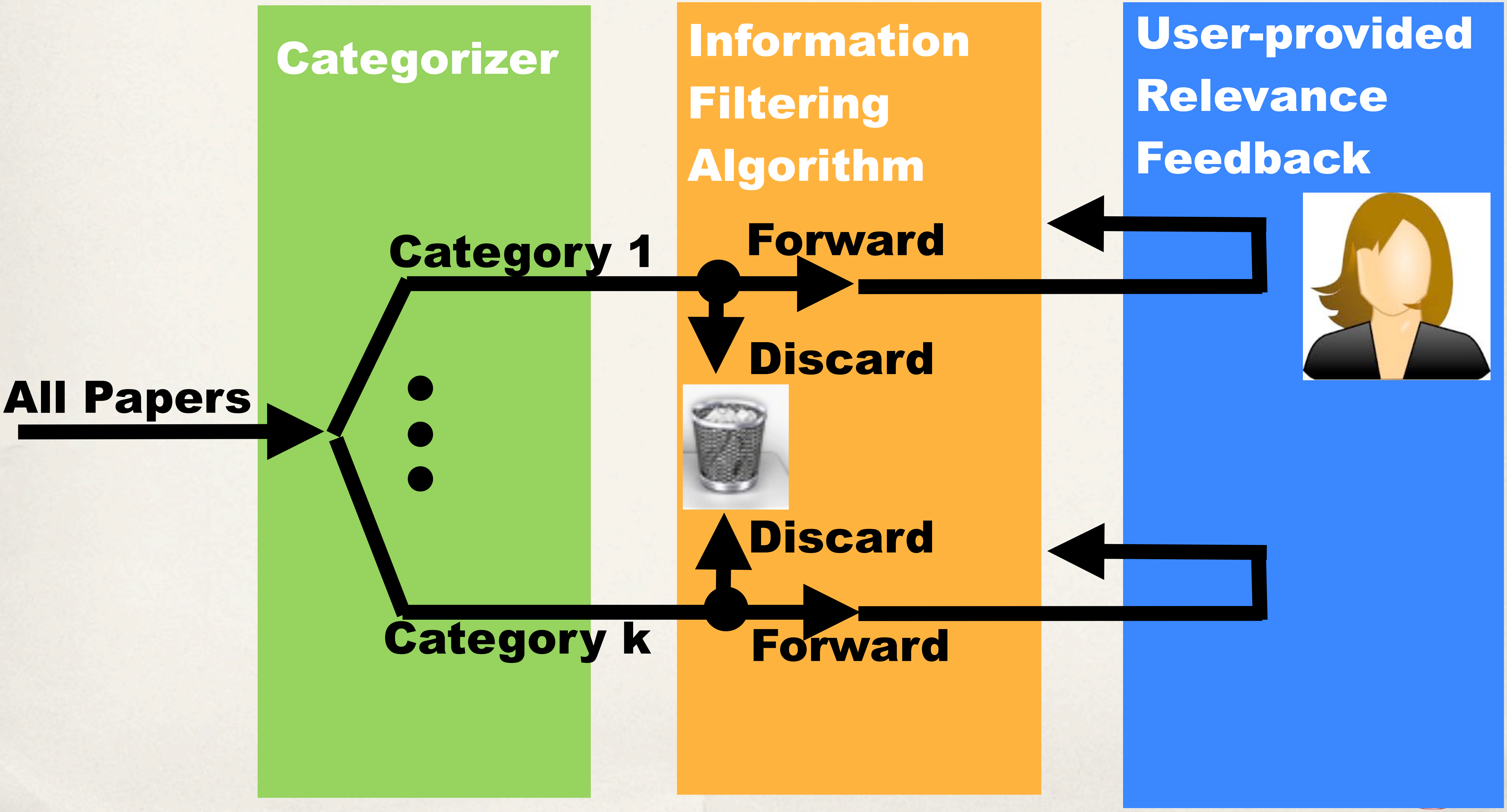
- If $U_{nx}=1$, we then observe Y_{nx}, which is 1 if the item was relevant to the user, and 0 if not.

$$Y_{nx}|\theta_{nx} \sim \text{Bernoulli}(\theta_x)$$

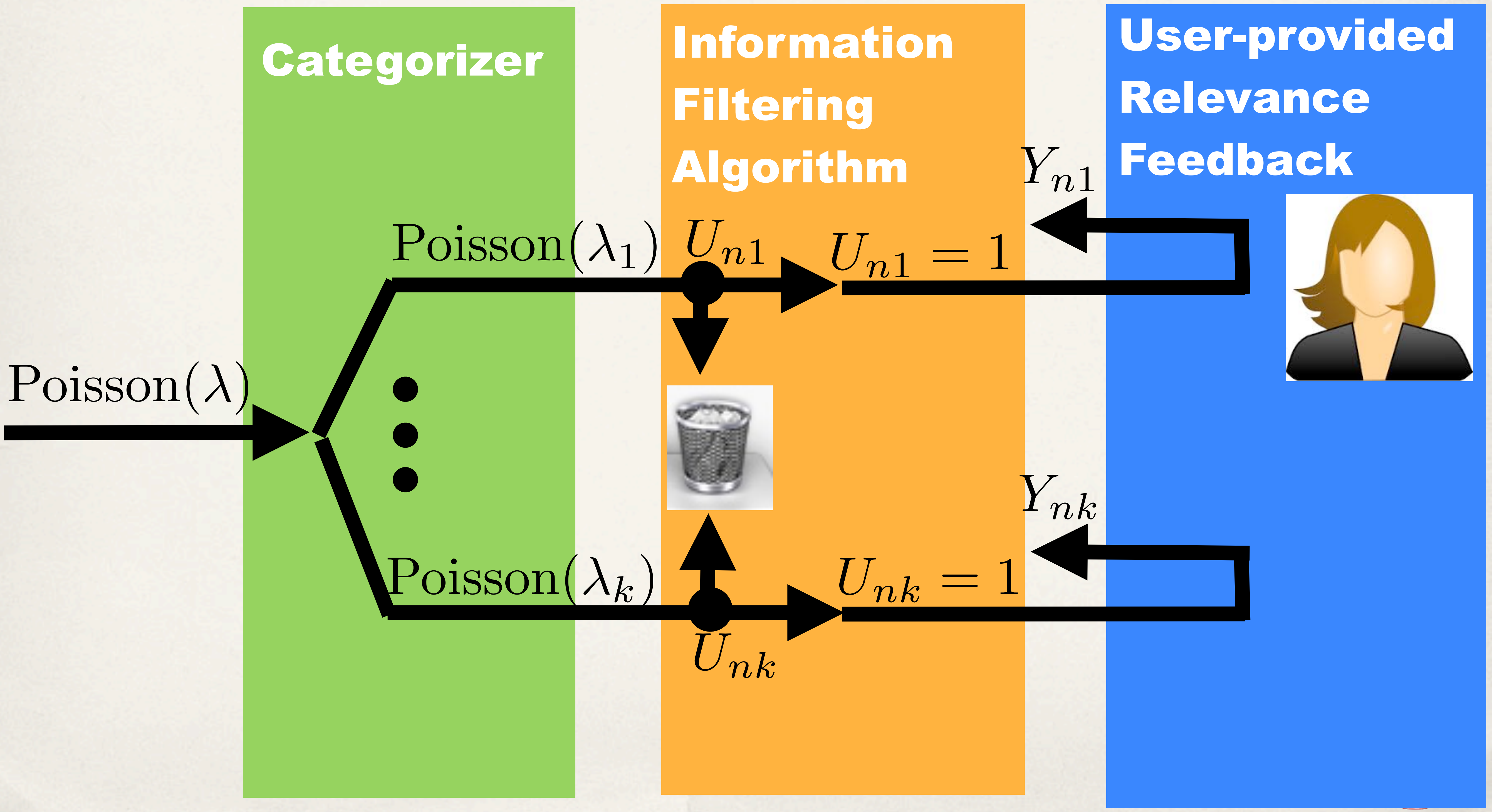
- We can then update our posterior distribution on θ_x, which will still be Beta-distributed (details later),

$$\theta_x|(Y_{mx} : m \leq n, U_{mx} = 1) \sim \text{Beta}(\alpha_{nx}, \beta_{nx})$$

Mathematical Model



Mathematical Model



Mathematical Model

- To model the cost of the user's time, we penalize ourselves with a cost c for forwarding an item. [more on the choice of c later]
- We give ourselves a reward of 1 for showing a relevant item.
- Our net reward is U_{nx} (Y_{nx}-c).
- Our goal is to design an algorithm π that maximizes

$$E^{\pi}\left[\sum_{x=1}^{k}\sum_{n=1}^{N_x} U_{nx}(Y_{nx}-c)\right]$$

Mathematical Model

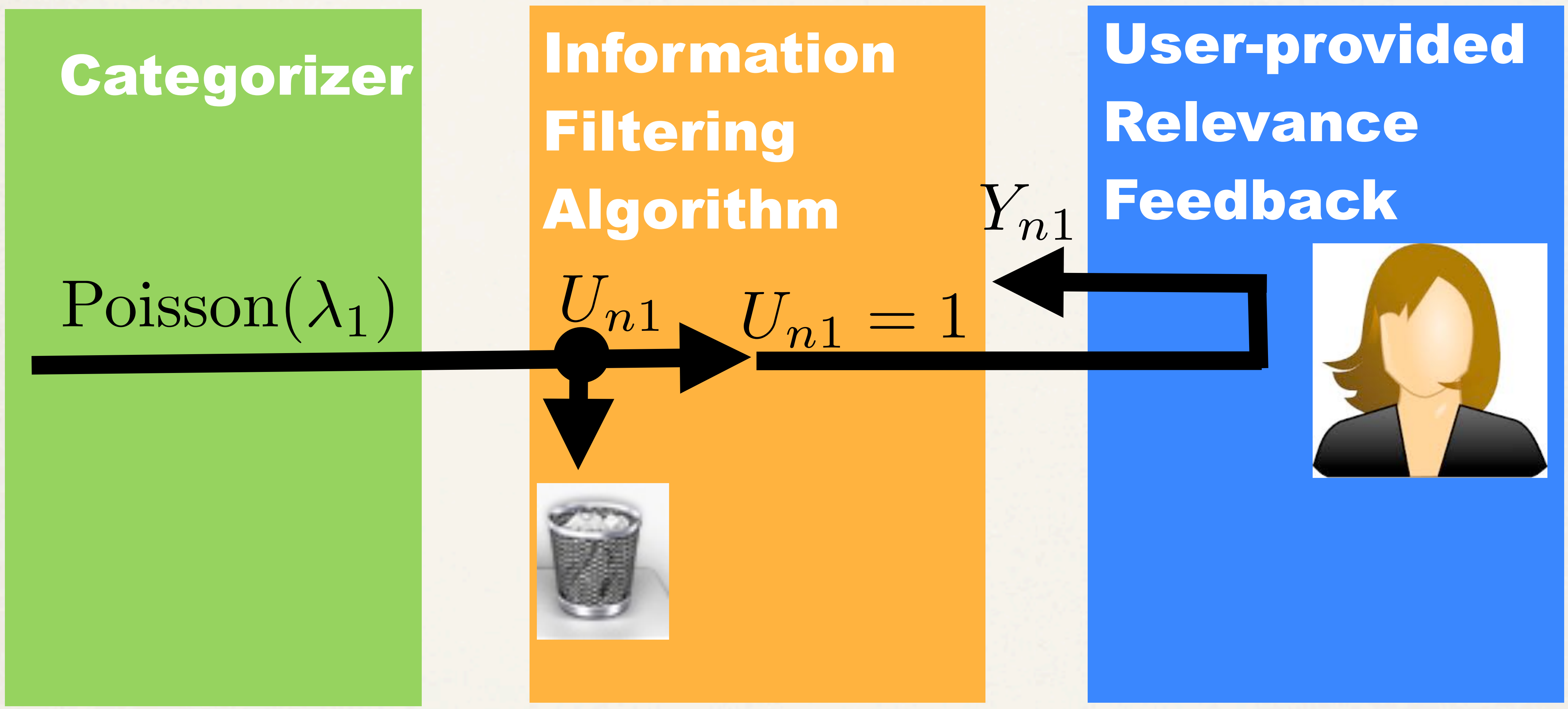
- Our goal is to solve:

$$\sup_{\pi} E^{\pi}\left[\sum_{x=1}^{k}\sum_{n=1}^{N_x} U_{nx}(Y_{nx}-c)\right]$$

- Here, $N_x = \sup\{n : t_{nx} \leq T\}$ is the number of items from category x seen by the user, up to some random time horizon T, and t_{nx} is the arrival time of the n^{th} item in category x. We construct T so that N_x is geometric.

- An algorithm π is a rule for choosing each U_{nx} based only on previously observed feedback $(Y_{mz} : U_{mz}=1, t_{mz} < t_{nx})$,

Let’s first solve the problem for a single category



Let's first solve the problem for a single category

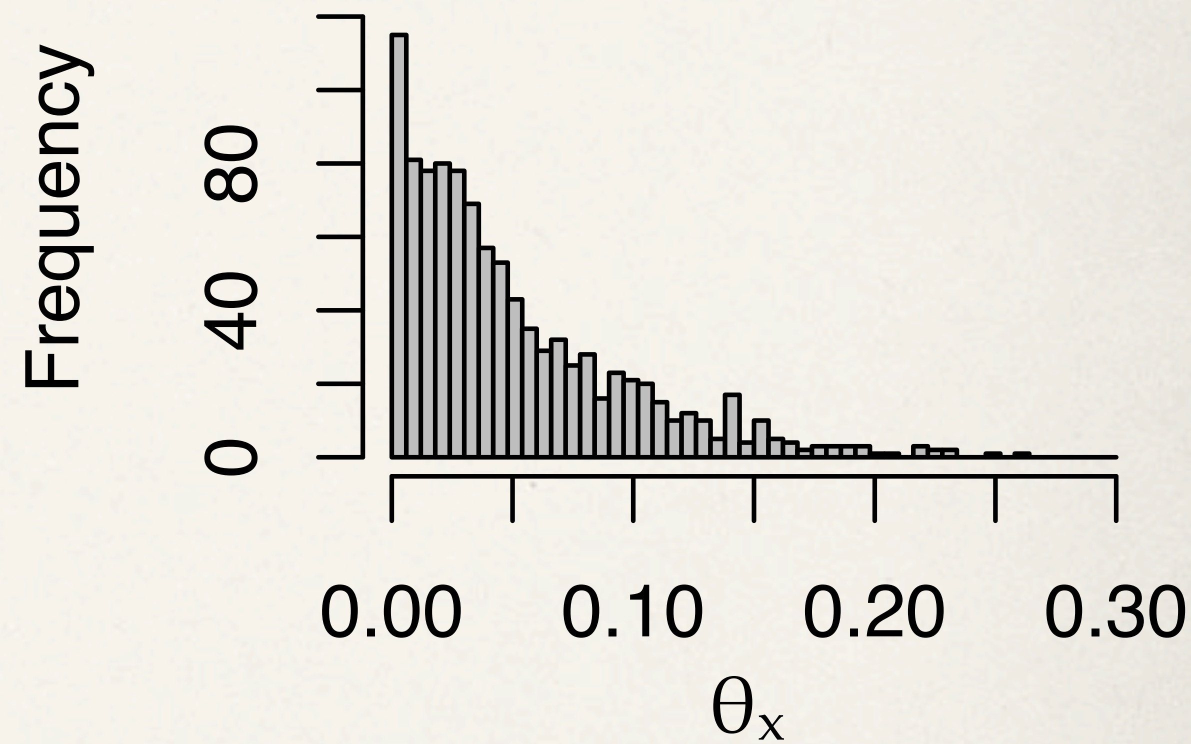
- For a given cluster x, let's figure out how to maximize the reward from just that cluster,

$$\sup_{\pi} E^{\pi}\left[\sum_{n=1}^{N_x} U_{nx}(Y_{nx} - c)\right]$$

- When choosing U_{nx}, it is sufficient to consider feedback only from previous items in our category x, $(Y_{mx} : U_{mx}=1, m<n)$

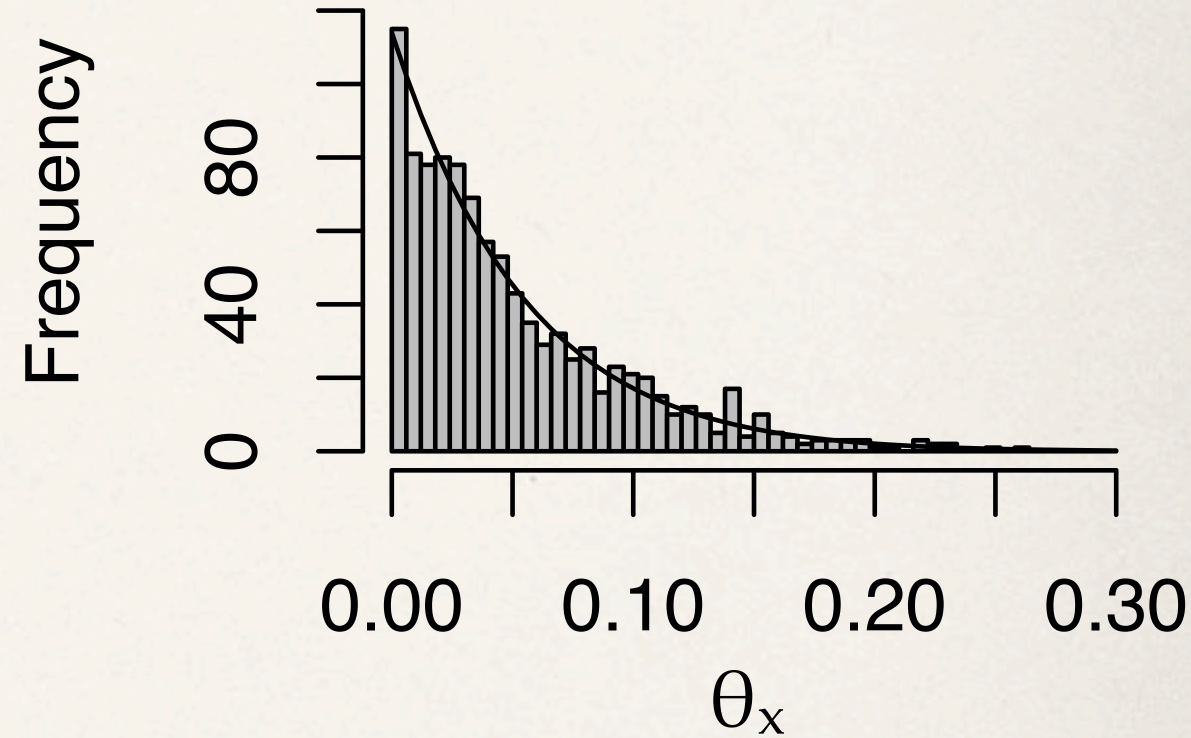
We use a standard Bayesian statistical model

- Recall that we model $\theta_x \sim \text{Beta}(\alpha_{0x}, \beta_{0x})$.
- Here's how we choose α_{0x} and β_{0x}.
  - We first find a few users with lots of historical data in this cluster.
  - We estimate θ_x for each of these users, using their average relevance feedback.
  - We then make a histogram.



