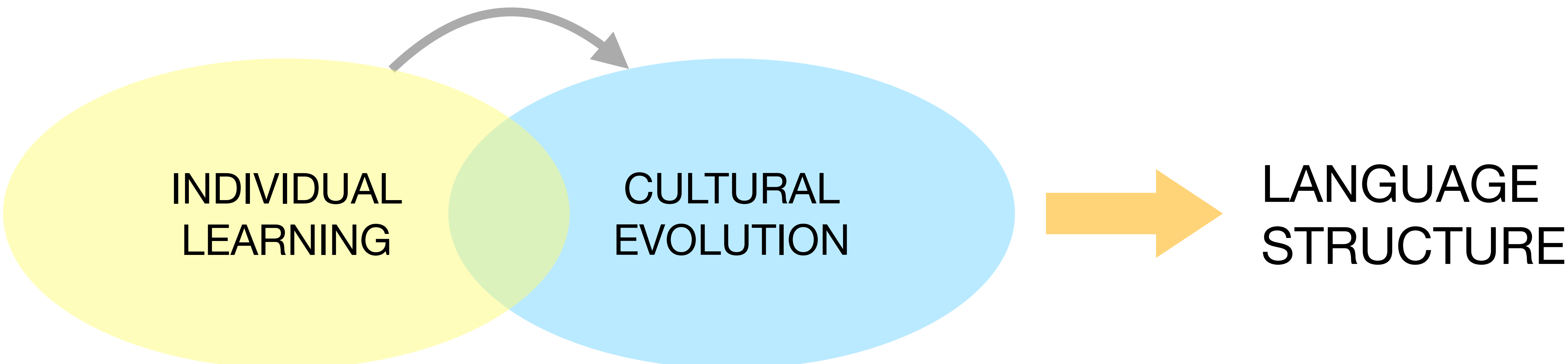
Simulating Language
8: Innateness and culture

Simon Kirby

simon.kirby@ed.ac.uk

Note to self: remember to start the recording!

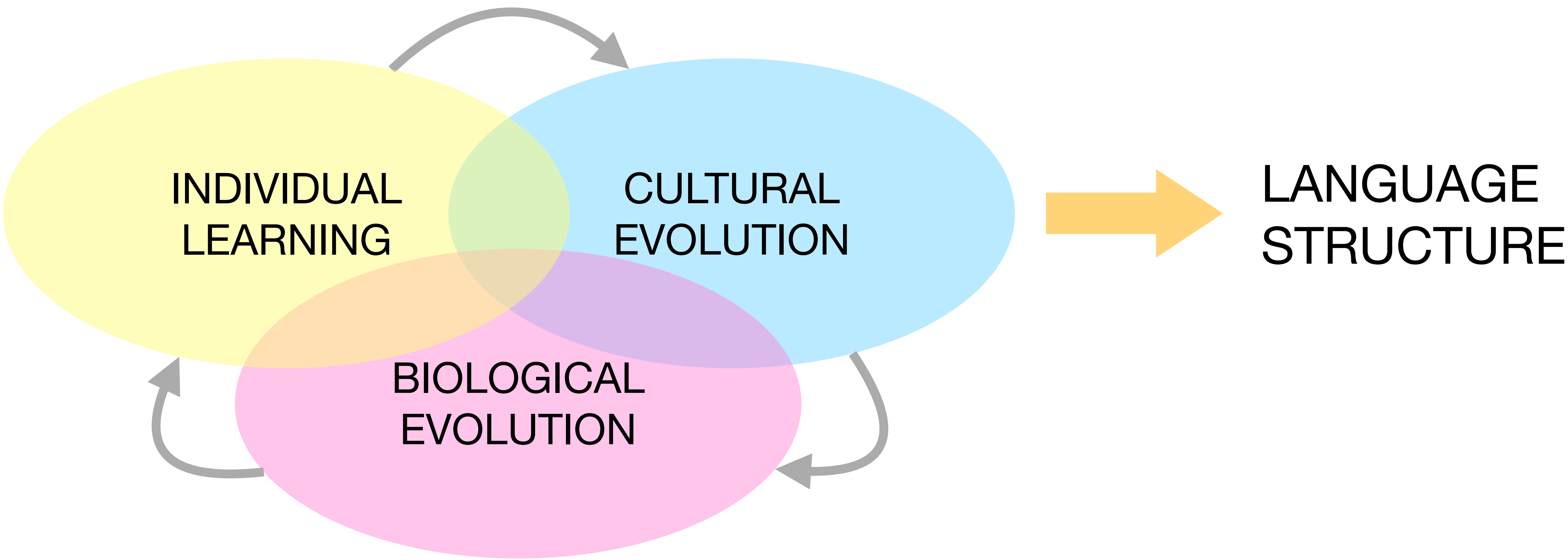
The story so far…



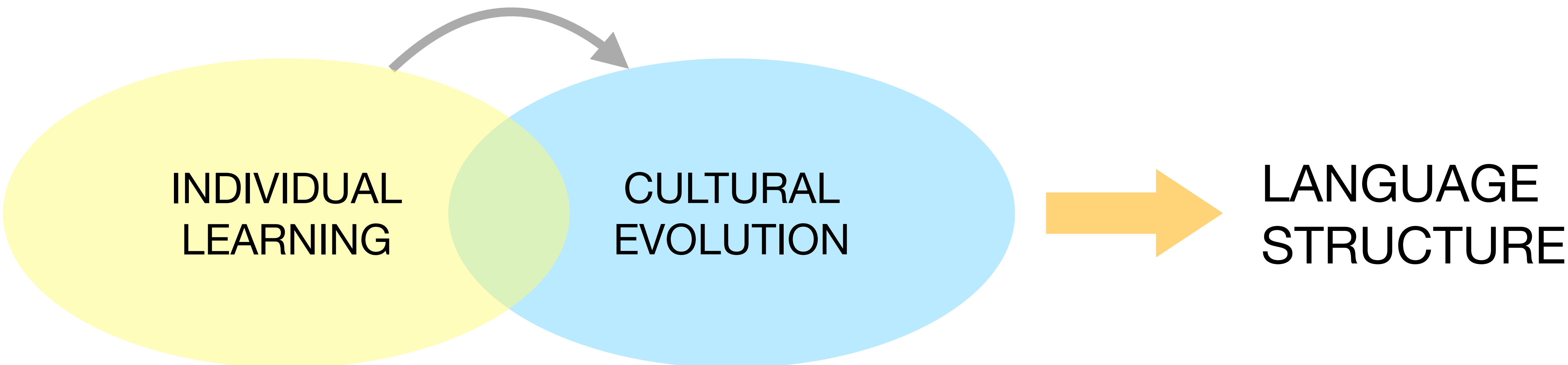
Prior bias in learners + Bottleneck in iterated learning → Universal properties of language

Where does this come from?

Where we're heading next…



But before we get there…



Prior bias in learners + $\boxed{\text{Bottleneck in iterated learning}}$ → Universal properties of language

What *exactly* does this do?

Remember the results from the compositionality models?

- We uncovered the importance of the *bottleneck* on cultural transmission
- It drives the evolution of structure because only structured languages can be stably transmitted through a bottleneck (without a bottleneck, language could stay holistic)
- This is a case of adaptation for learnability by a culturally evolving language
    - Although, in addition, other aspects of the bottleneck like communicative pressures can limit the evolution of language towards the simplest (degenerate) solution
- But it would be nice to understand how all these things interact in a bit more detail…

Thorough analysis of Iterated Bayesian Learning (Griffiths & Kalish 2007)

- Try out different models of language, different bottlenecks, different amounts of noise

- See how the process of cultural transmission takes the prior bias of the learner and gives rise to the actual resulting patterns of language

- What would you predict, based on the models you have seen so far?

- **The types of languages we see should:**

  A. be completely unconstrained by the biases of language learners

  B. reflect the biases of language learners, but in an interestingly complex way (e.g. effect of bottleneck etc. on outcome)

  C. directly reflect the biases of language learners and nothing more

Thorough analysis of Iterated Bayesian Learning (Griffiths & Kalish 2007)

- Try out different models of language, different bottlenecks, different amounts of noise

- See how the process of cultural transmission takes the prior bias of the learner and gives rise to the actual resulting patterns of language

- Their result:

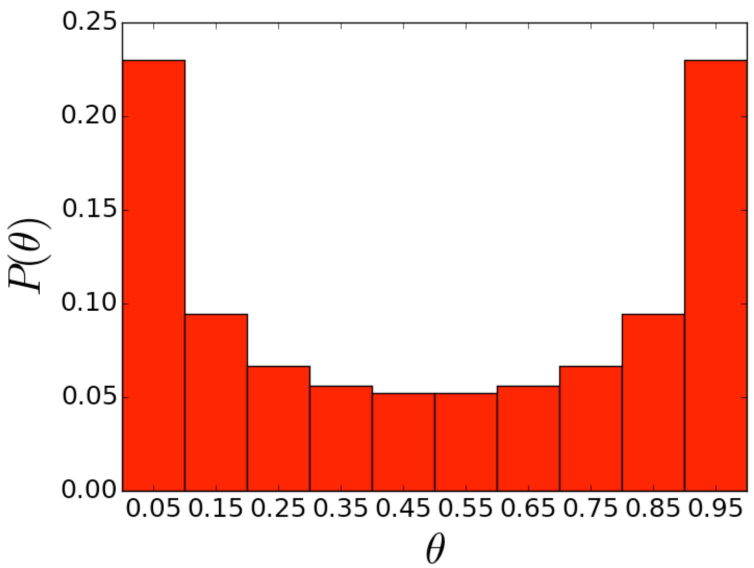
  Bottleneck does nothing
  Noise does nothing
  Details of language model do nothing

- Given enough time, the end result of cultural evolution **always reflects the prior bias and nothing else**