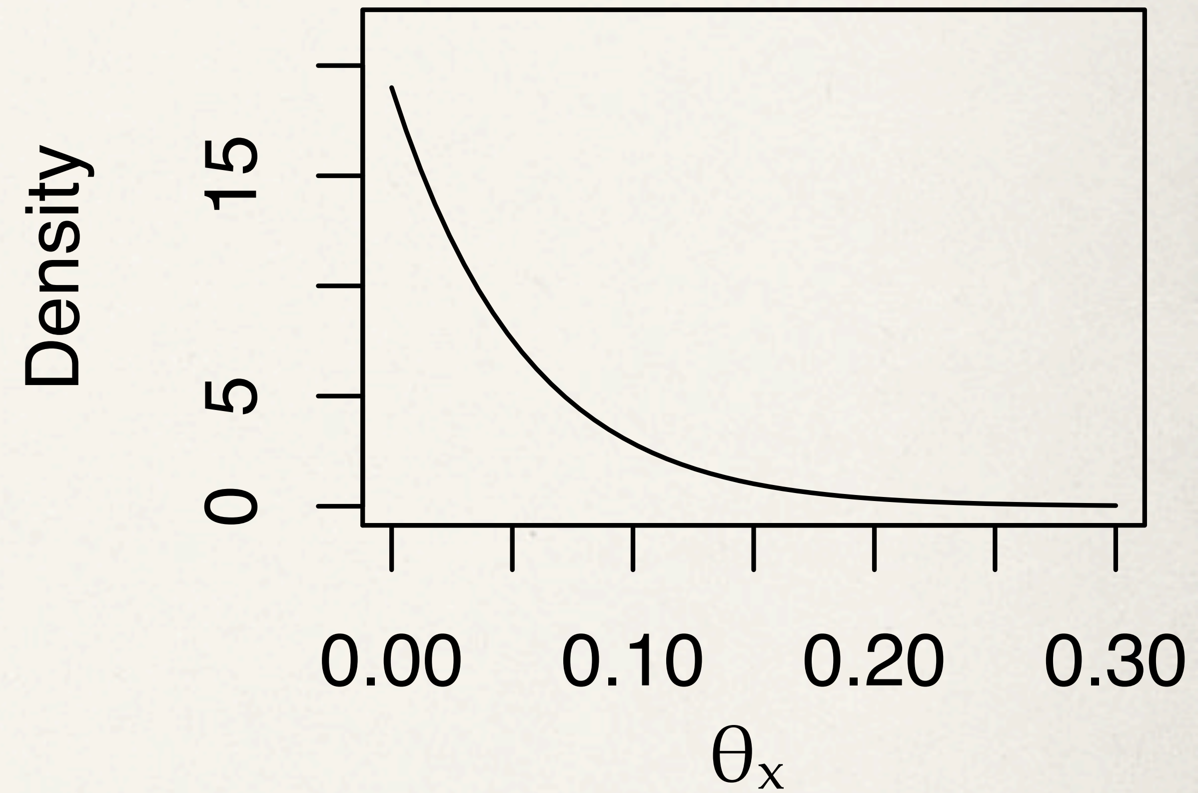
We use a standard Bayesian statistical model

- Recall that we model $\theta_x \sim \mathrm{Beta}(\alpha_{0x}, \beta_{0x})$.
- Here's how we choose α_{0x} and β_{0x}.
    - We then fit a beta density to this empirical distribution, using maximum likelihood estimation.
    - We set α_{0x} and β_{0x} to their values from the fitted distribution.
    - A beta distribution is analytically convenient, and fits the data well.



We use a standard Bayesian statistical model

- Recall that we model $\theta_x \sim \text{Beta}(\alpha_{0x}, \beta_{0x})$.
- Here's how we choose α_{0x} and β_{0x}.
    - We then fit a beta density to this empirical distribution, using maximum likelihood estimation.
    - We set α_{0x} and β_{0x} to their values from the fitted distribution.
    - A beta distribution is analytically convenient, and fits the data well.



We use a standard Bayesian statistical model

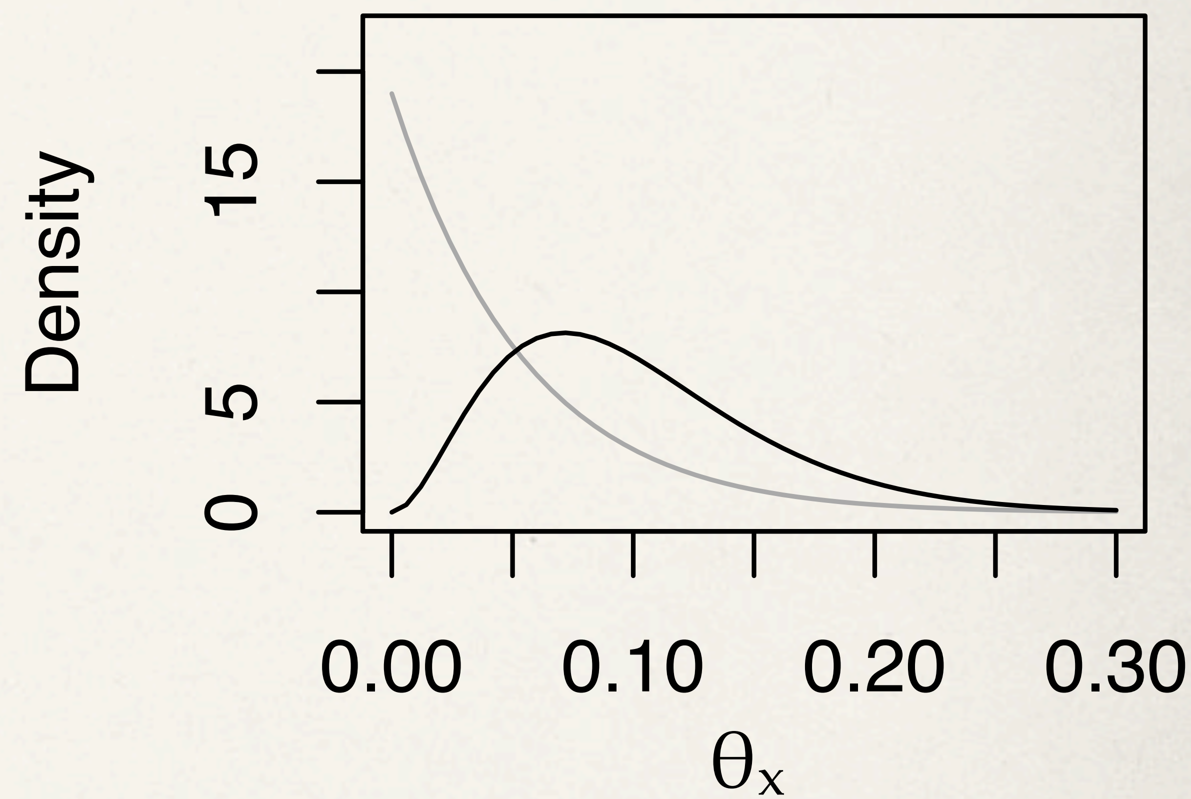
- After observing our data, we update our prior to obtain a posterior distribution using Bayes rule.

$$\theta_x | (Y_{mx} : m \leq n, U_{mx} = 1)$$
$$\sim \text{Beta}(\alpha_{nx}, \beta_{nx})$$

- Here, α_{nx} and β_{nx} count the effective numbers of relevant and irrelevant items shown:

$$\alpha_{nx} = \alpha_{0x} + \sum_{m=1}^{n} U_{mx} Y_{mx}$$

$$\beta_{nx} = \beta_{0x} + \sum_{m=1}^{n} U_{mx}(1 - Y_{mx})$$



We use a standard Bayesian statistical model

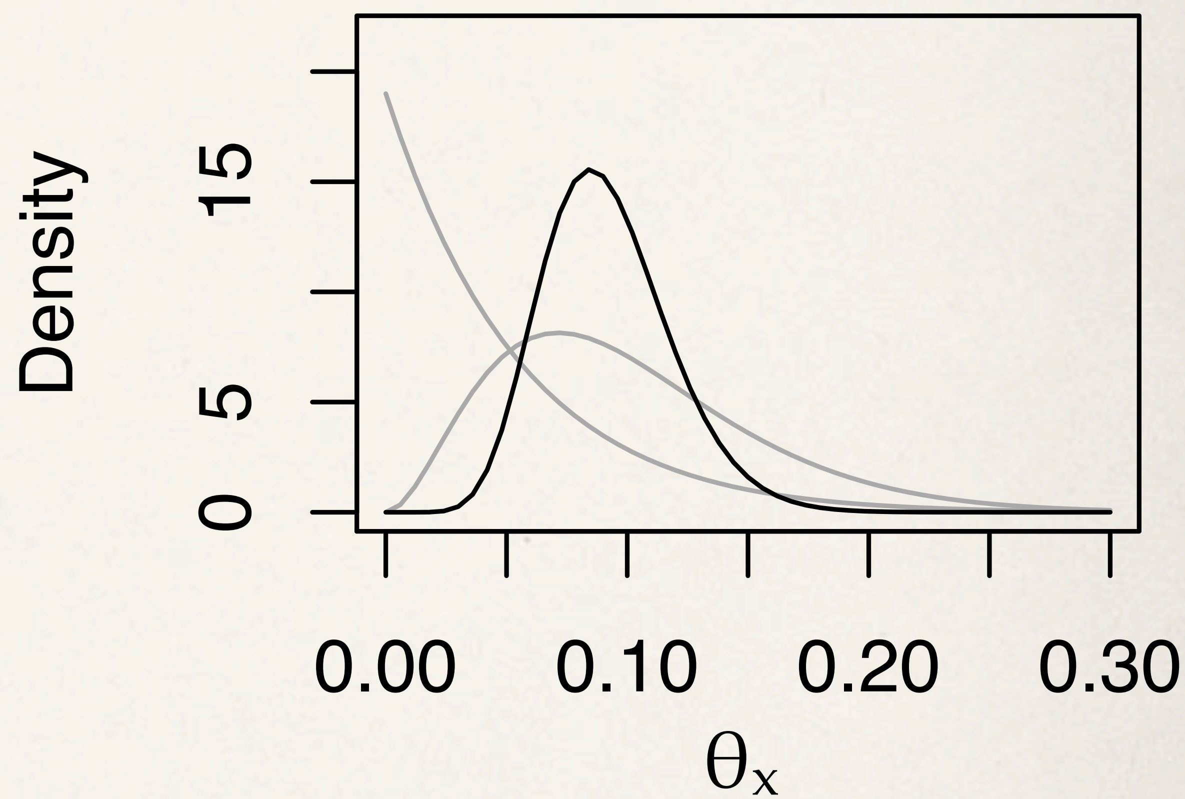
- After observing our data, we update our prior to obtain a posterior distribution using Bayes rule.

$$\theta_x | (Y_{mx} : m \leq n, U_{mx} = 1)$$
$$\sim \mathrm{Beta}(\alpha_{nx}, \beta_{nx})$$

- Here, α_{nx} and β_{nx} count the effective numbers of relevant and irrelevant items shown:

$$\alpha_{nx} = \alpha_{0x} + \sum_{m=1}^{n} U_{mx} Y_{mx}$$

$$\beta_{nx} = \beta_{0x} + \sum_{m=1}^{n} U_{mx}(1 - Y_{mx})$$



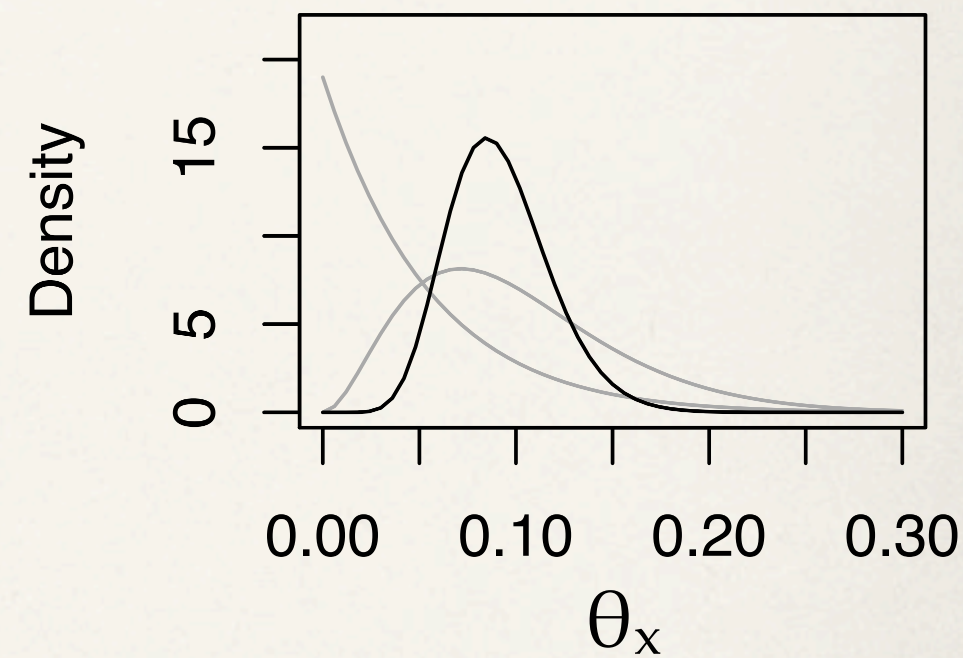
We use a standard Bayesian statistical model