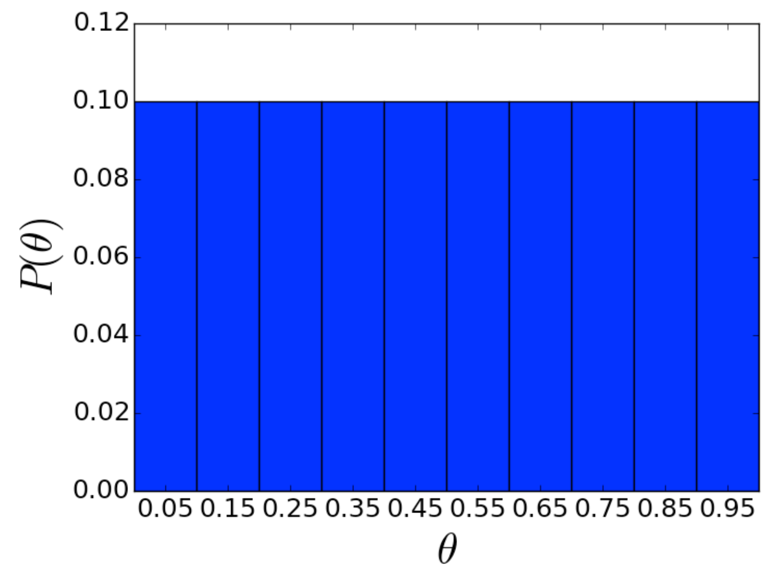
You have already seen this result

- Cast your mind back to the iterated beta-binomial model, with learners estimating frequencies of two competing linguistic variants