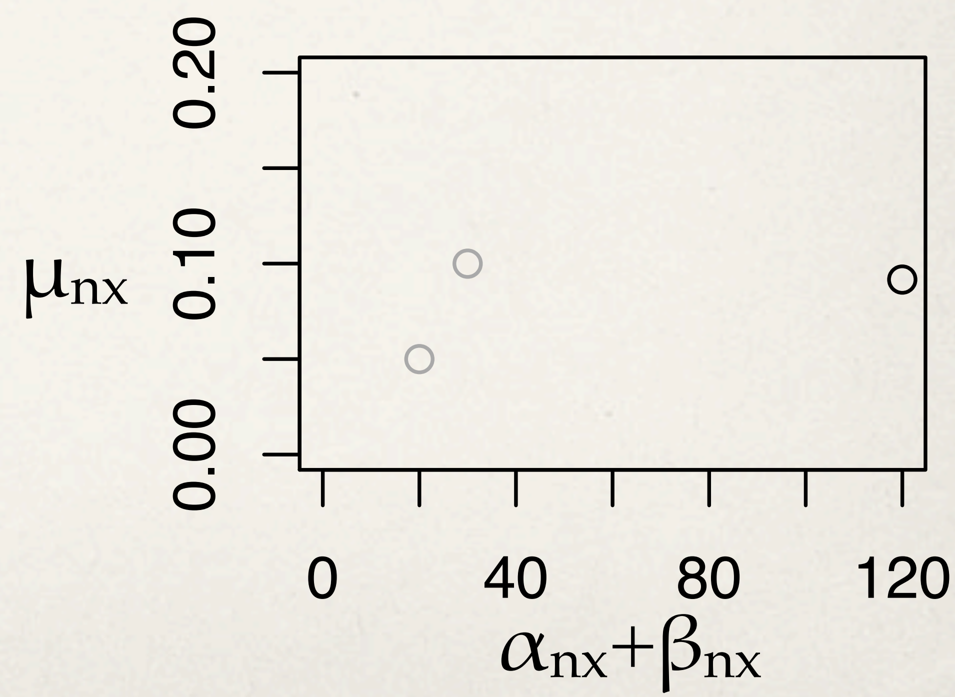
- Our posterior is

$$\theta_x|(Y_{mx} : m \leq n, U_{mx} = 1)$$
$$\sim \text{Beta}(\alpha_{nx}, \beta_{nx})$$

- We can parameterize this posterior with (μ_{nx}, $\alpha_{nx}+\beta_{nx}$) where

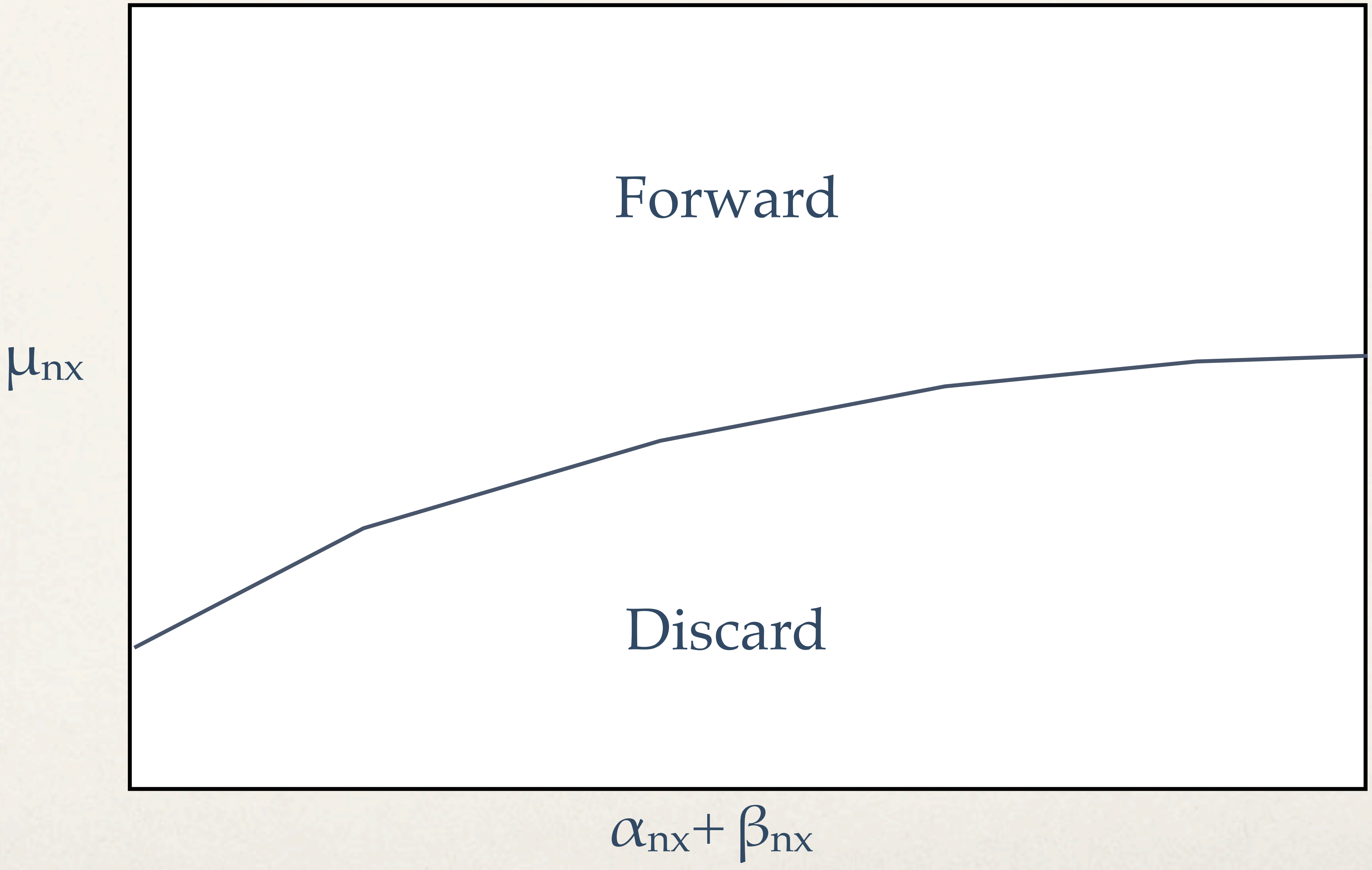
$$\mu_{nx} = E_n\left[\theta_x\right] = \frac{\alpha_{nx}}{\alpha_{nx} + \beta_{nx}}$$





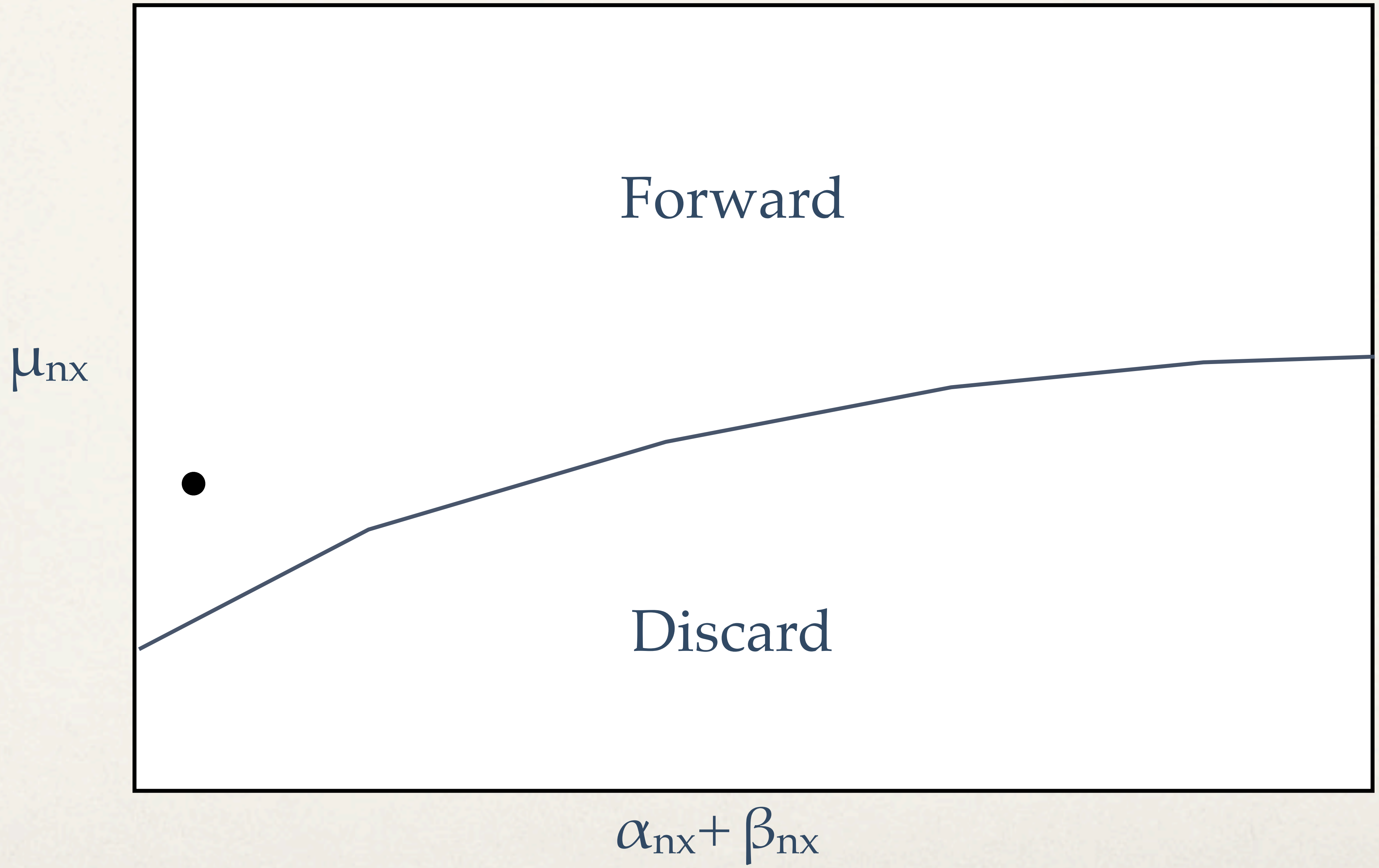
An algorithm partitions the space of posteriors into “Forward” and “Discard”

- Here is one possible algorithm:



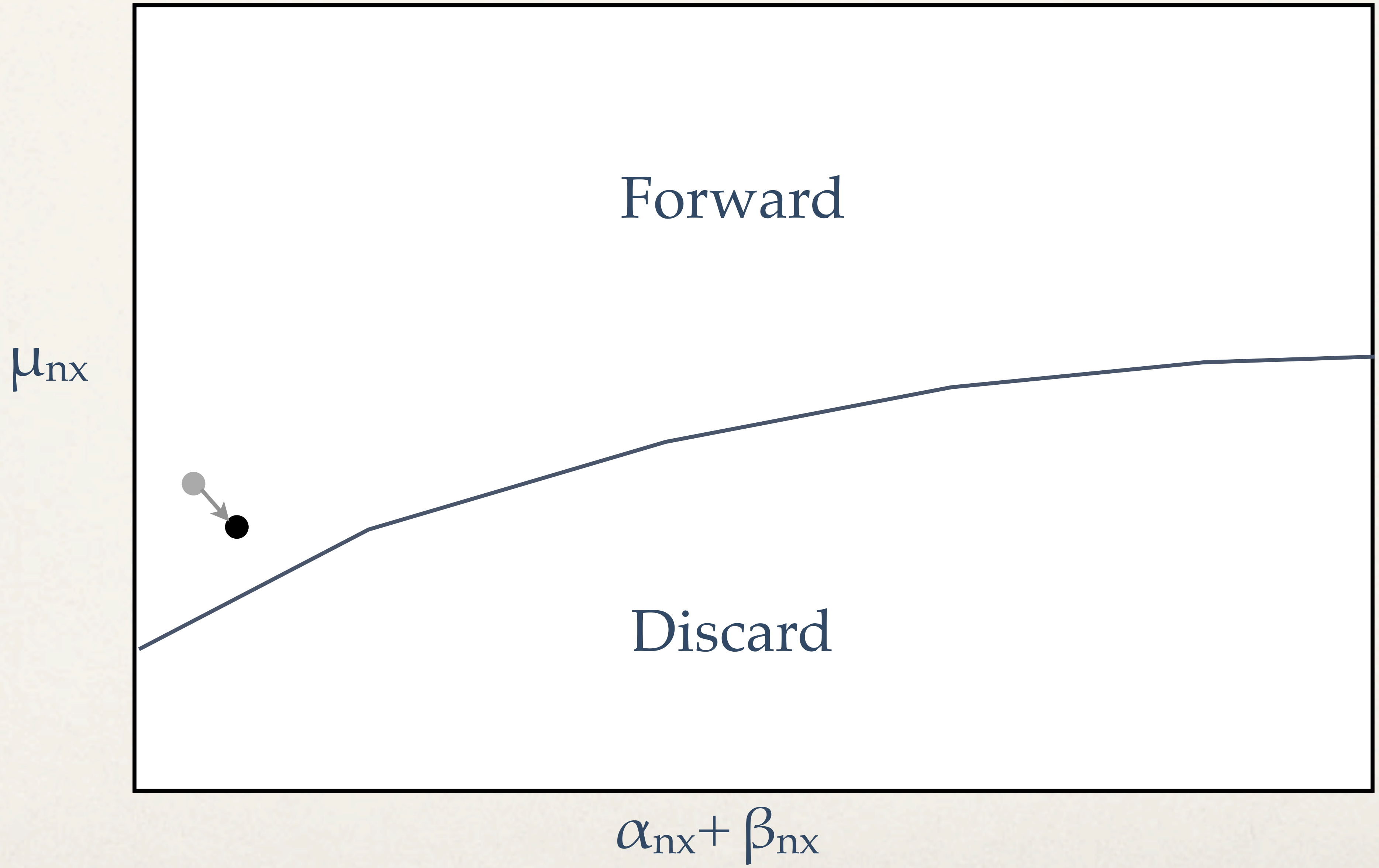
An algorithm partitions the space of posteriors into “Forward” and “Discard”

- Here is one possible algorithm:



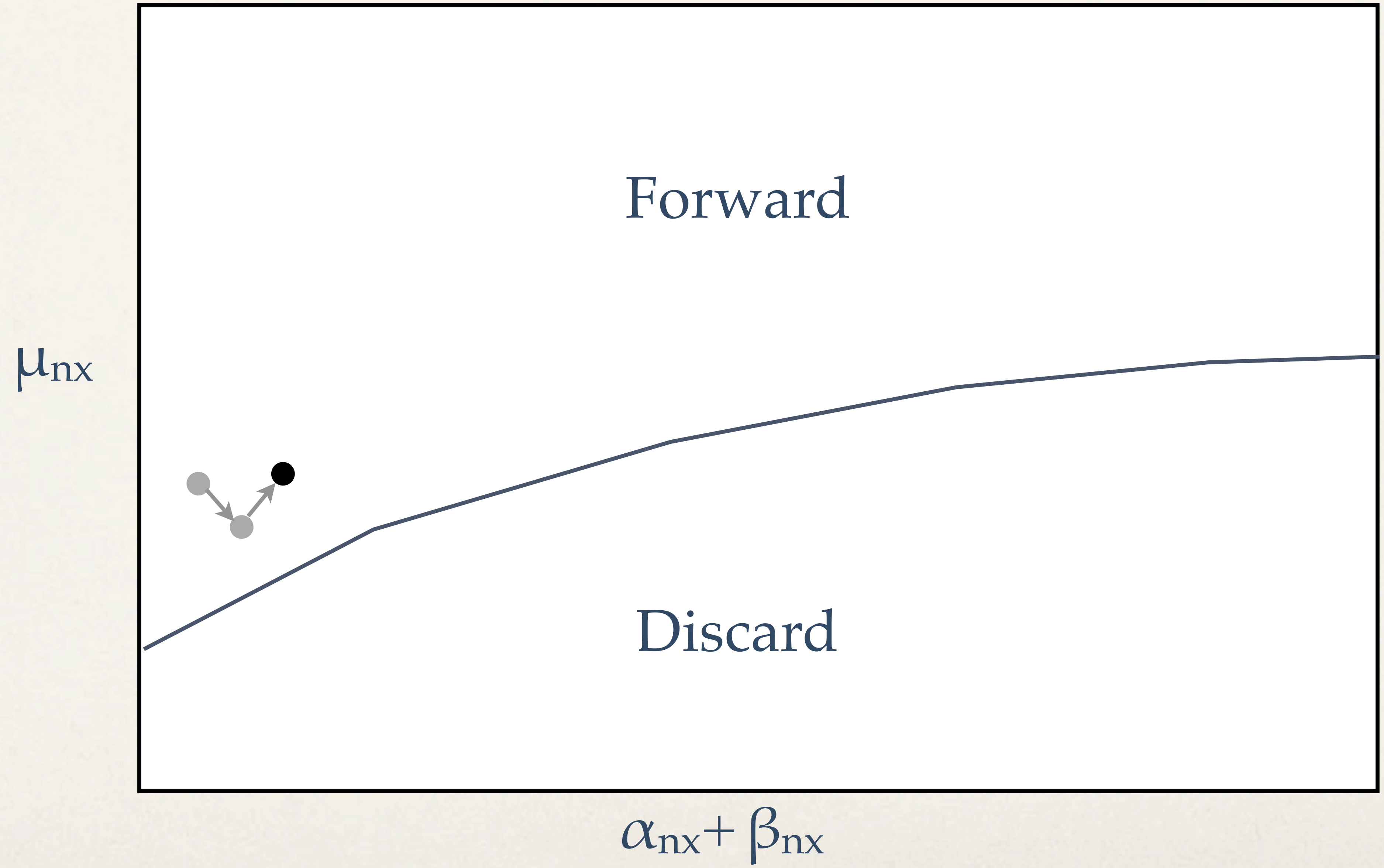
An algorithm partitions the space of posteriors into "Forward" and "Discard"