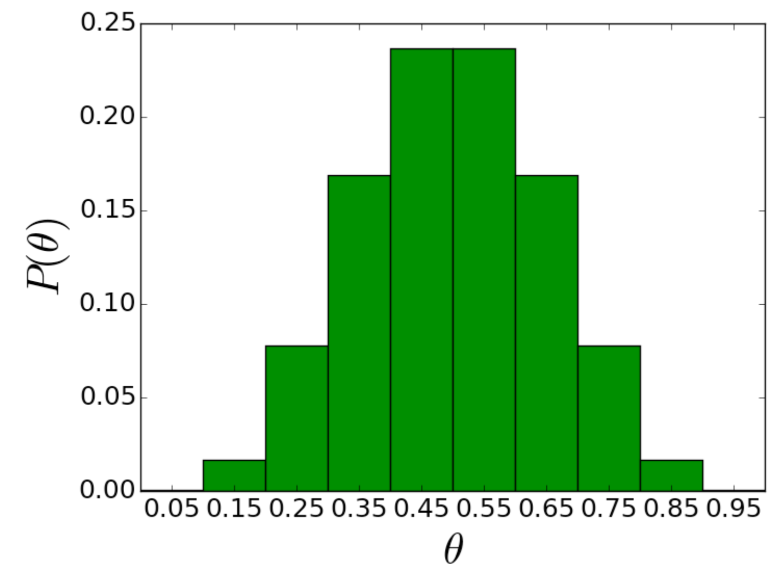
α = 0.1

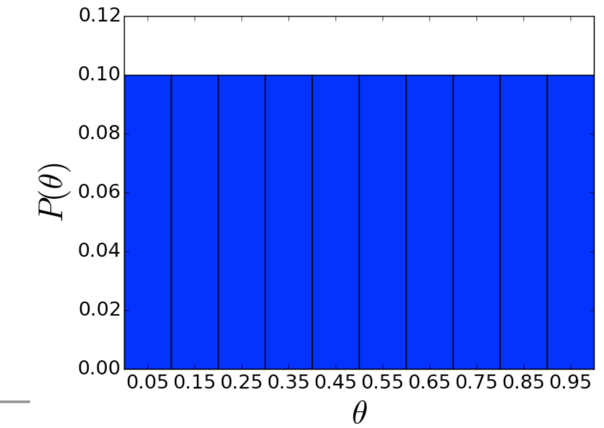


α = 1

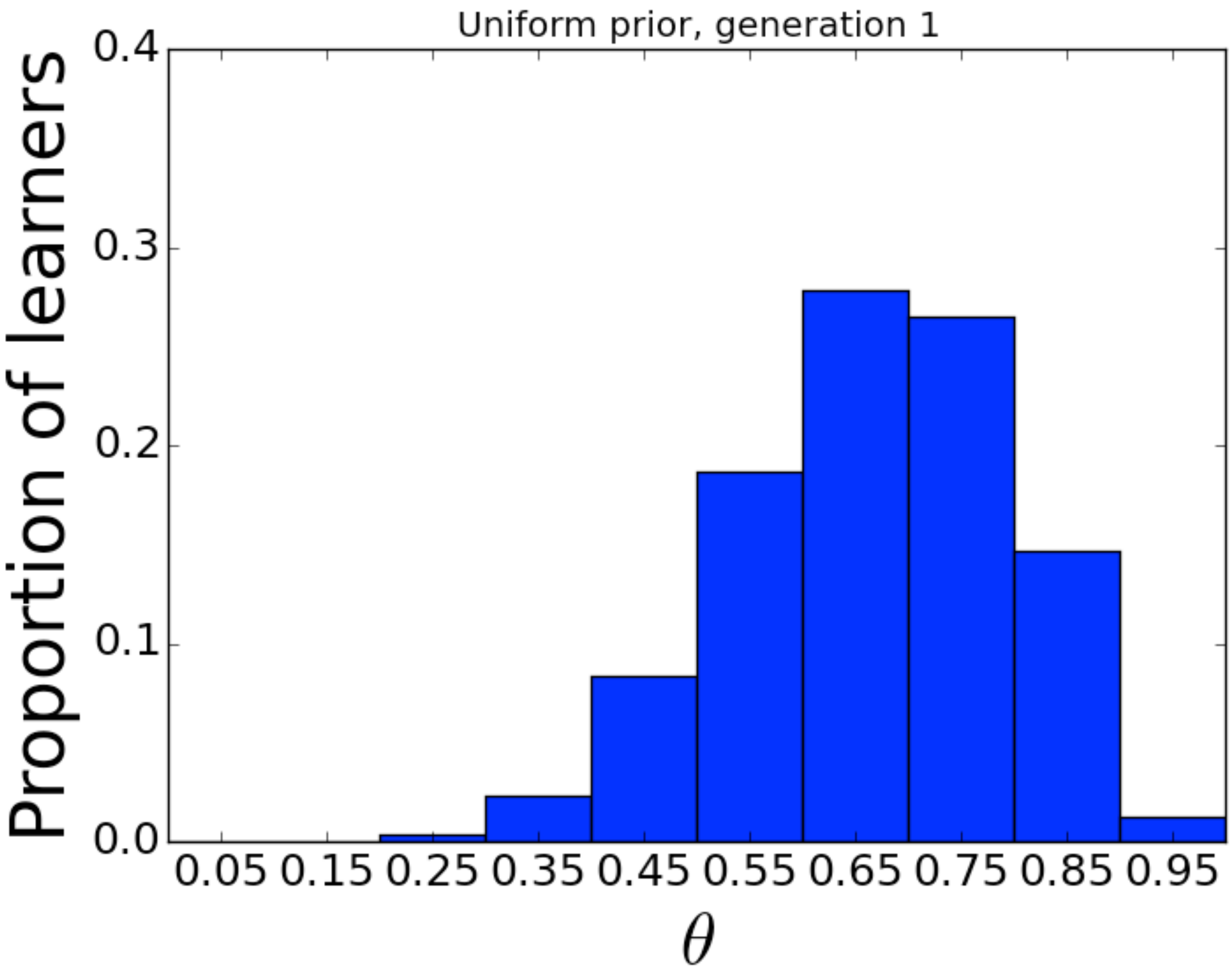


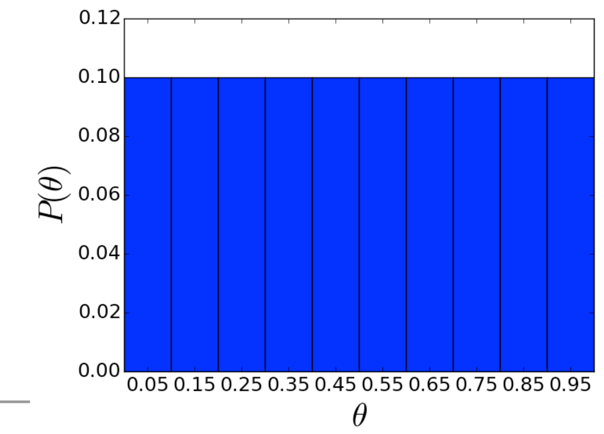
α = 5



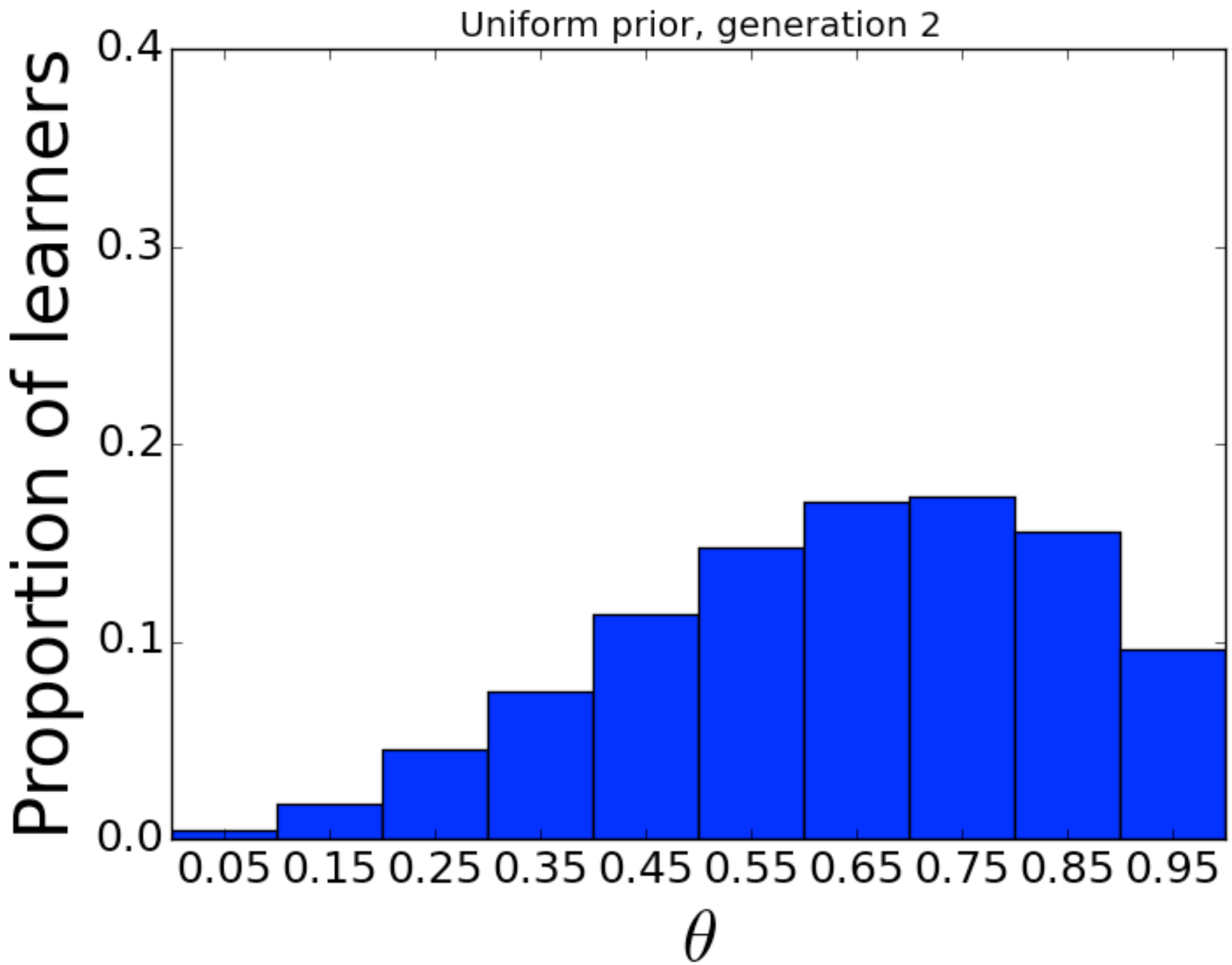


Watching the prior reveal itself

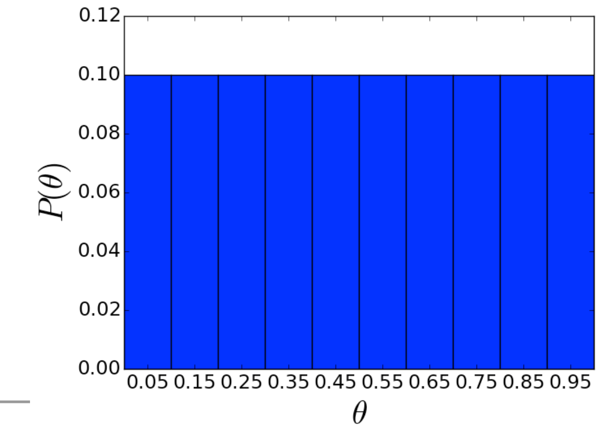