- Here is one possible algorithm:



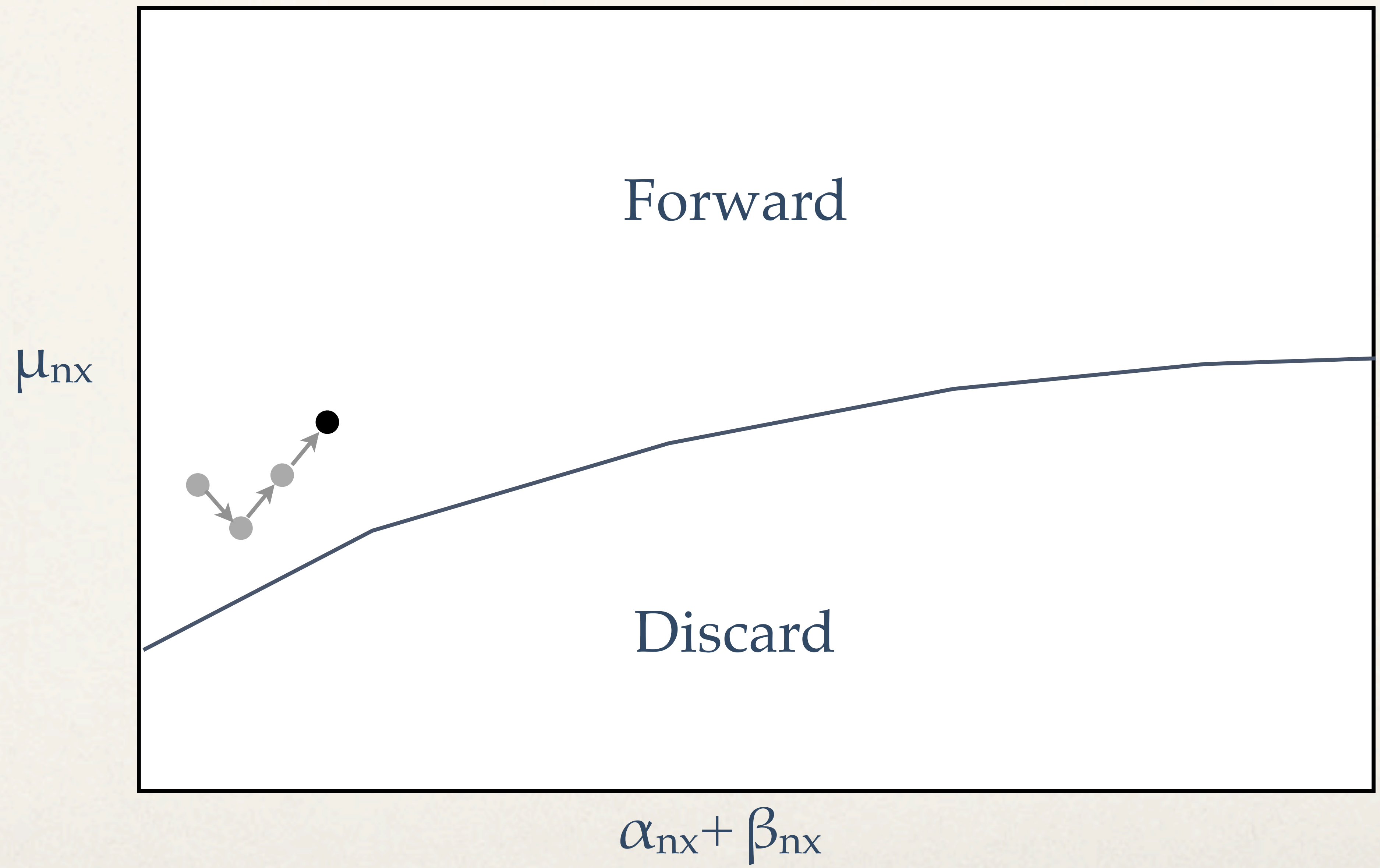
An algorithm partitions the space of posteriors into “Forward” and “Discard”

- Here is one possible algorithm:



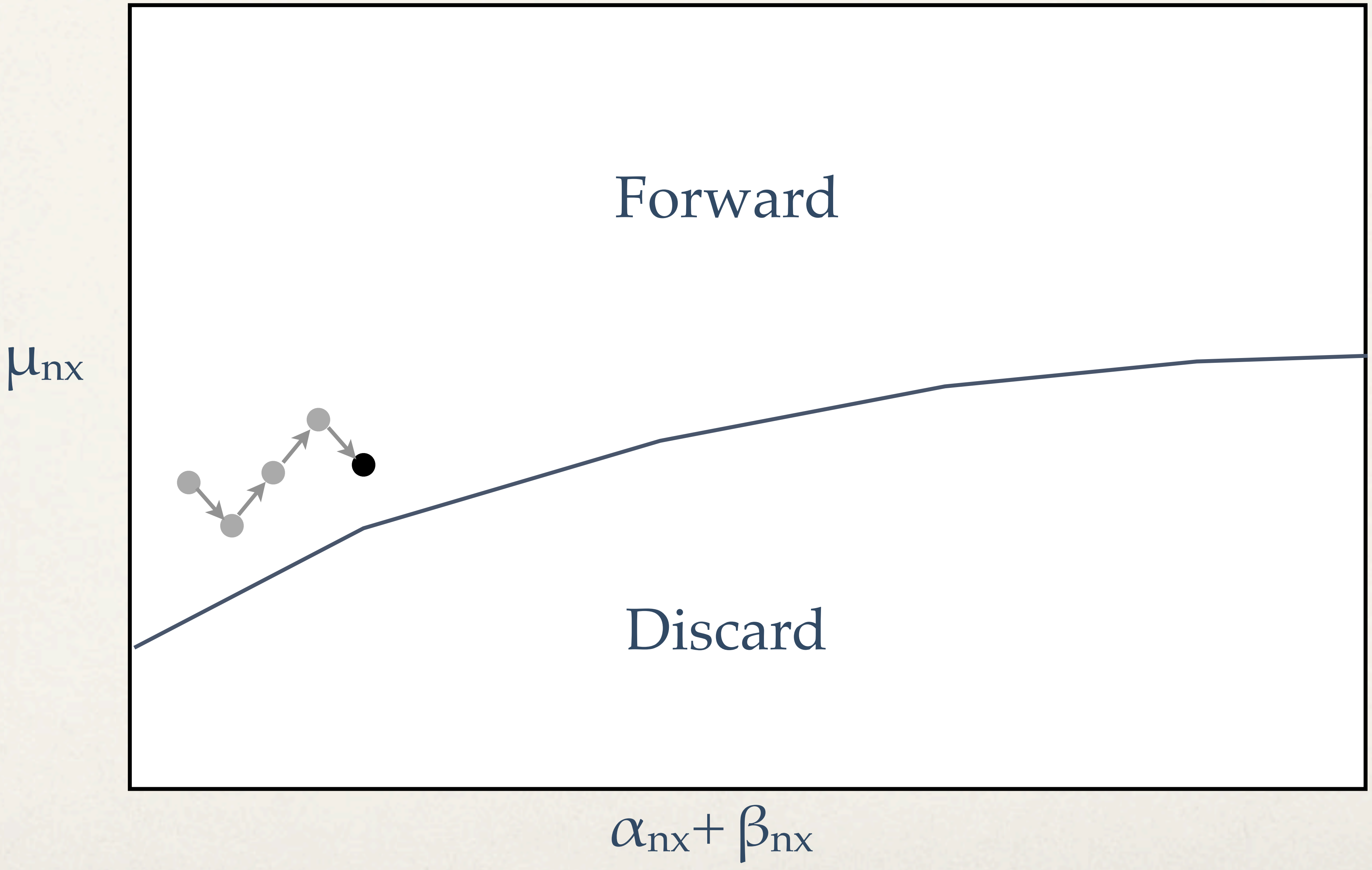
An algorithm partitions the space of posteriors into “Forward” and “Discard”

- Here is one possible algorithm:



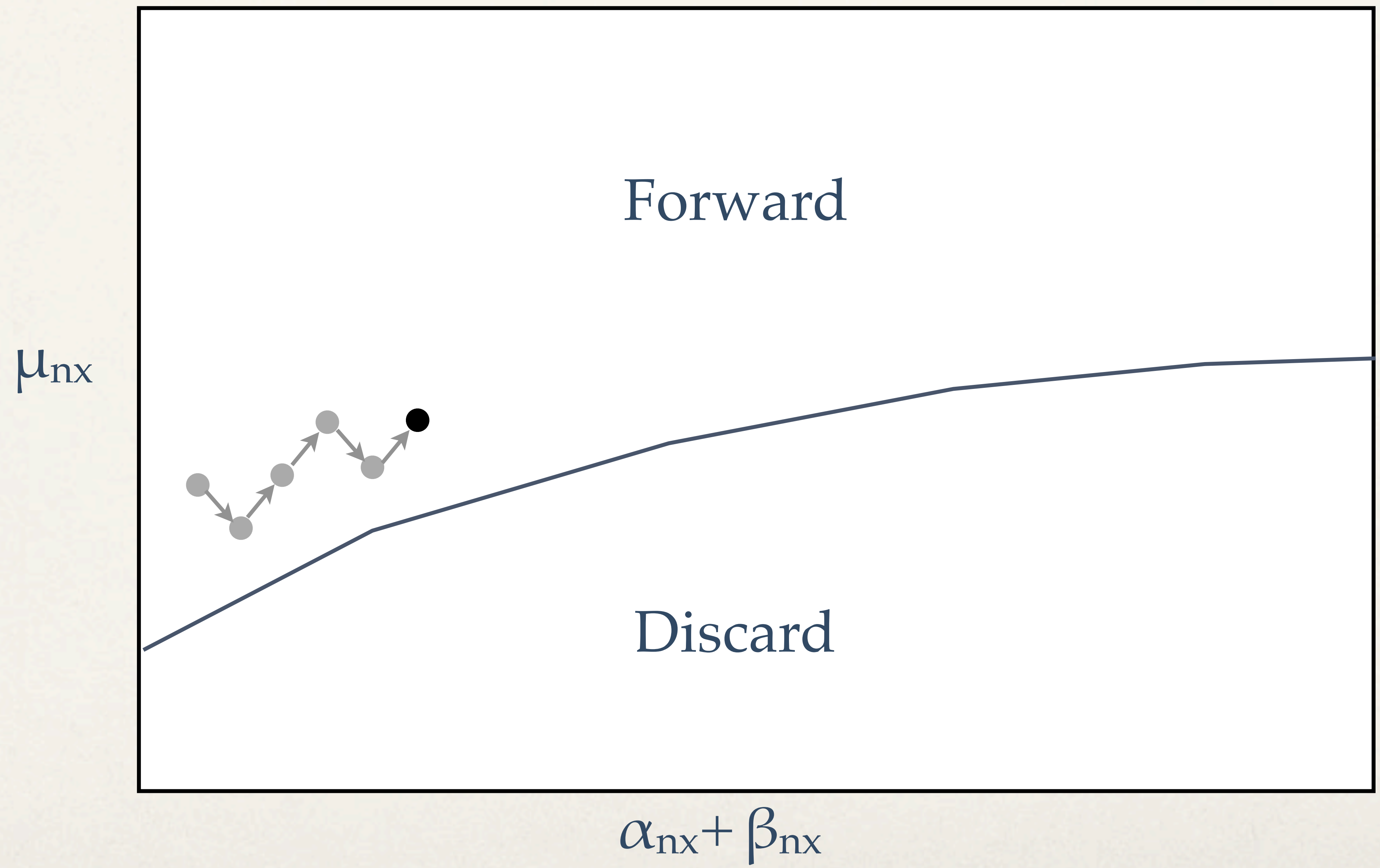
An algorithm partitions the space of posteriors into "Forward" and "Discard"

- Here is one possible algorithm:



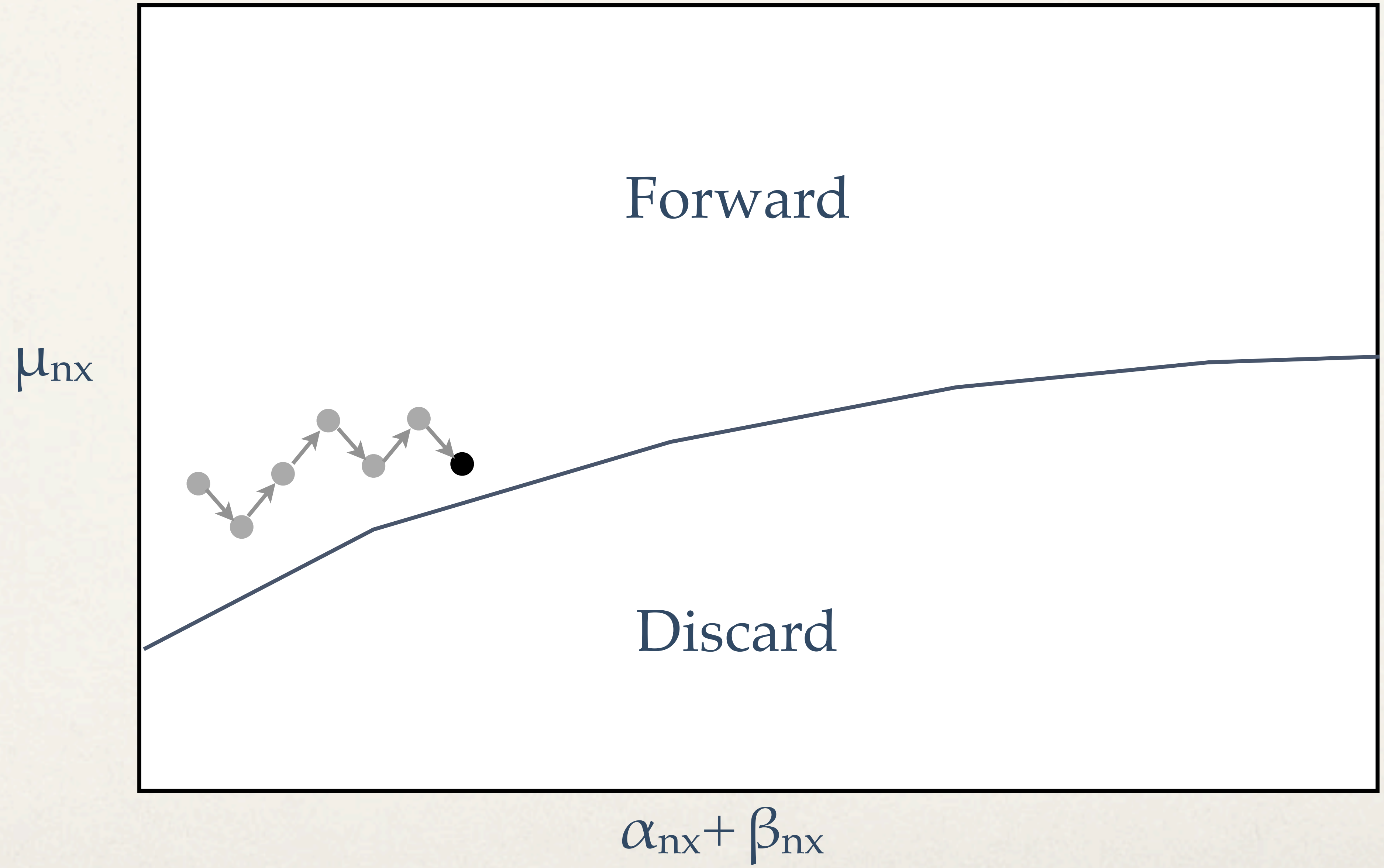
An algorithm partitions the space of posteriors into "Forward" and "Discard"

- Here is one possible algorithm:



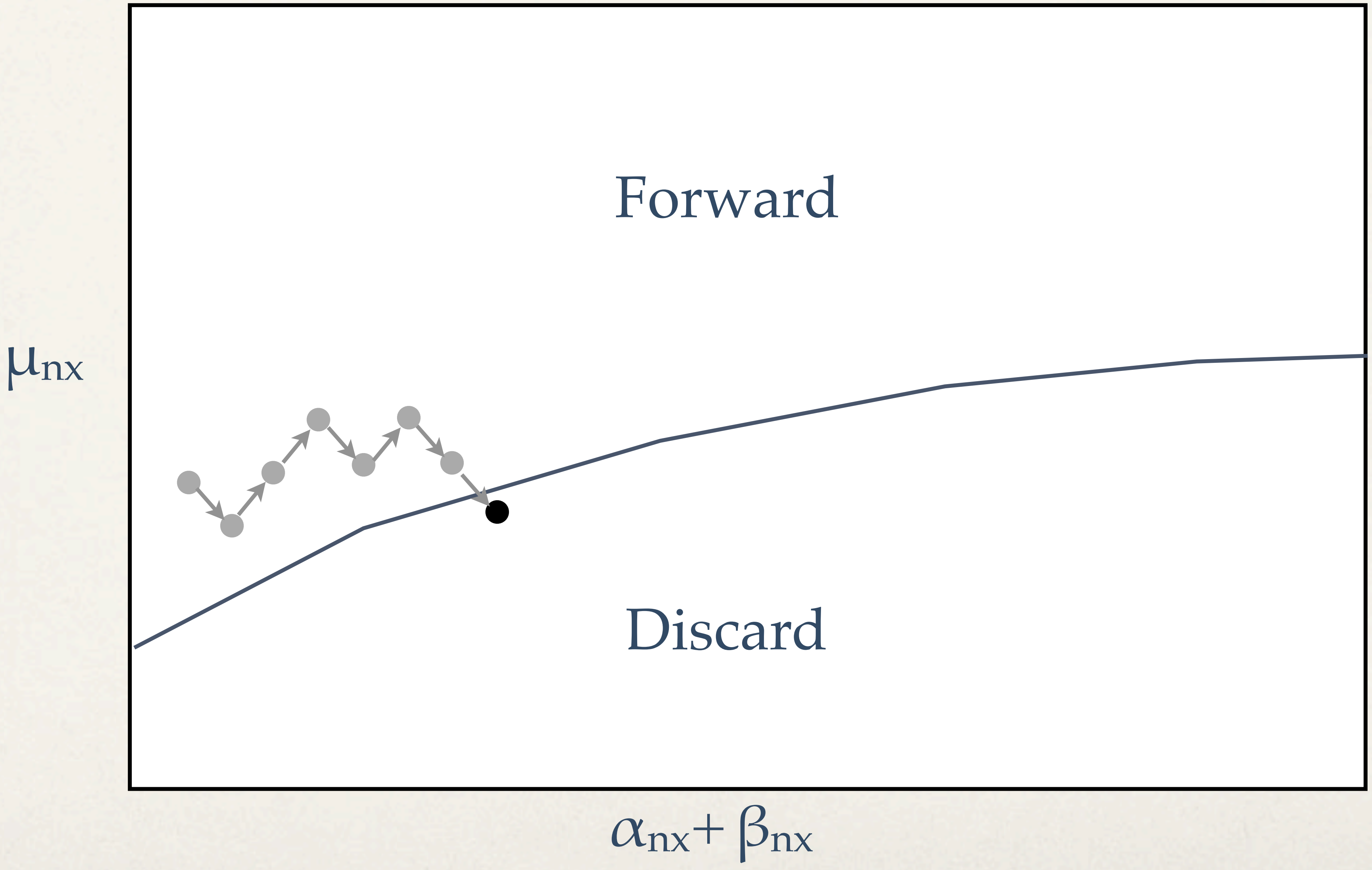
An algorithm partitions the space of posteriors into “Forward” and “Discard”

- Here is one possible algorithm:



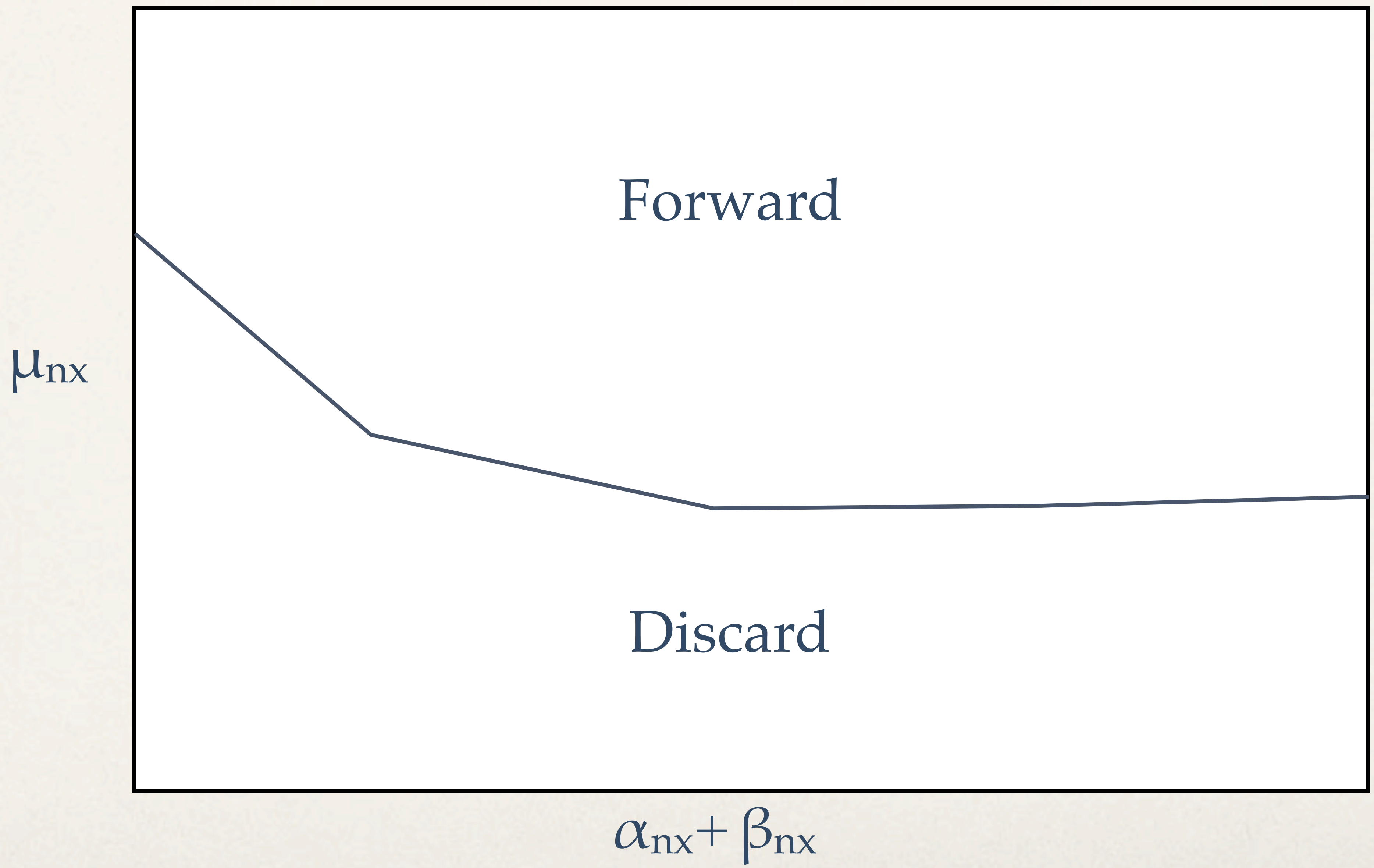
An algorithm partitions the space of posteriors into "Forward" and "Discard"

- Here is one possible algorithm:



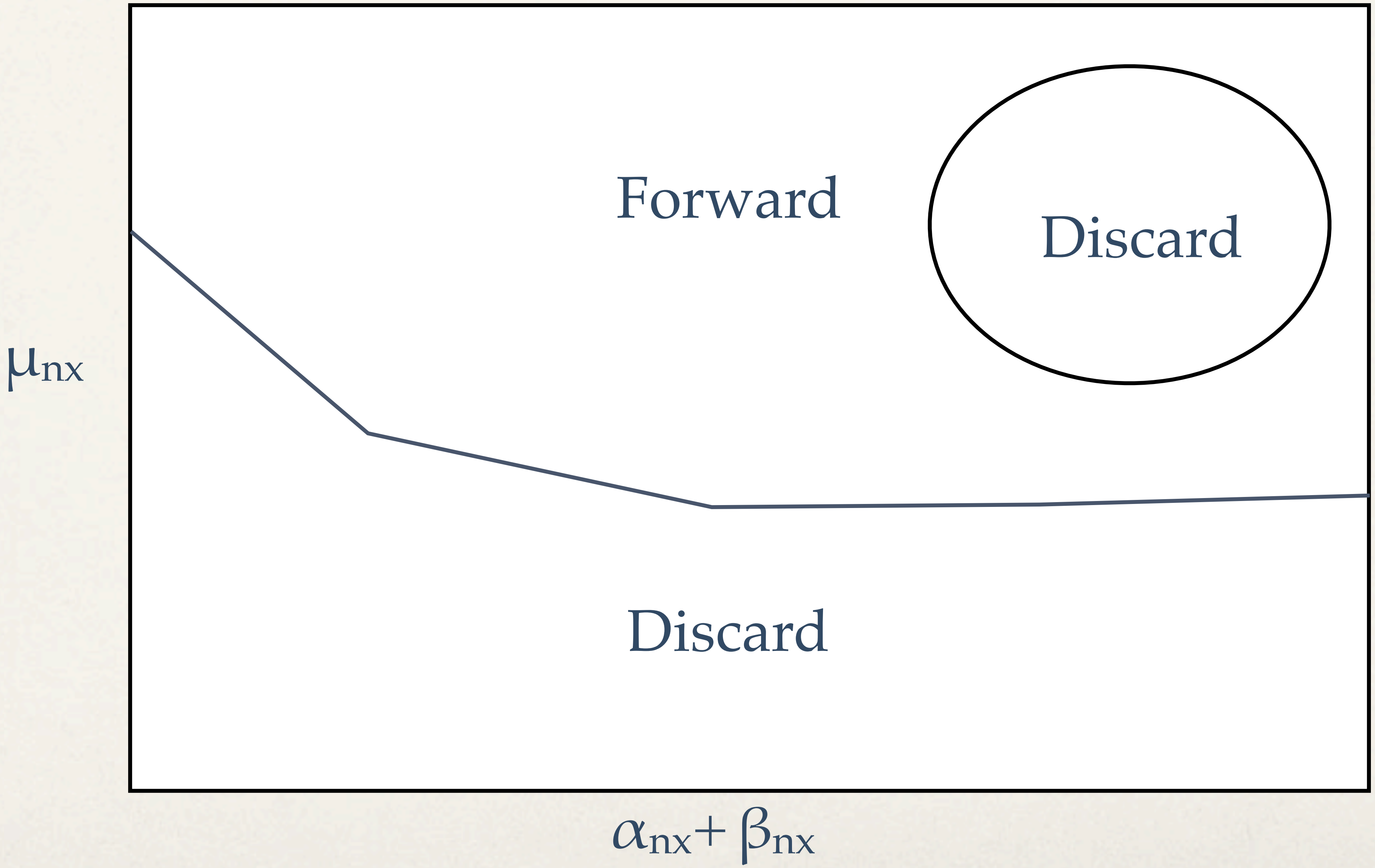
An algorithm partitions the space of posteriors into “Forward” and “Discard”

- Here is another possible algorithm:



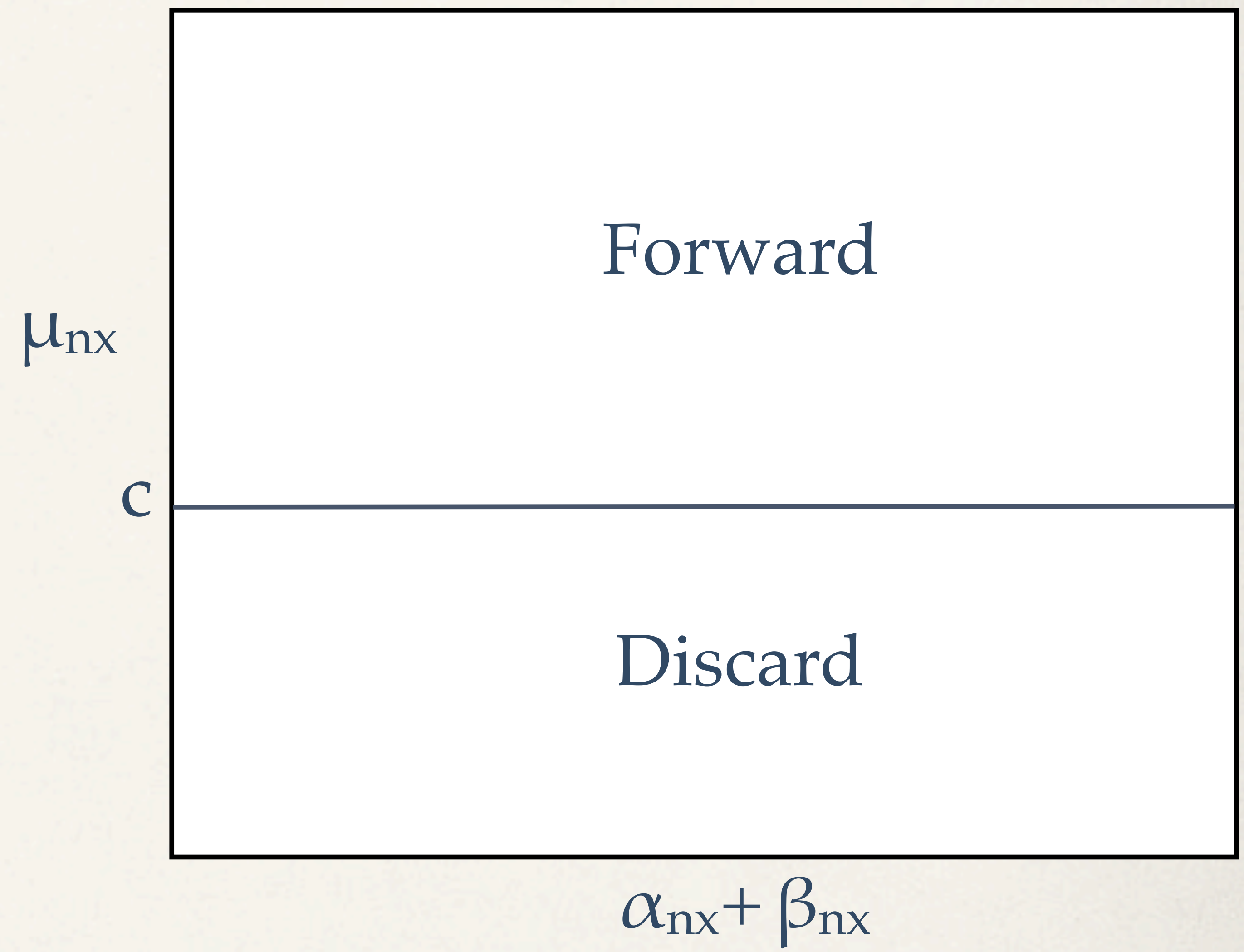
An algorithm partitions the space of posteriors into "Forward" and "Discard"

- Here is yet another possible algorithm:



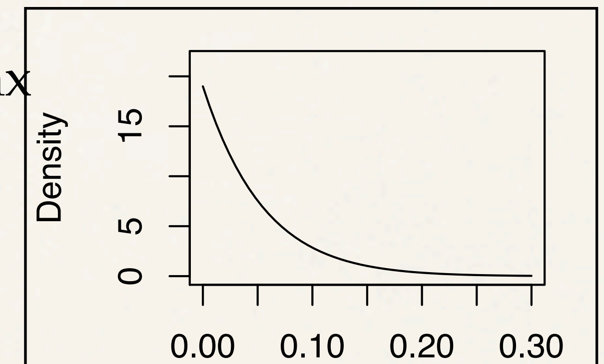
The myopic algorithm can be expressed in this way.