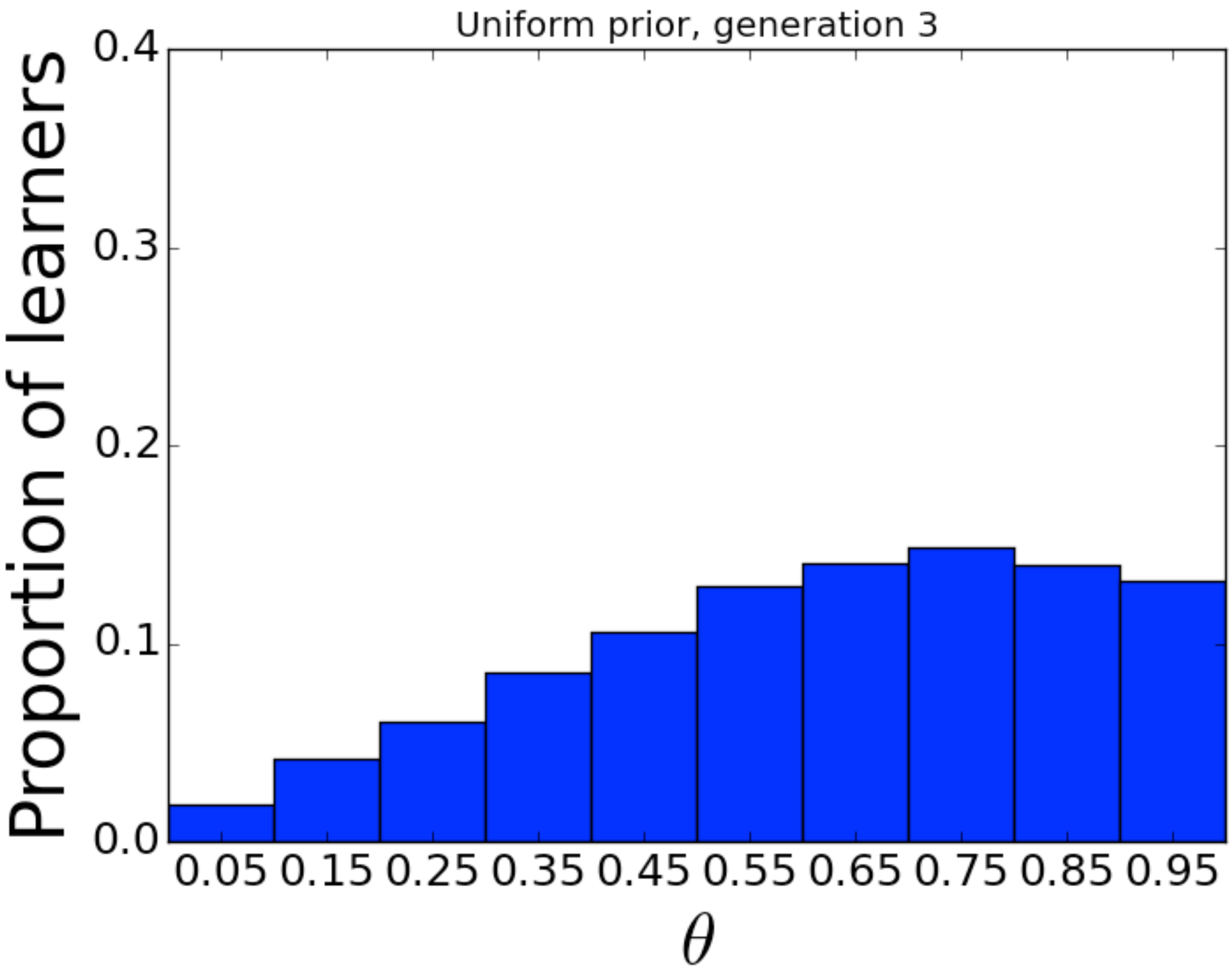


Watching the prior reveal itself

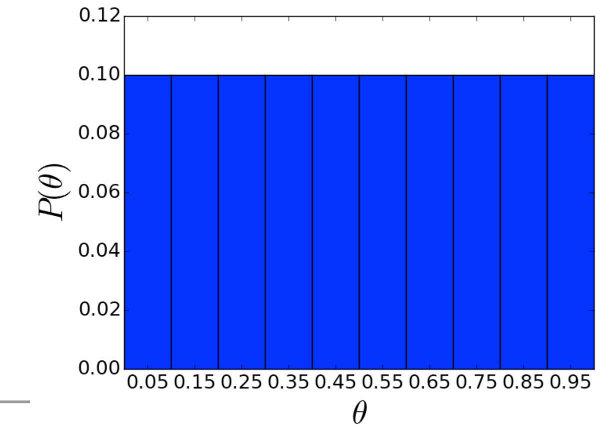


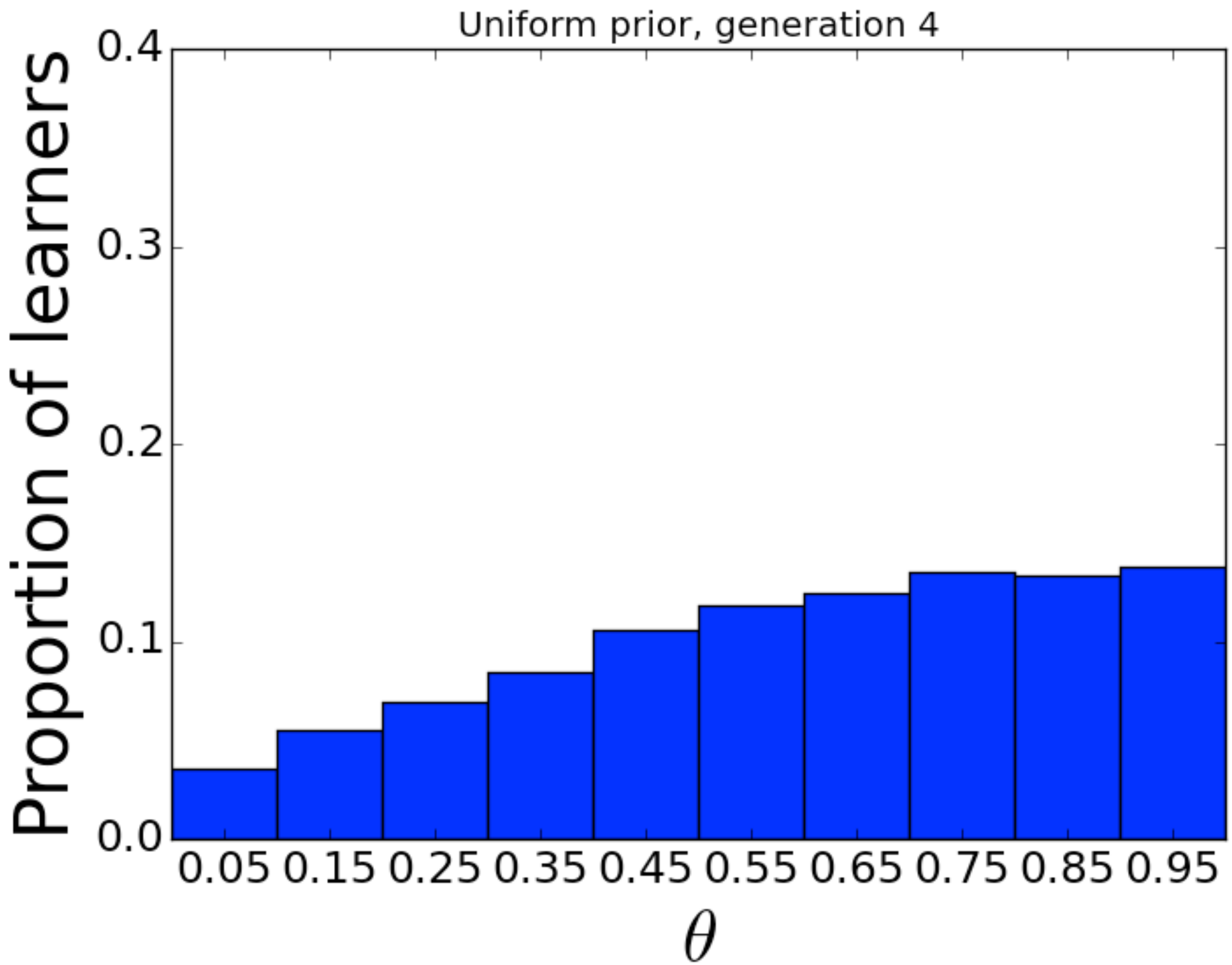
Watching the prior reveal itself



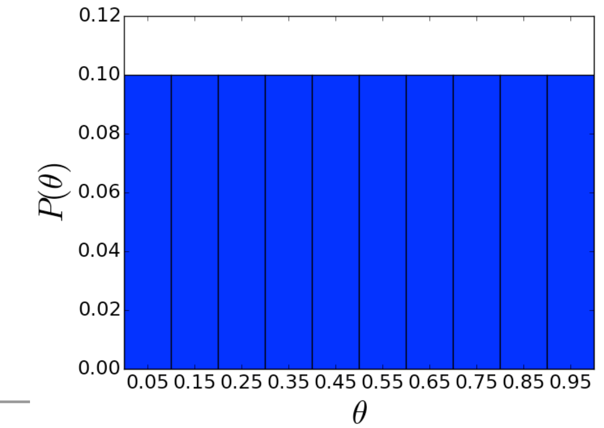


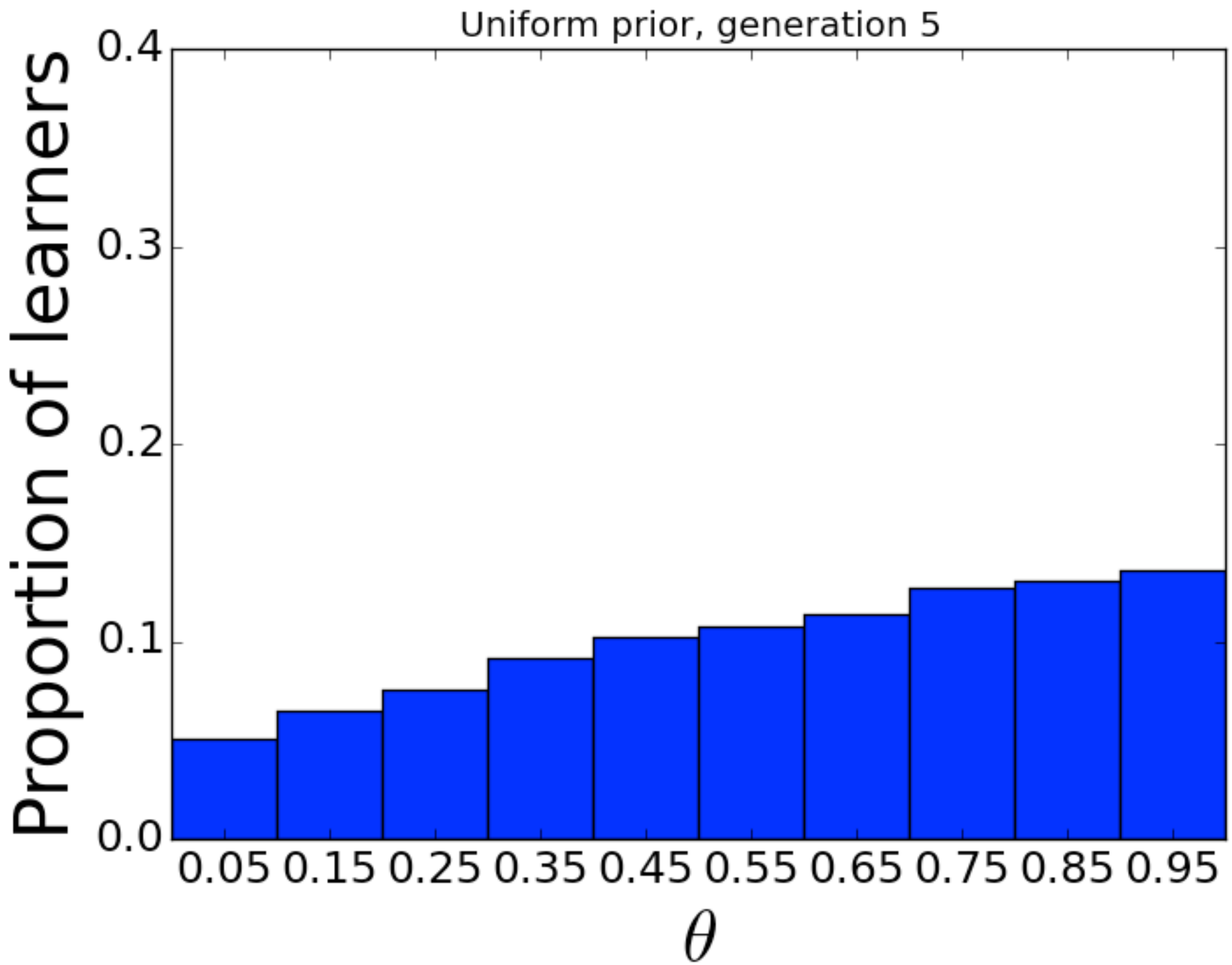
Watching the prior reveal itself