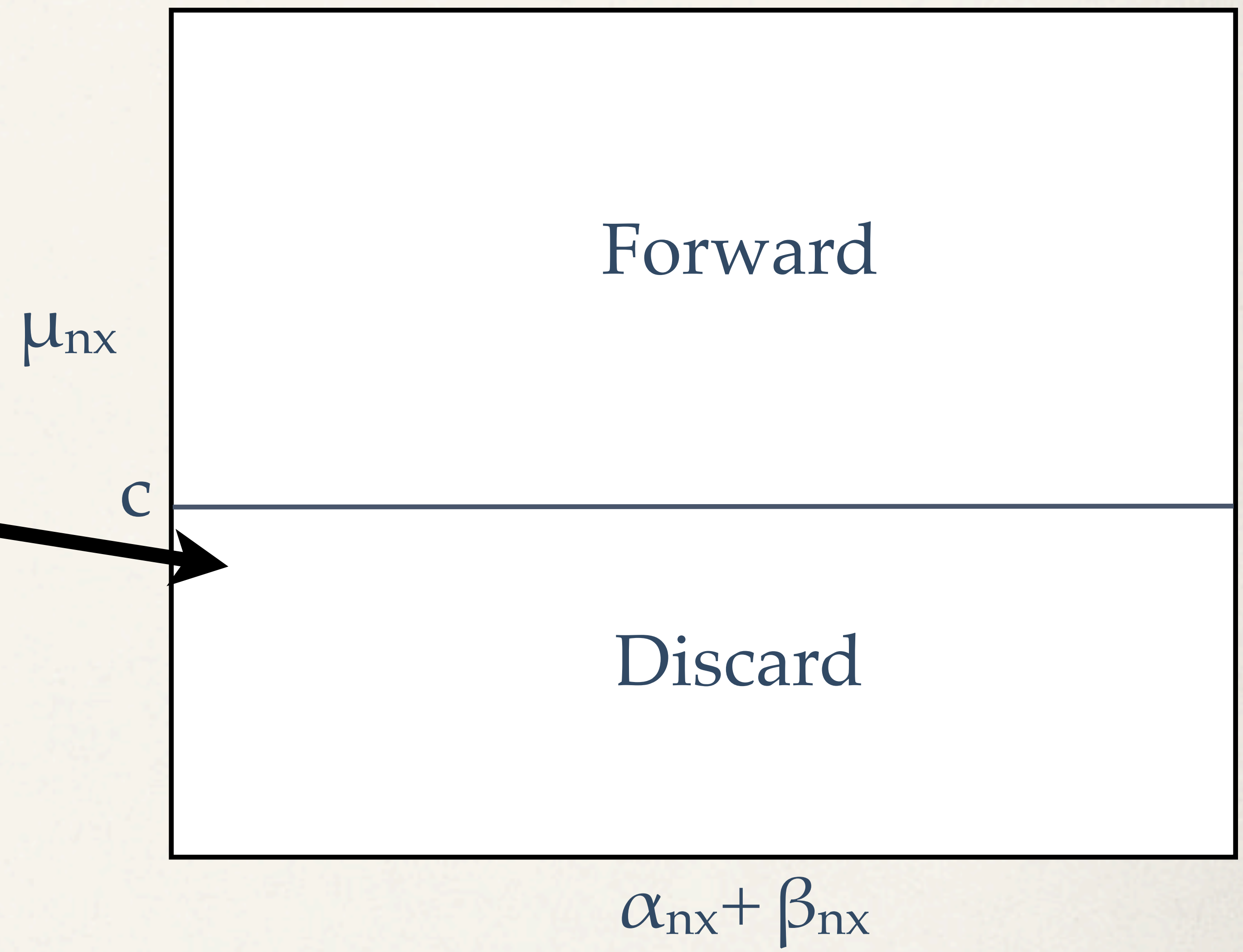
- The expected *immediate* payoff of forwarding is $E_n[\theta_x - c] = \mu_{nx} - c$
- The expected immediate payoff of discarding is 0.
- The rule that maximizes expected immediate reward is:
    - Forward if $\mu_{nx} > c$
    - Discard if not.



The myopic algorithm ignores the value of **exploring**

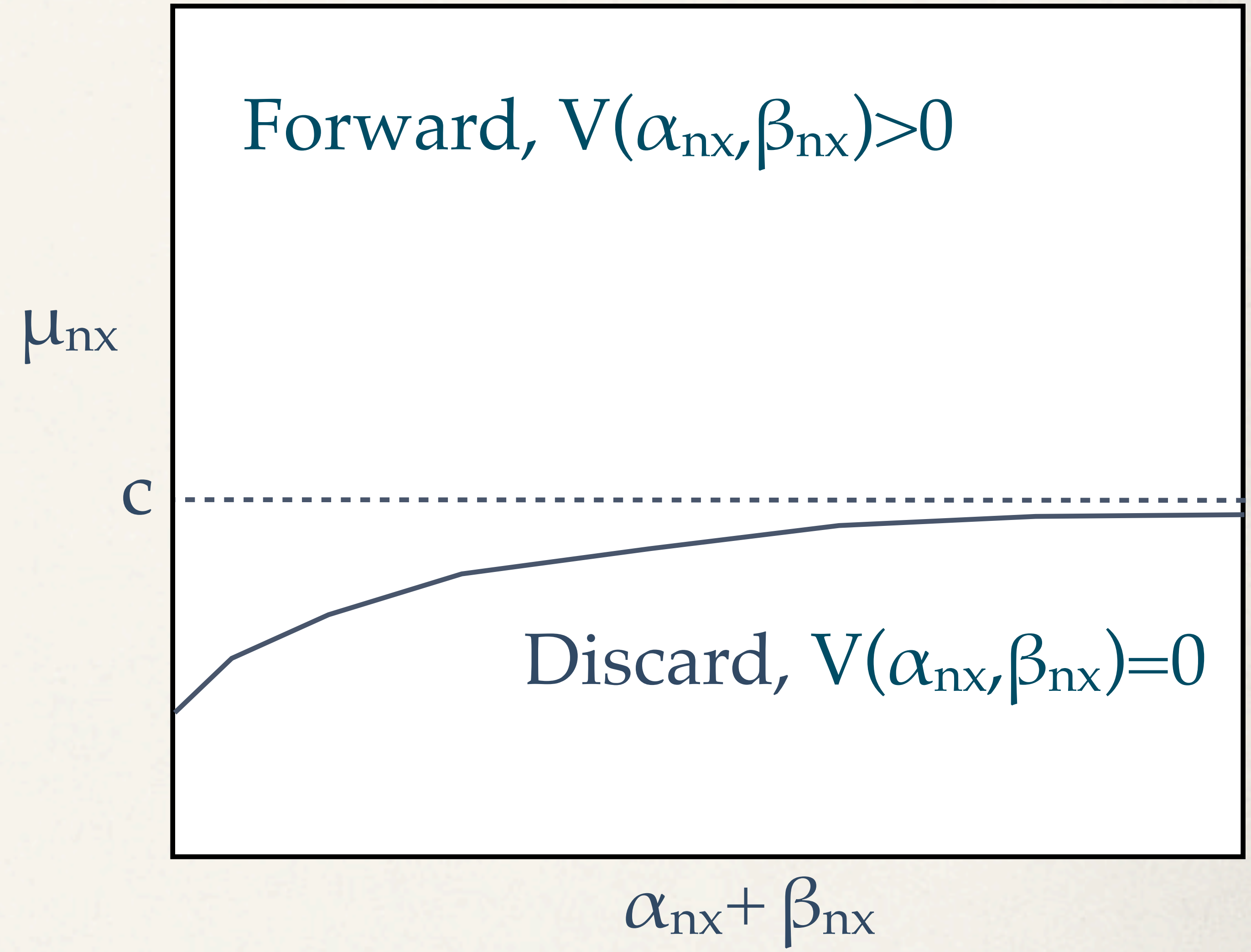
- If our current posterior has:
  - small $\alpha_{nx} + \beta_{nx}$
  - μ_{nx} close to c





- then it might be worth forwarding, just to learn more about θ_x.
- If it turns out θ_x >c, we can take advantage of this in future forwarding decisions.

We can compute the optimal algorithm through stochastic dynamic programming



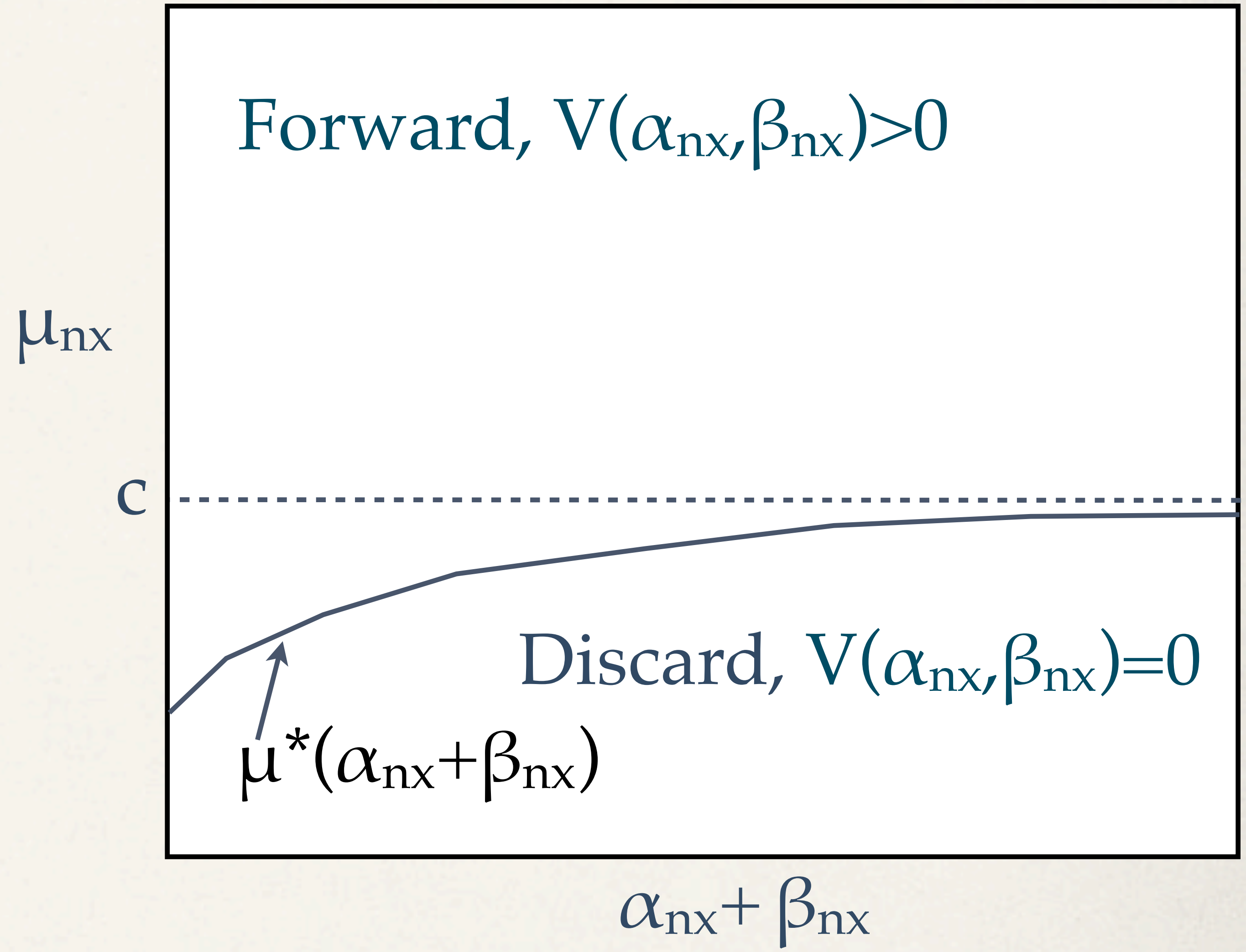
- Let $V(\alpha_{nx},\beta_{nx})$ be the expected future reward under the optimal policy, given n documents of history.
- V satisfies the dynamic programming recursion:

$$V(\alpha_{nx},\beta_{nx})=P(N_x>n)\max(0,\mu_{nx}-c+E_n[V(\alpha_{n+1,x},\beta_{n+1,x})])$$

The optimal algorithm trades **exploration vs. exploitation**

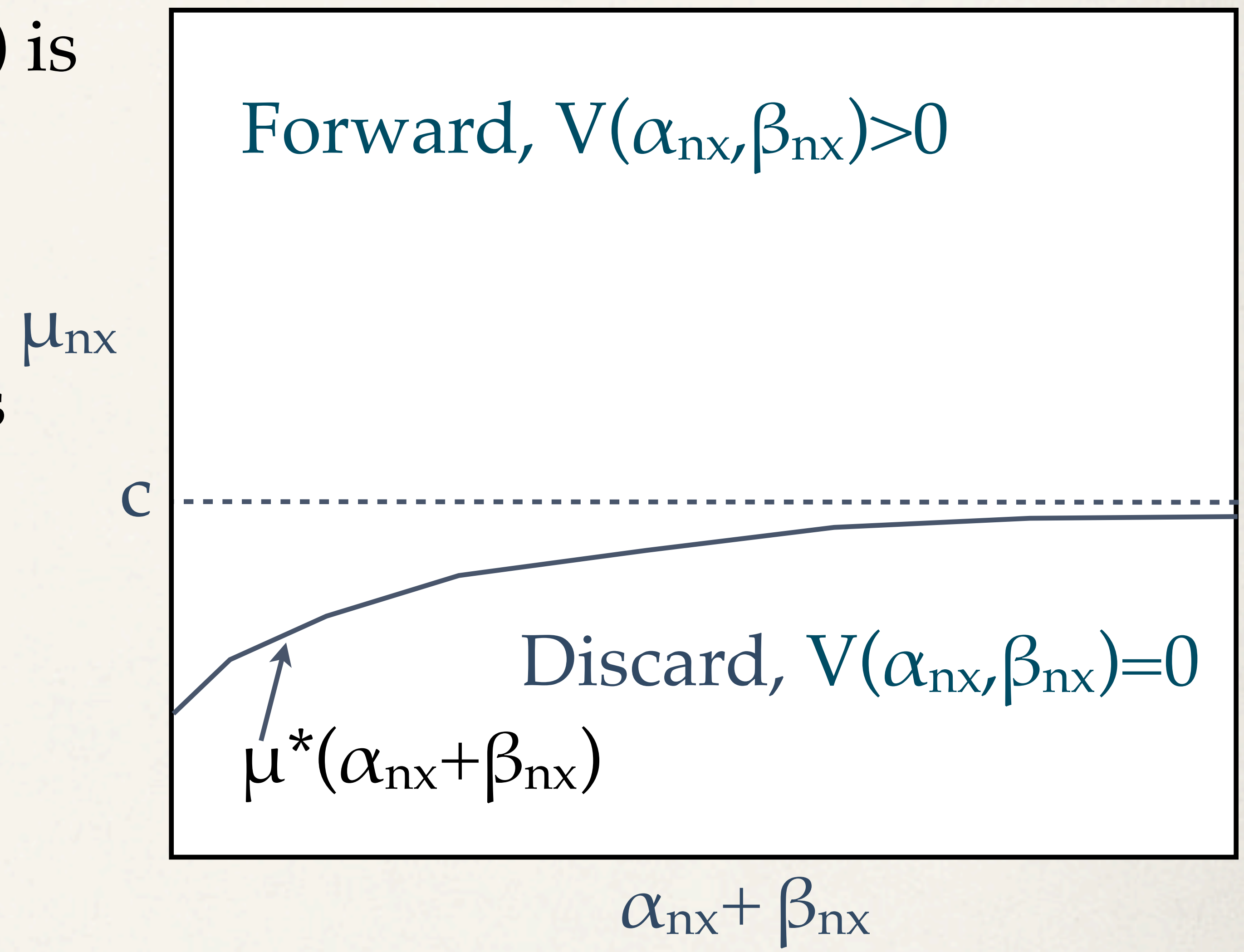
- **Theorem 1**: There exists a function $\mu^*(\alpha+\beta)$ such that it is optimal to forward when $\mu_{nx} \geq \mu^*(\alpha+\beta)$ and to discard otherwise.

- **Theorem 2**: $\mu^*(\alpha+\beta)$ has the following properties:
  - it is bounded above by c;
  - it is increasing in $\alpha+\beta$;
  - it and goes to c as $\alpha+\beta \rightarrow \infty$.