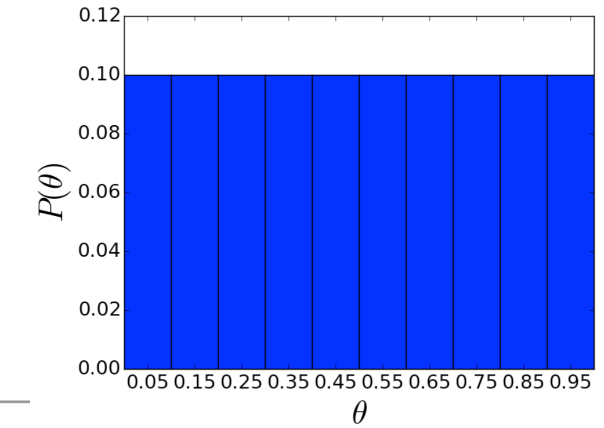


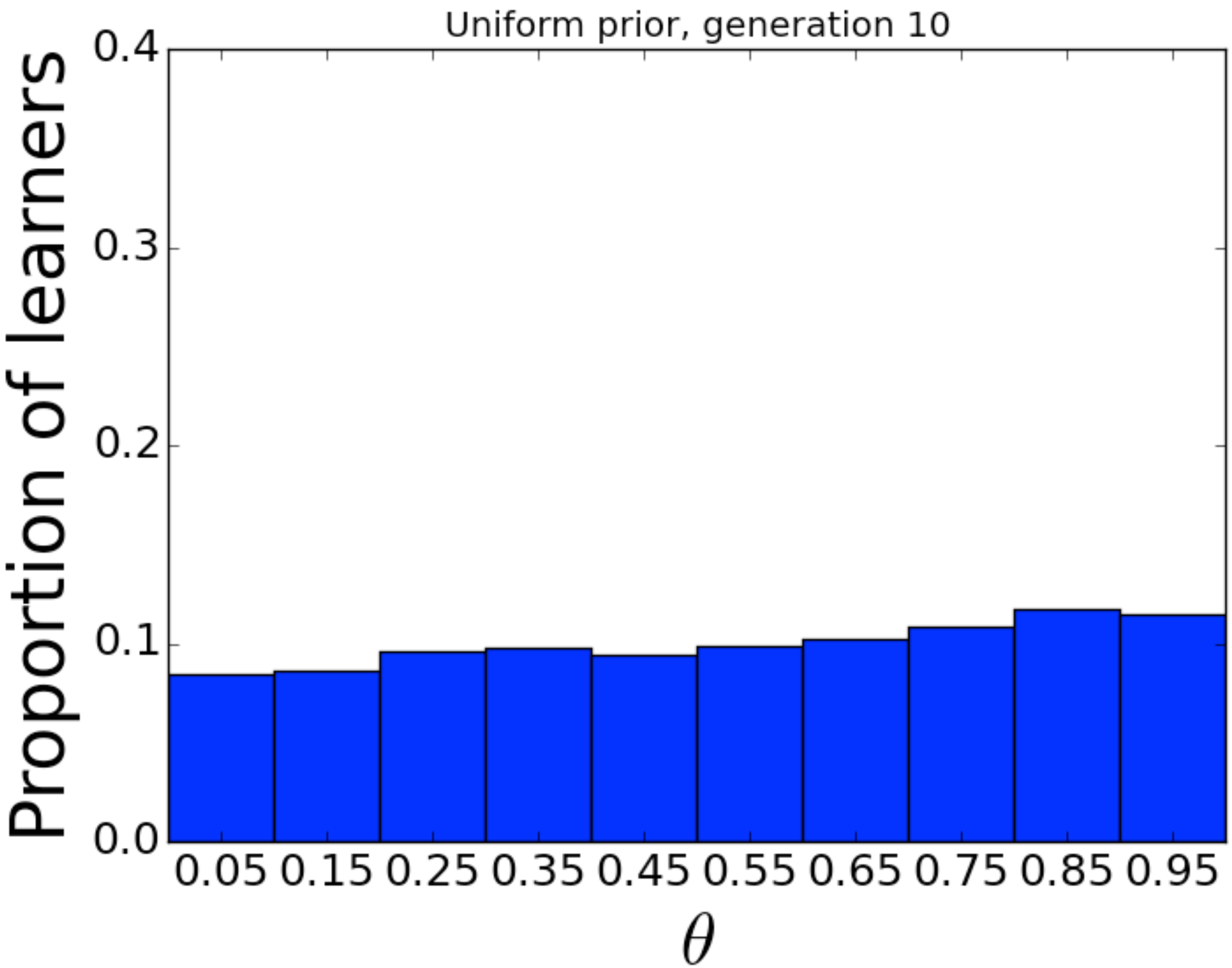
Watching the prior reveal itself



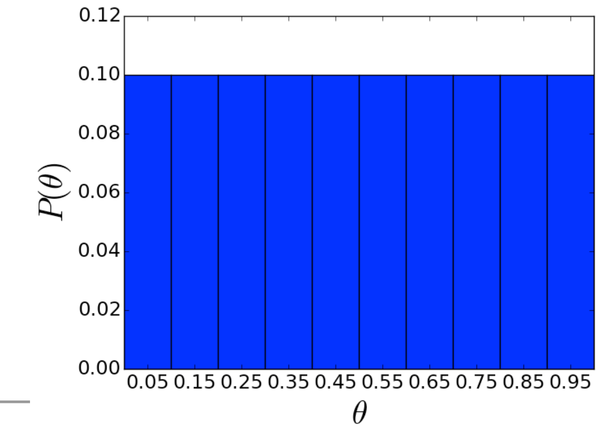


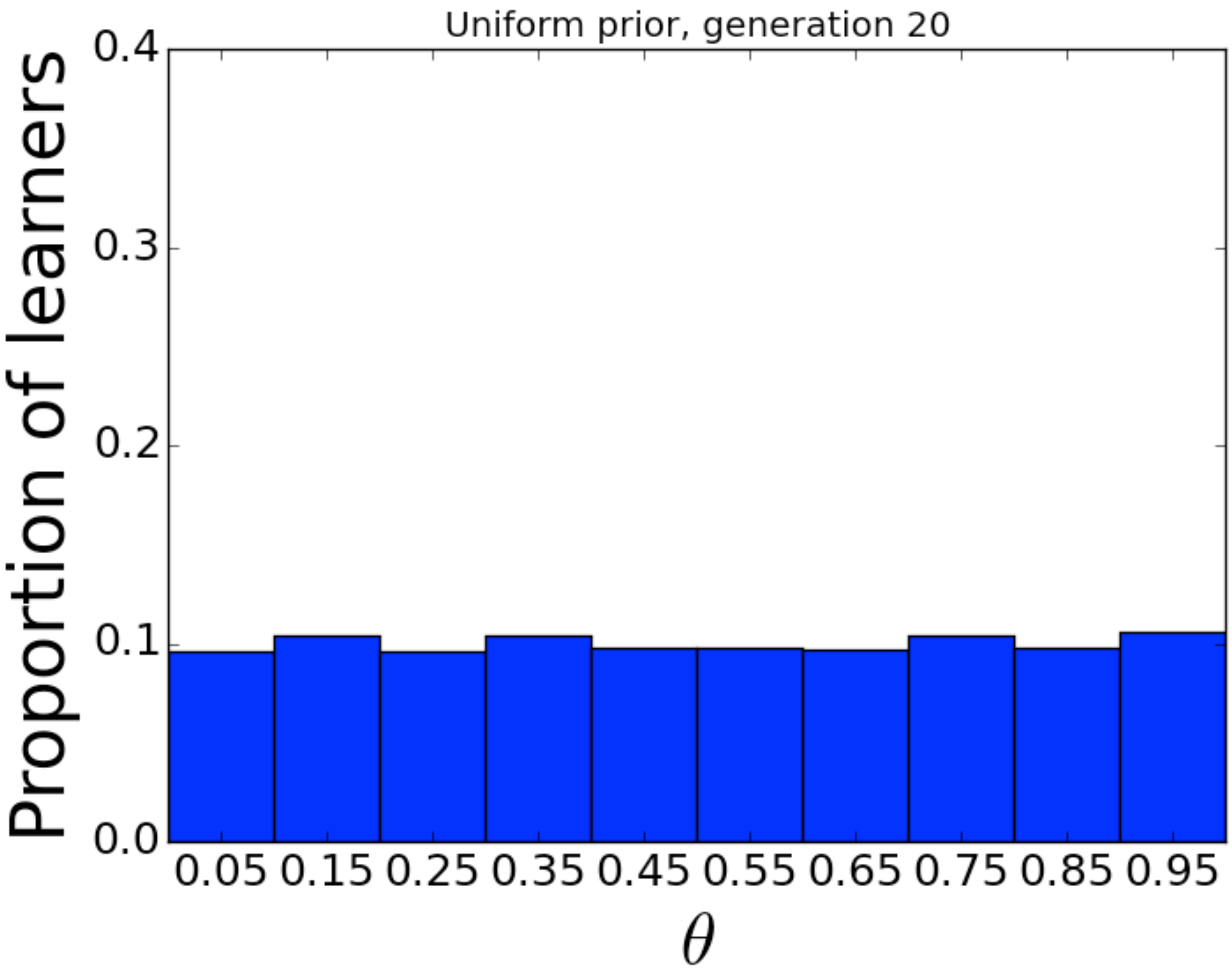
Watching the prior reveal itself



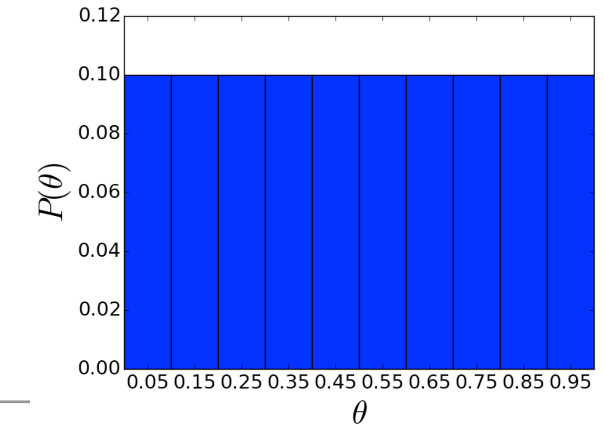


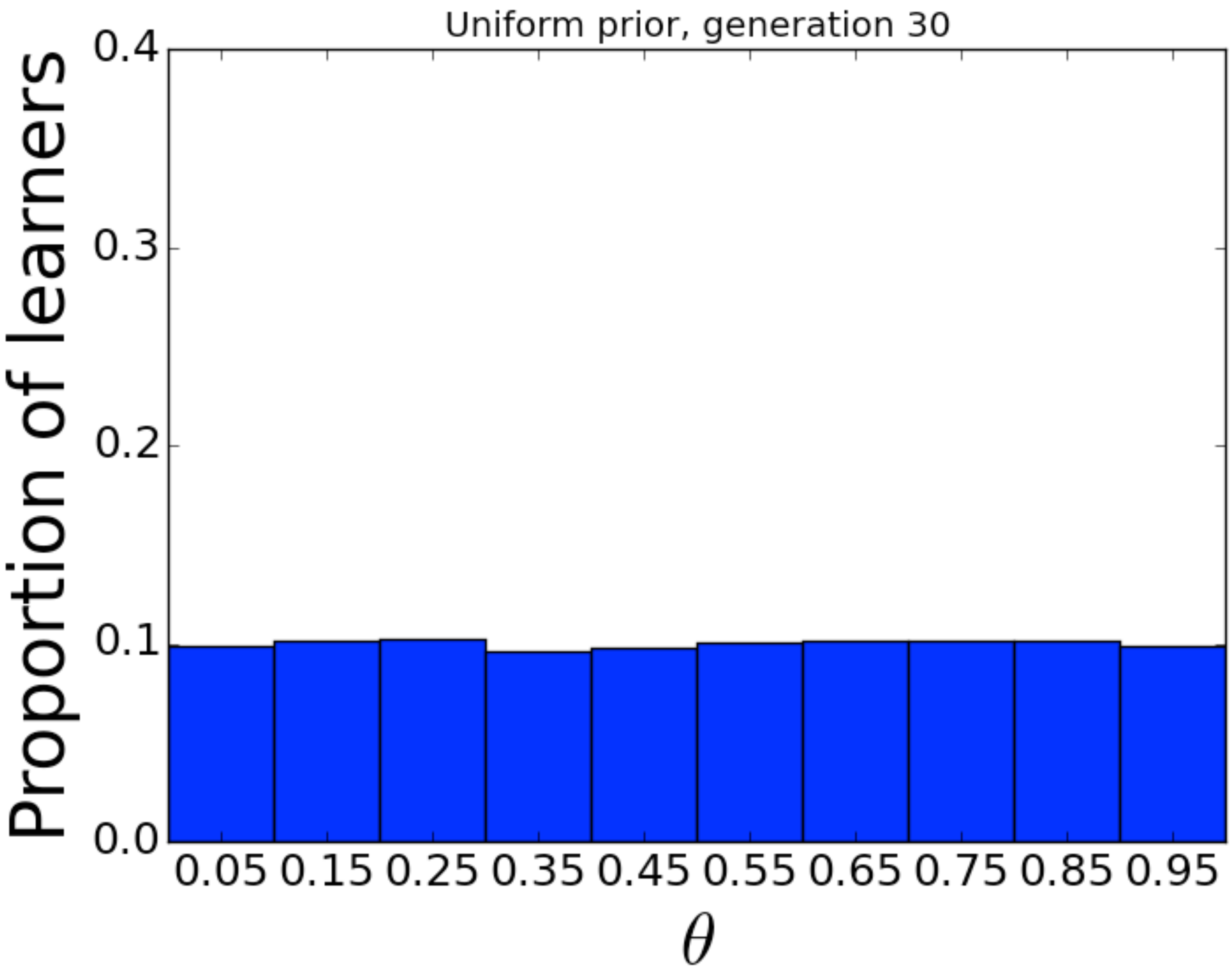