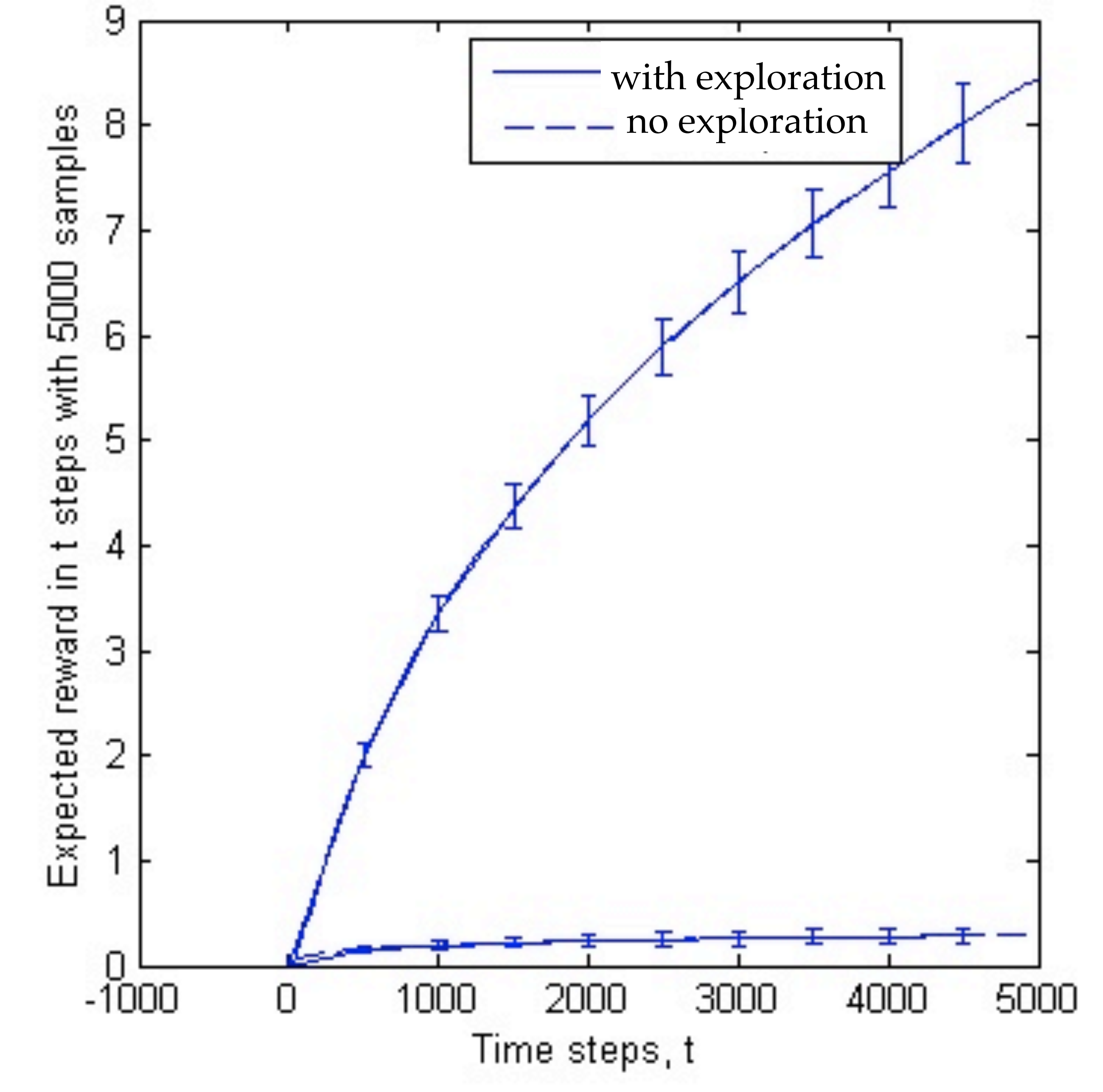


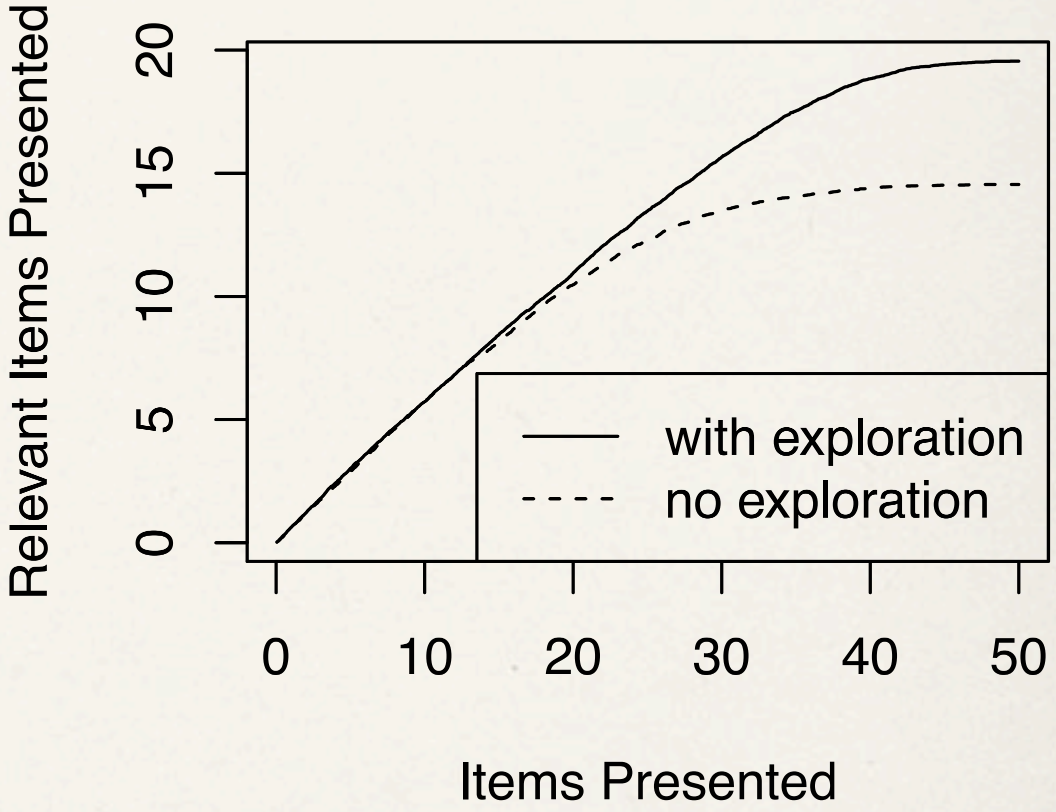
The optimal algorithm trades **exploration vs. exploitation**

- When $\alpha_{nx}+\beta_{nx}$ is small, $\mu^*(\alpha_{nx}+\beta_{nx})$ is much less than c, and we favor **exploration**.
- When $\alpha_{nx}+\beta_{nx}$ is big, $\mu^*(\alpha_{nx}+\beta_{nx})$ is close to c, and we favor **exploitation**.



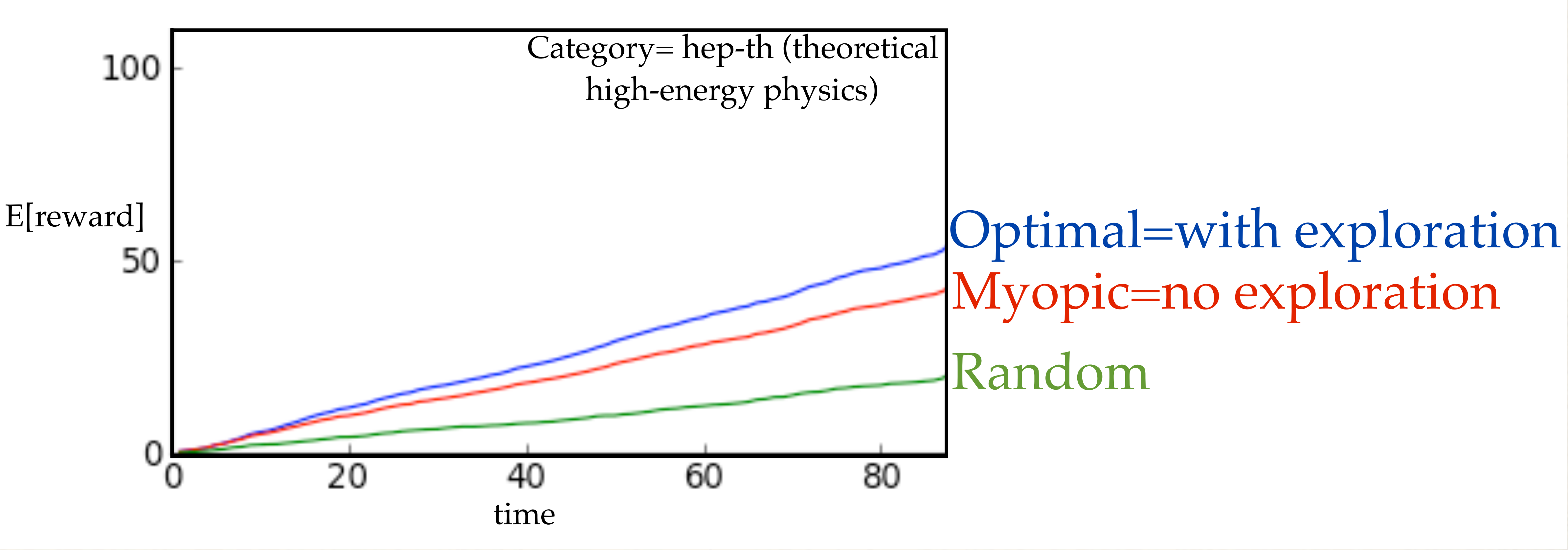
Optimal outperforms myopic (with simulated users)



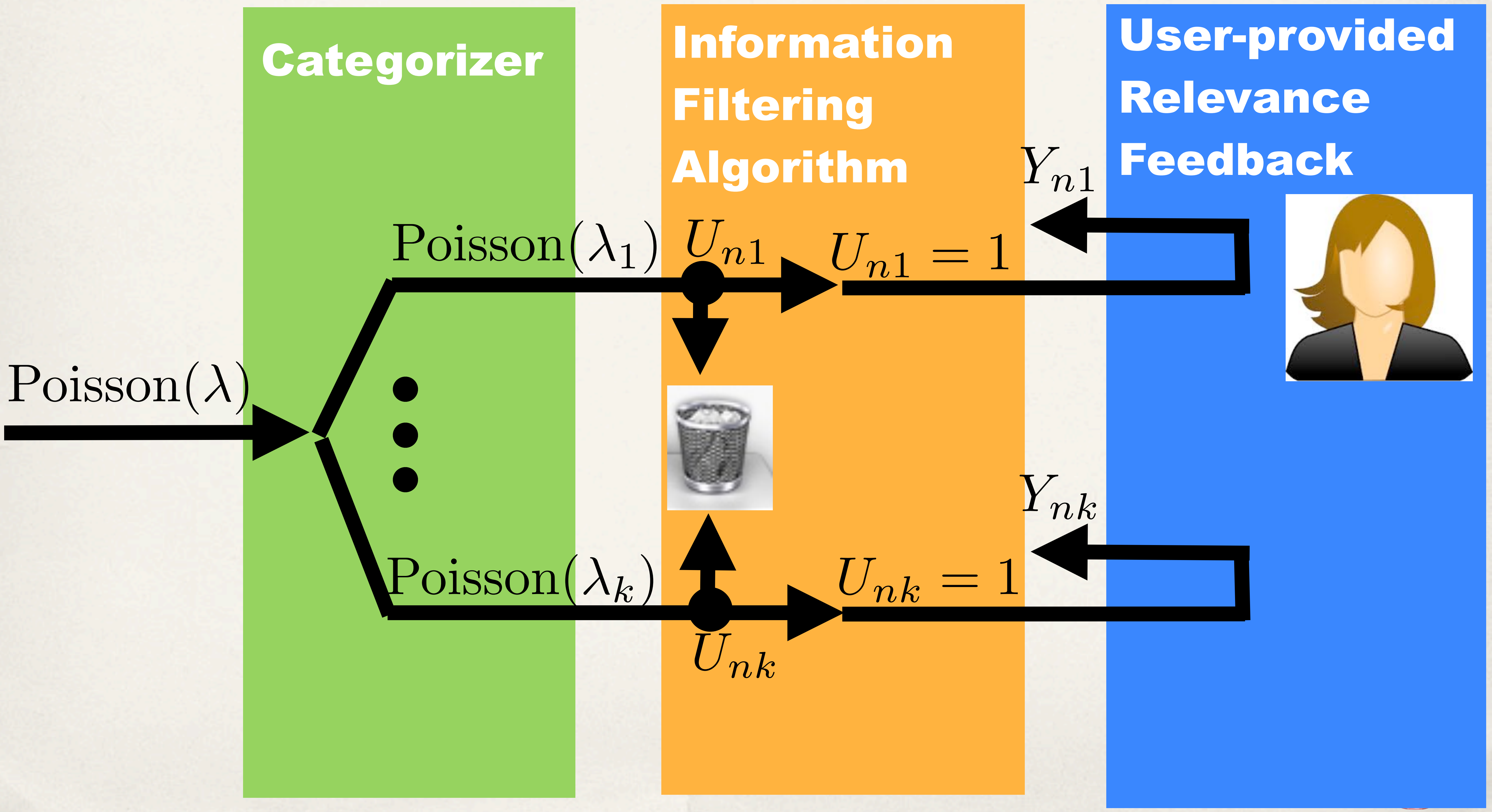


In these graphs,
"with exploration"=optimal
"no exploration"=myopic

Optimal outperforms myopic (in backtesting with historical data)



Combining single-category solutions solves the multi-category problem



Combining single-category solutions solves the multi-category problem

- We know the optimal forwarding/discarding strategy for a single category.
- To deal with multiple categories, simply apply this strategy independently to each individual category.
- The value of this optimal multi-category strategy is the sum of the values of the optimal single-category strategies:

$$\sup_{\pi} E^{\pi}\left[\sum_{x=1}^{k}\sum_{n=1}^{N_x} U_{nx}(Y_{nx}-c)\right] = \sum_{x=1}^{k}\sup_{\pi} E^{\pi}\left[\sum_{n=1}^{N_x} U_{nx}(Y_{nx}-c)\right]$$

Outline

- Categorizing items
- Mathematical Model
- Extensions & Tweaks

Extension #1: Periodic review

- In the arxiv, users do not respond instantaneously. Instead they visit arxiv periodically (once per day) to read papers.
- We allow papers to accumulate in a queue until the user arrives.
- When the user arrives, we decide which papers to forward/discard.
- The analysis is still tractable using stochastic dynamic programming.

Extension #2: Unknown costs

- In reality, we do not know the cost c for each forwarded document.
- To address this, we:
  - Compute c^* for each paper, which is the largest cost c such that we would be willing to forward this paper.
  - We present papers in a ranked list in decreasing order of c^*.
- Optimality analysis:
  - If we model the user as knowing his own c, and looking at all papers with $c^*>c$, then this algorithm **is optimal**.
  - If we model the user as looking at the top n papers in the list each time, this algorithm **is not optimal** in general, but we can obtain tractable upper and lower bounds.
  - If $n=1$, this algorithm **is optimal**, and is equivalent to the Gittins index policy for multi-armed bandits.

Extension #3: Time-varying user preferences

- User preferences change over time.
- Our Bayesian statistical model may be extended to allow θ_x to change over time.
- The analysis is still tractable.

Extension #4: Correlated prior distributions

- Our model assumed an independent prior on θ_x.
- In the data, a user's strong interest in one category (e.g., theoretical high-energy physics) may make a strong interest in another category more likely (e.g., experimental high-energy physics).
- We can model this with a correlated prior on θ_x.
- The dynamic program is no longer tractable, but we can compute $\mu^*(\alpha_{nx}+\beta_{nx})$ using independence, but update our posterior using a correlated prior.

Conclusion

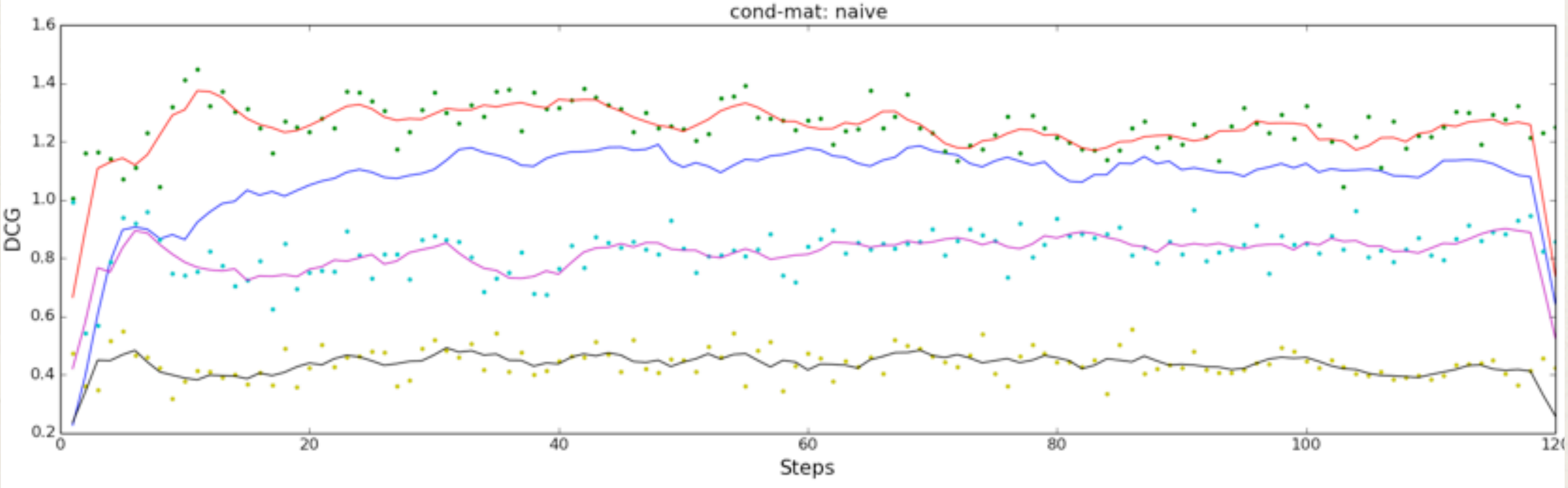
- We have presented a mathematical model that captures the exploration vs. exploitation tradeoff in information filtering.
- If the posterior mean is just a bit below c, and the number of samples is low, the optimal algorithm forwards, while the myopic algorithm does not.
- We are deploying an algorithm based on this analysis to my.arxiv.org

Thanks to my collaborators!