Watching the prior reveal itself



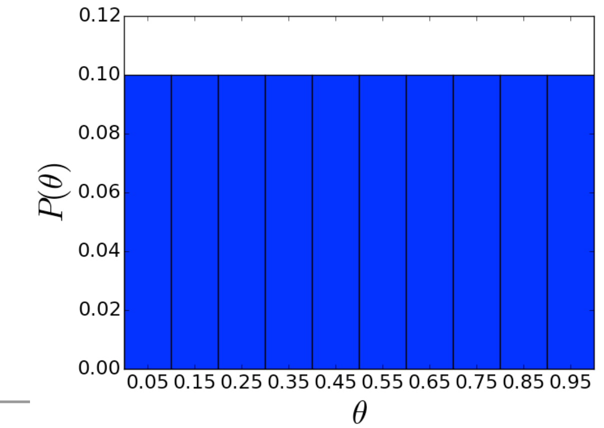


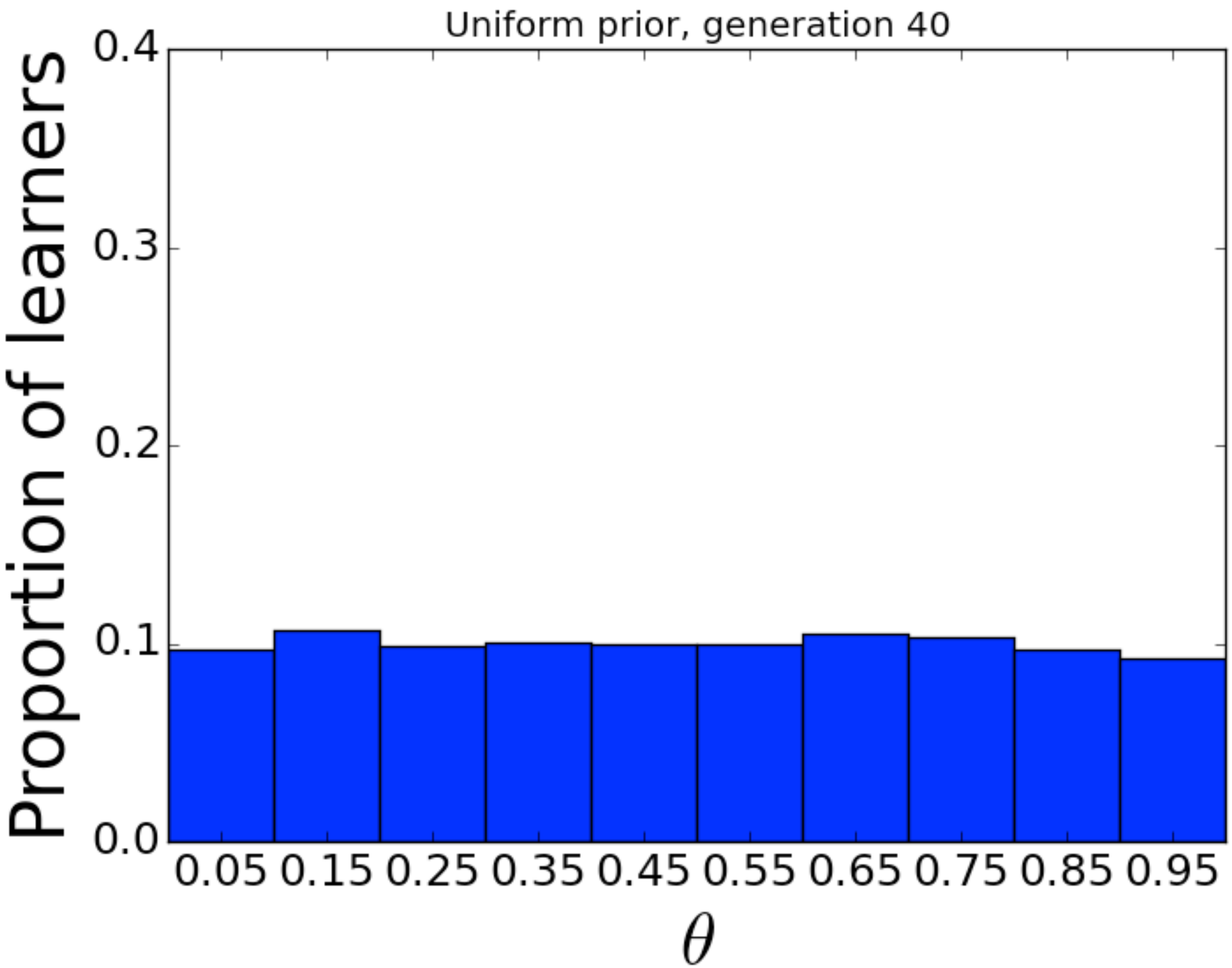
Watching the prior reveal itself



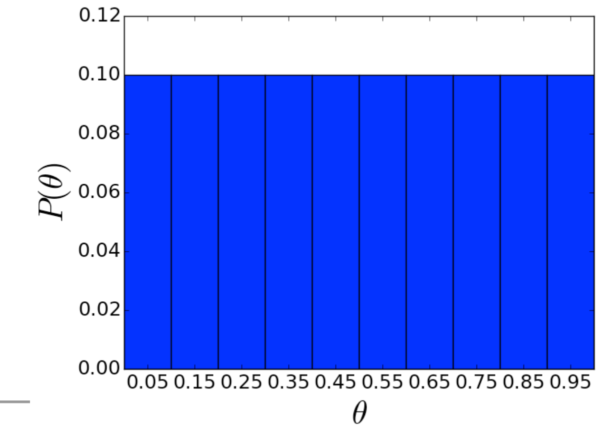


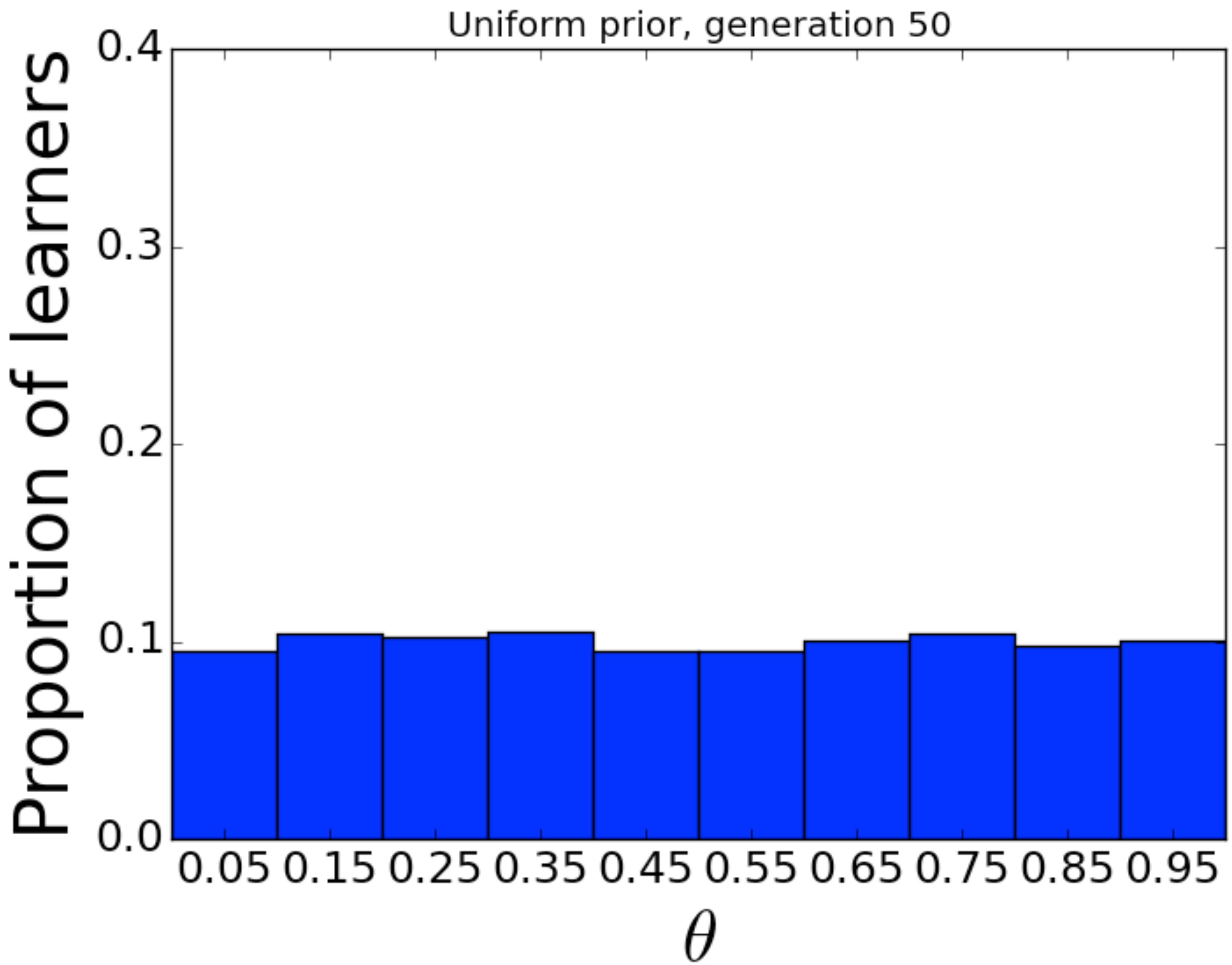
Watching the prior reveal itself