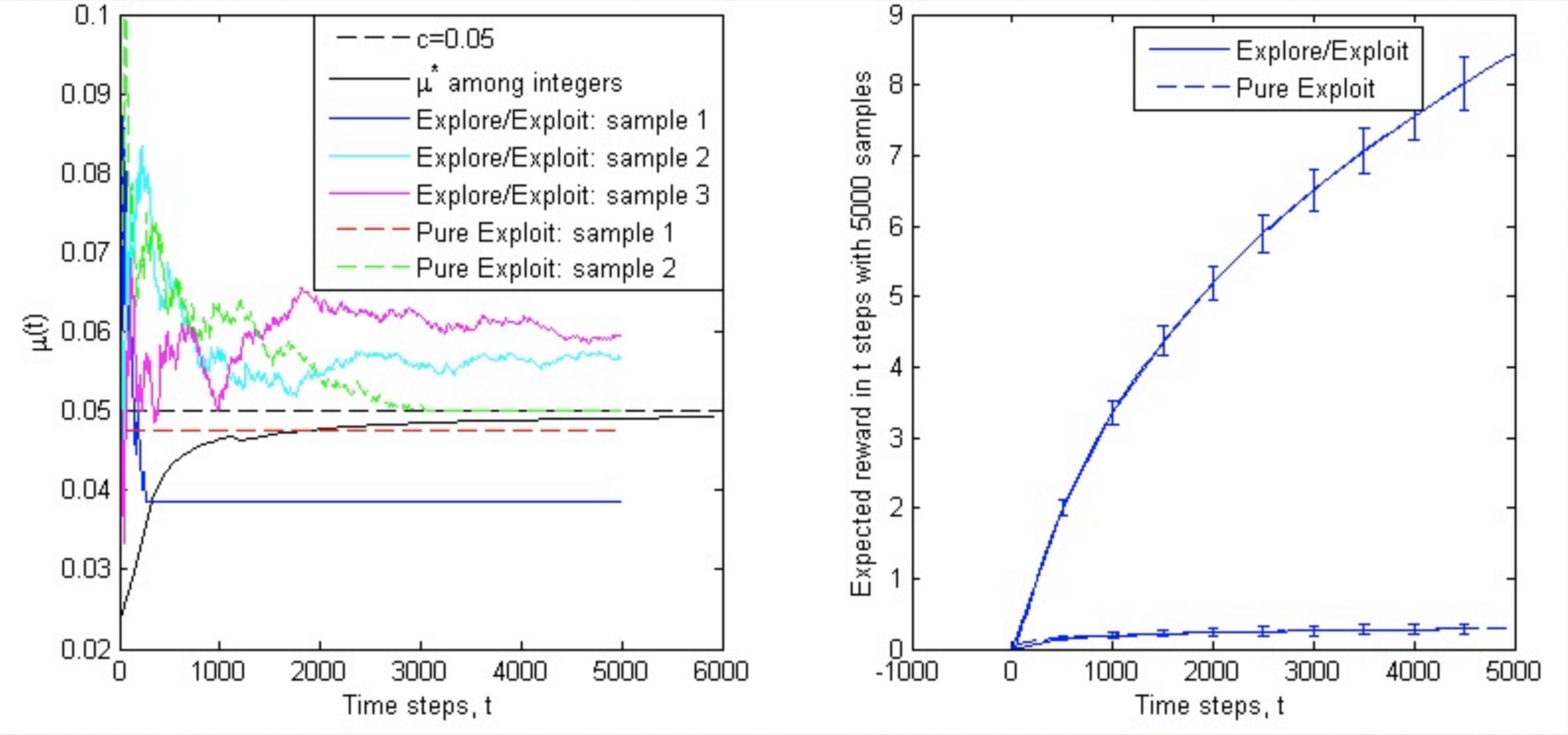
- This project is part of a larger collaboration on recommender systems for the arxiv, with faculty & students in CS, Operations Research, and Information Science at Cornell, Princeton, & Rutgers.
  - Paul Ginsparg, Thorsten Joachims, Xiaoting Zhao, Darlin Alberto, Karthik Raman, Ziyu Fan, Akilesh Potti (**Cornell**)
  - Paul Kantor & Vladimir Menkov (**Rutgers**)
  - Dave Blei & Laurent Charlin (**Princeton**)

Thanks for your attention!

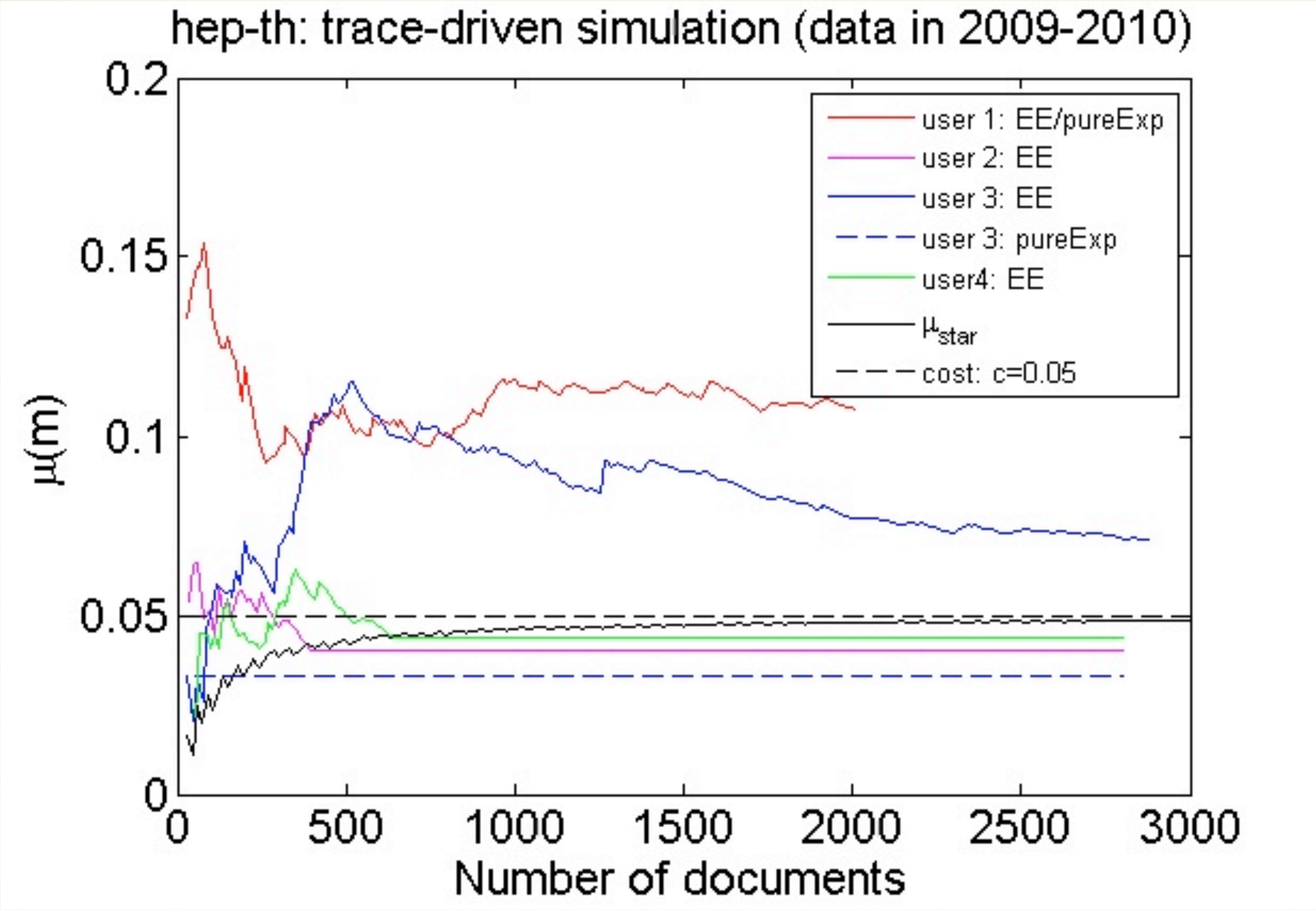
- Any questions?

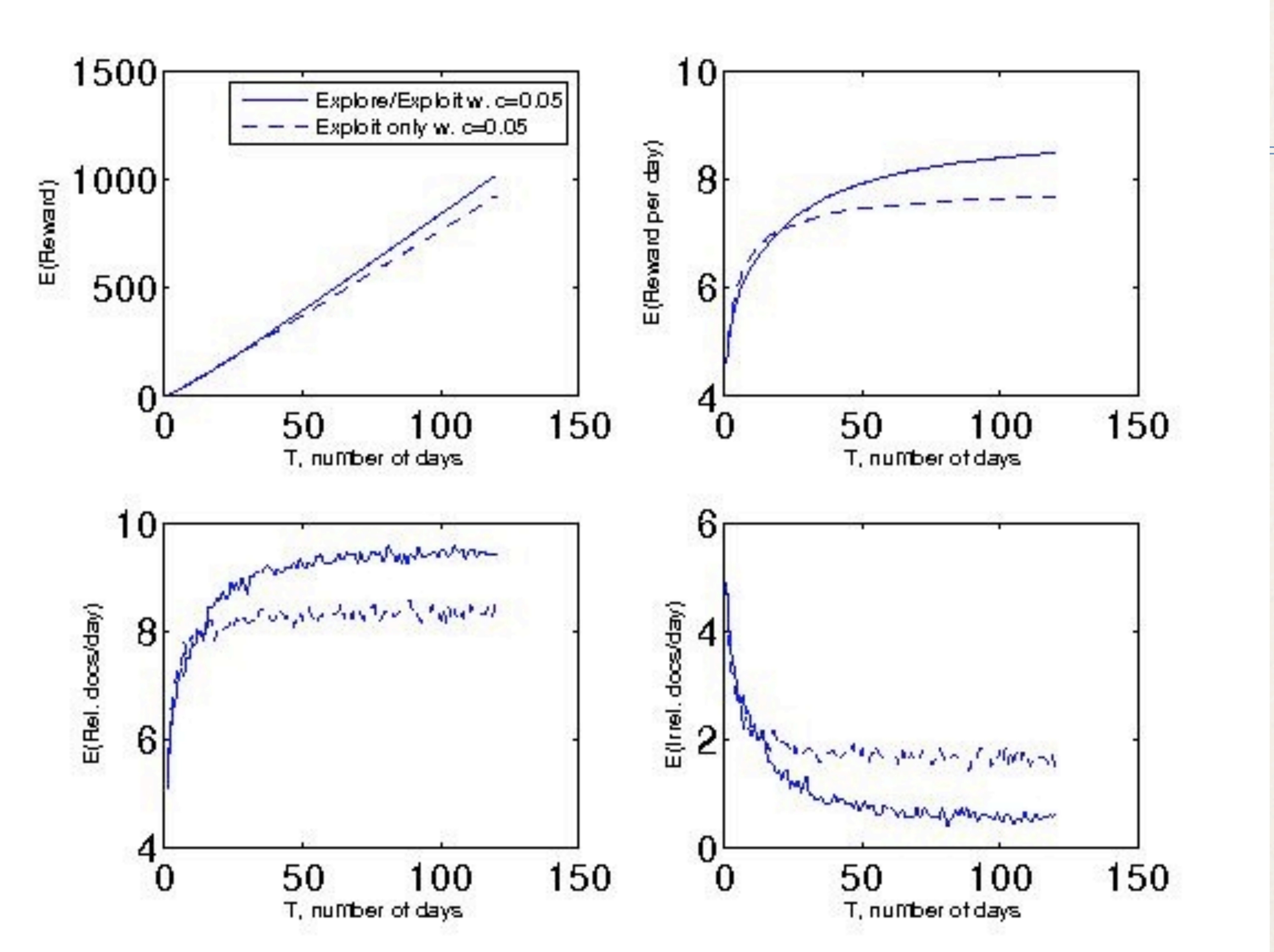


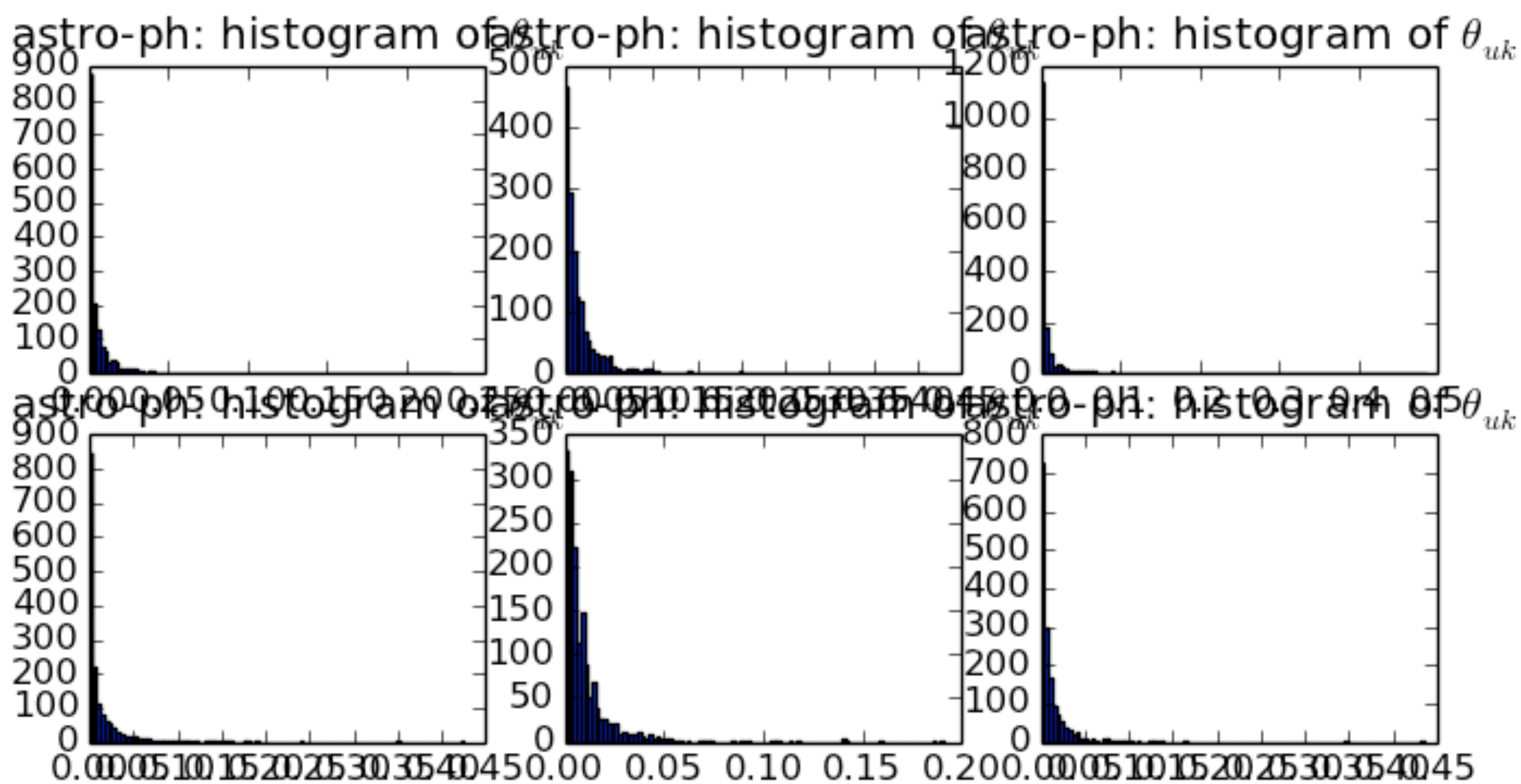
cond-mat: naive
DCG
Steps
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
20
40
60
80
100
120



c=0.05
μ* among integers
Explore/Exploit: sample 1
Explore/Exploit: sample 2
Explore/Exploit: sample 3
Pure Exploit: sample 1
Pure Exploit: sample 2
μ(t)
Time steps, t
Explore/Exploit
Pure Exploit
Expected reward in t steps with 5000 samples
Time steps, t


hep-th: trace-driven simulation (data in 2009-2010)
0.2
0.15
0.1
0.05
0
μ(m)
0
500
1000
1500
2000
2500
3000
Number of documents
user 1: EE/pureExp
user 2: EE
user 3: EE
user 3: pureExp
user4: EE
μstar
cost: c=0.05


Explore/Exploit w. c=0.05
Exploit only w. c=0.05
E(Reward)
E(Reward per day)
E(Rel. docs/day)
E(Irrel. docs/day)
T, number of days
T, number of days
T, number of days
T, number of days


astro-ph: histogram of θ_{uk}
0.00 0.05 0.10 0.15 0.20
0.0 0.1 0.2 0.3 0.4 0.5