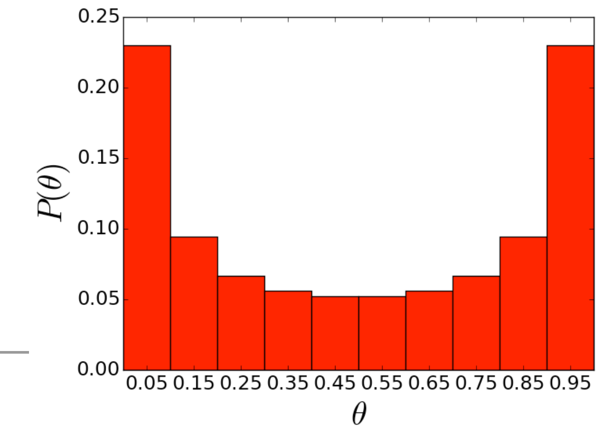


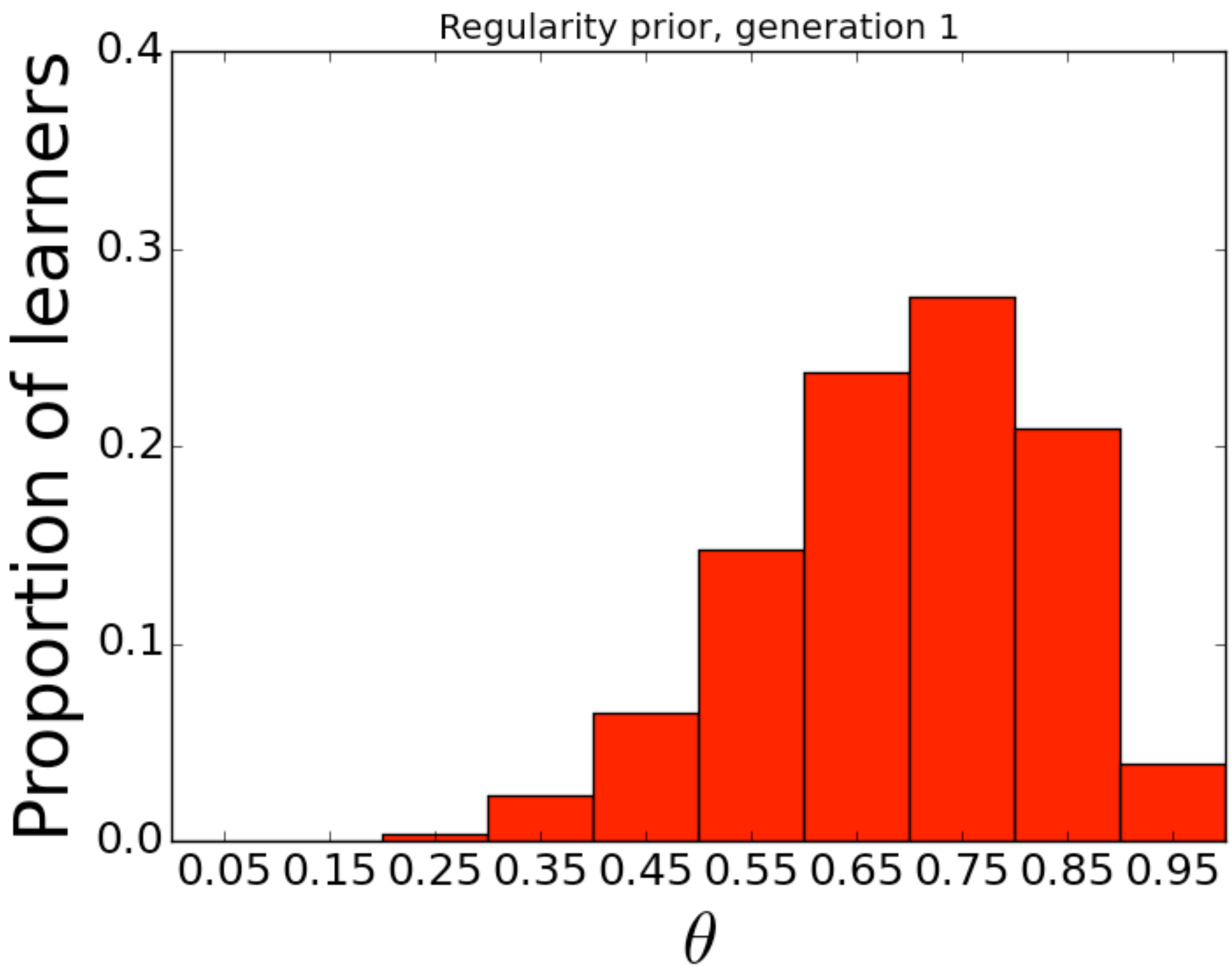
Watching the prior reveal itself



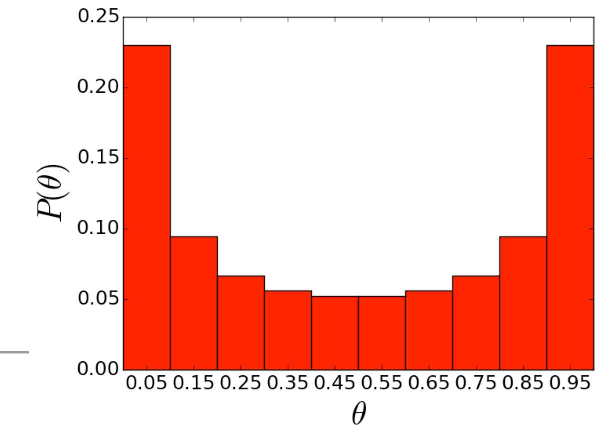


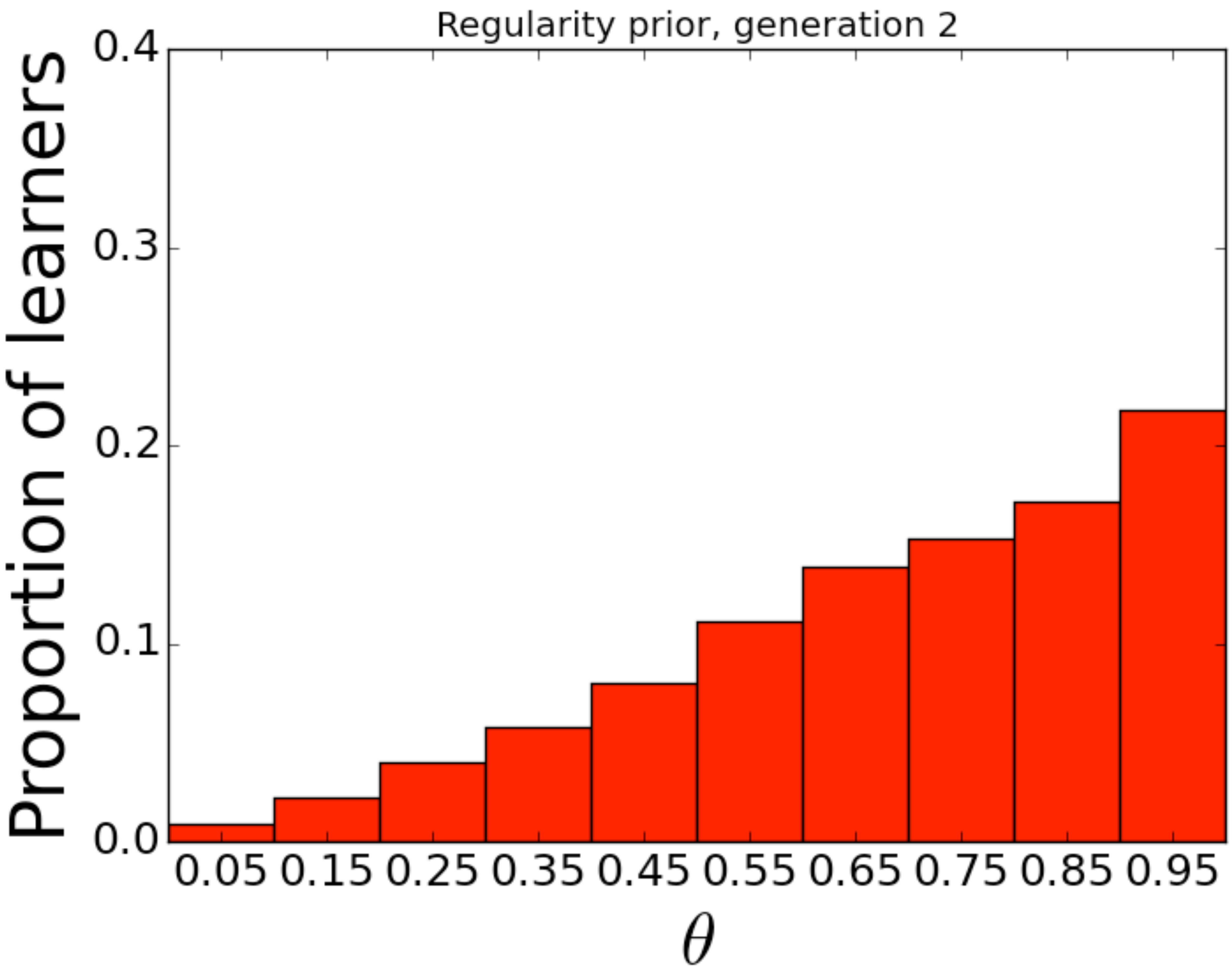
Watching the prior reveal itself



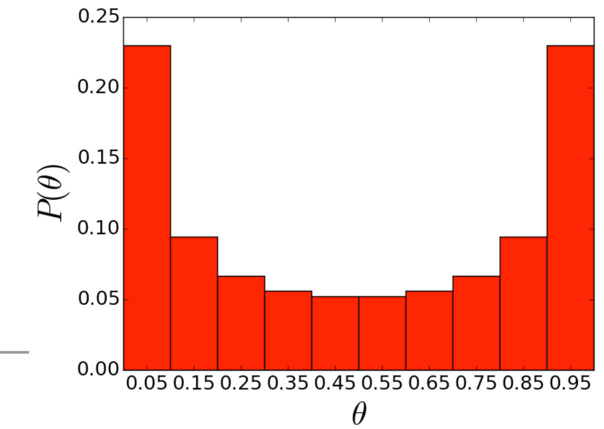


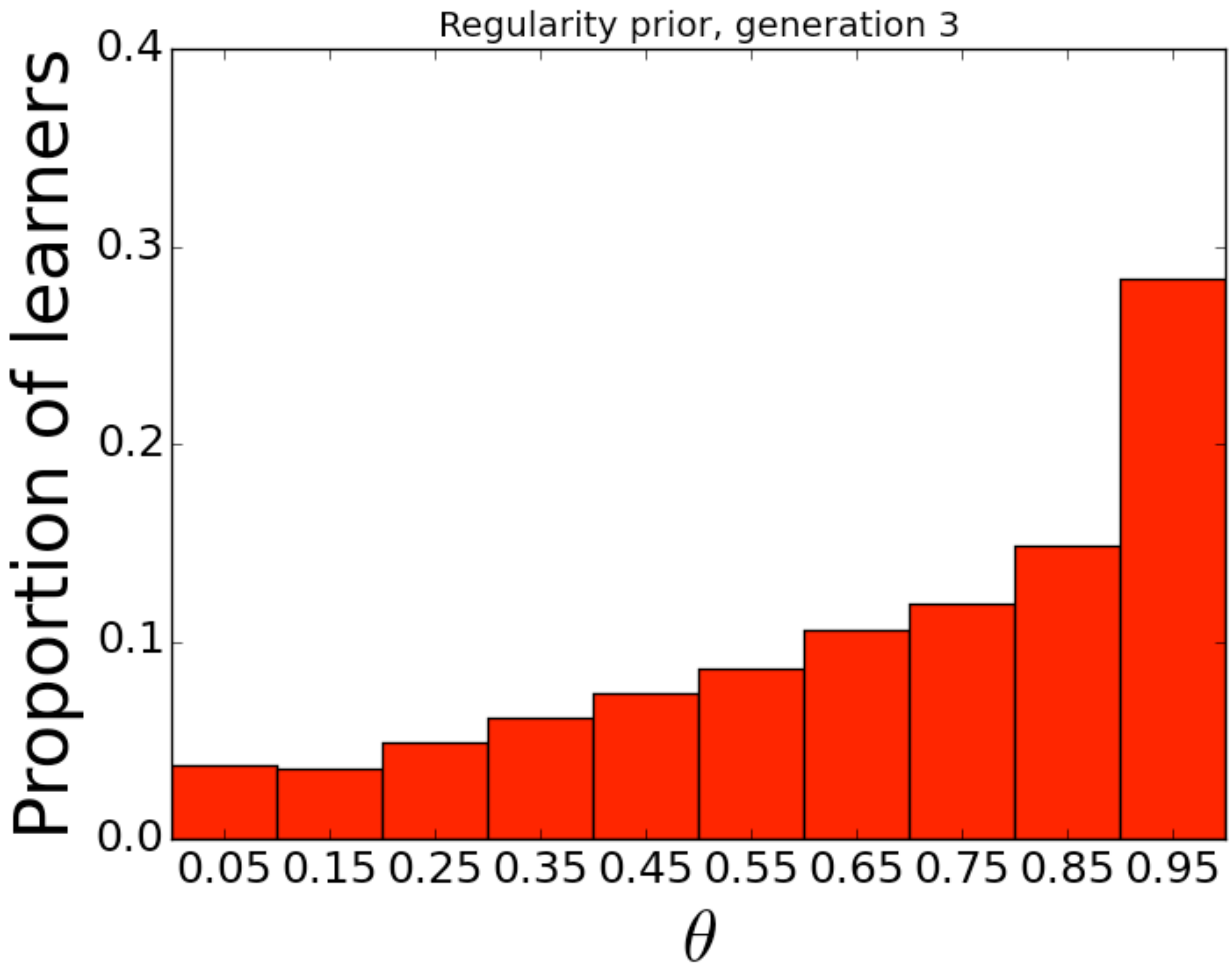
Watching the prior reveal itself



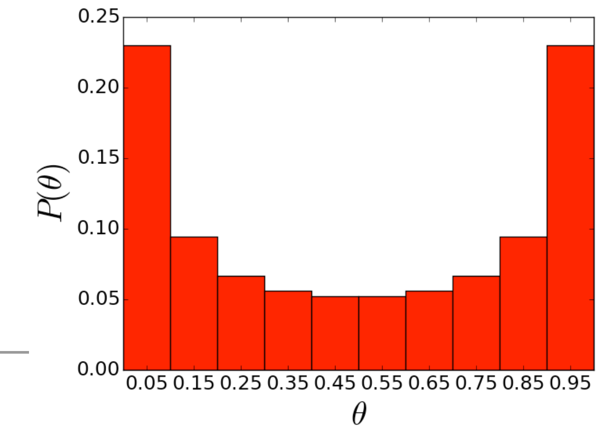


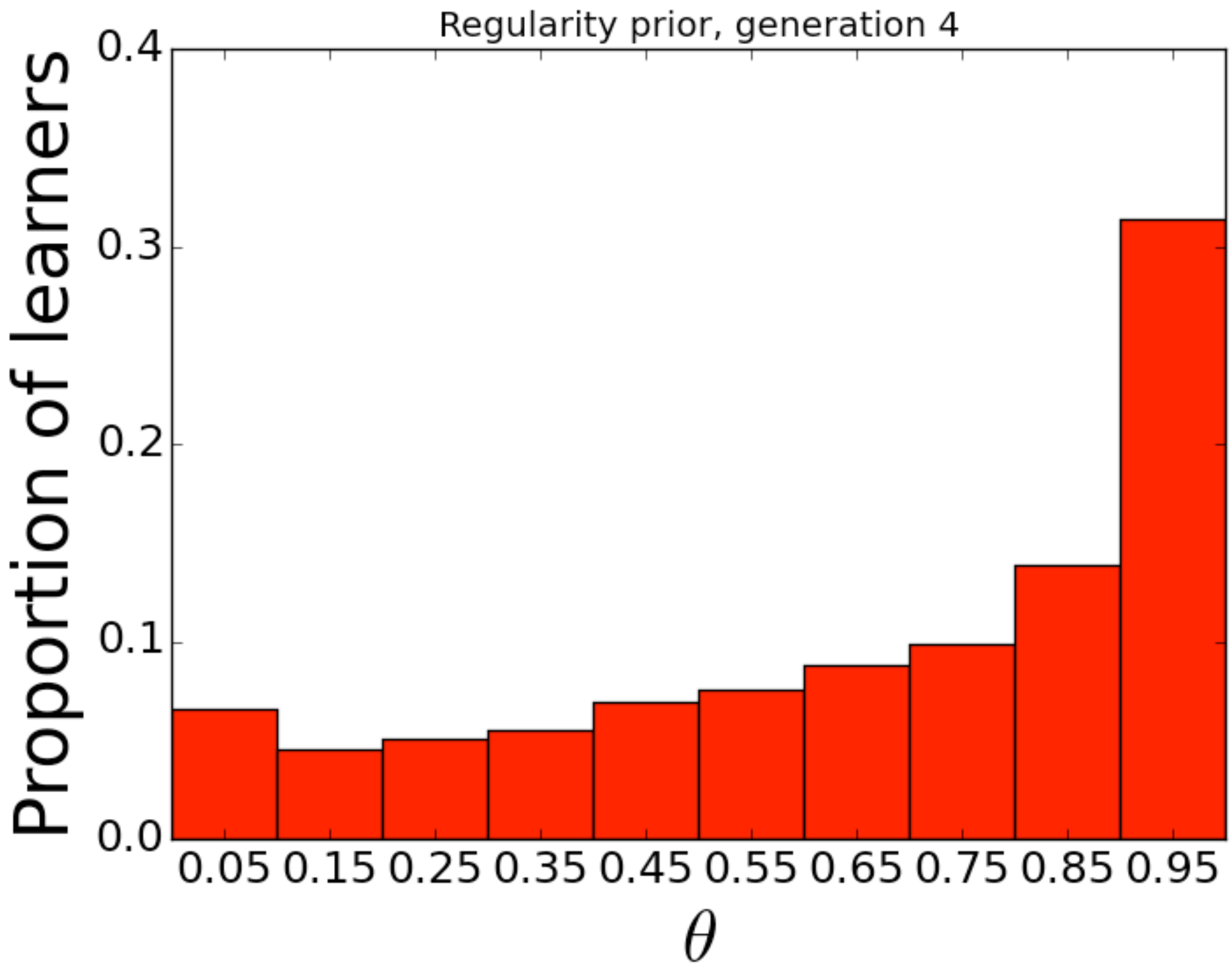
Watching the prior reveal itself



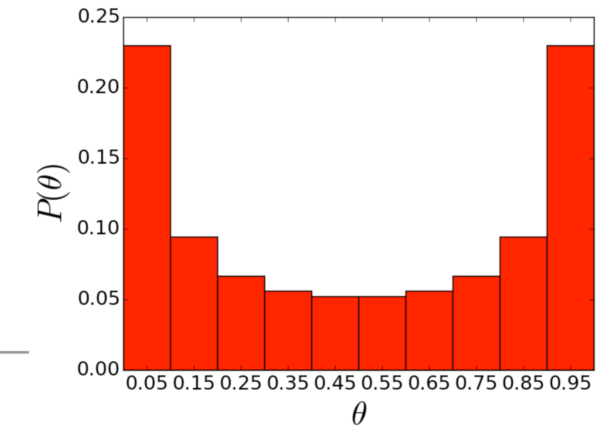


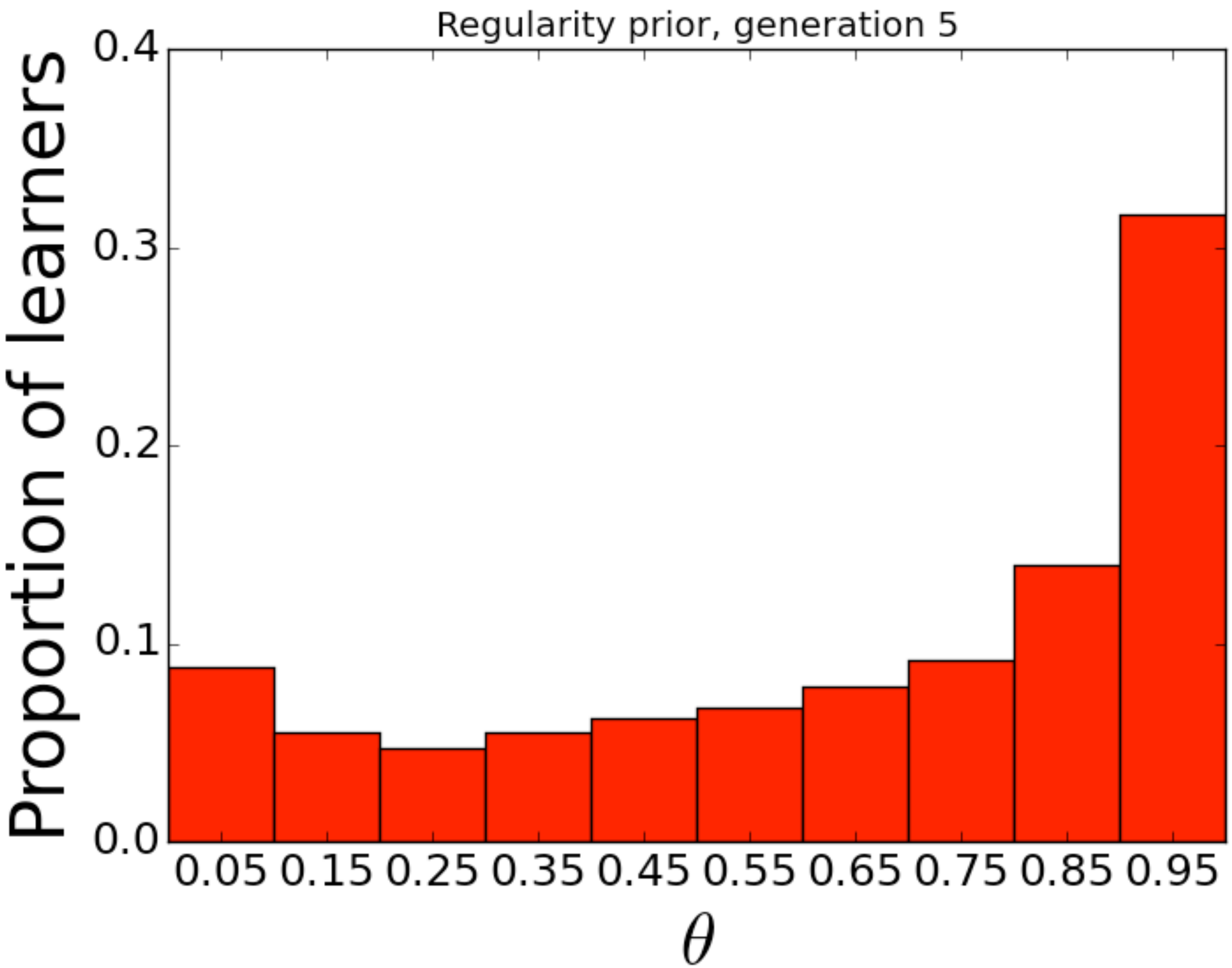
Watching the prior reveal itself



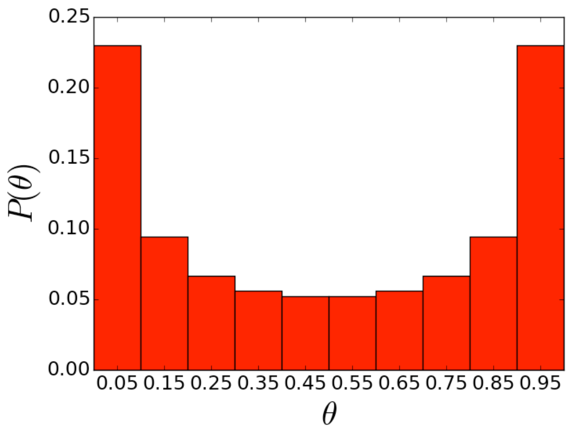


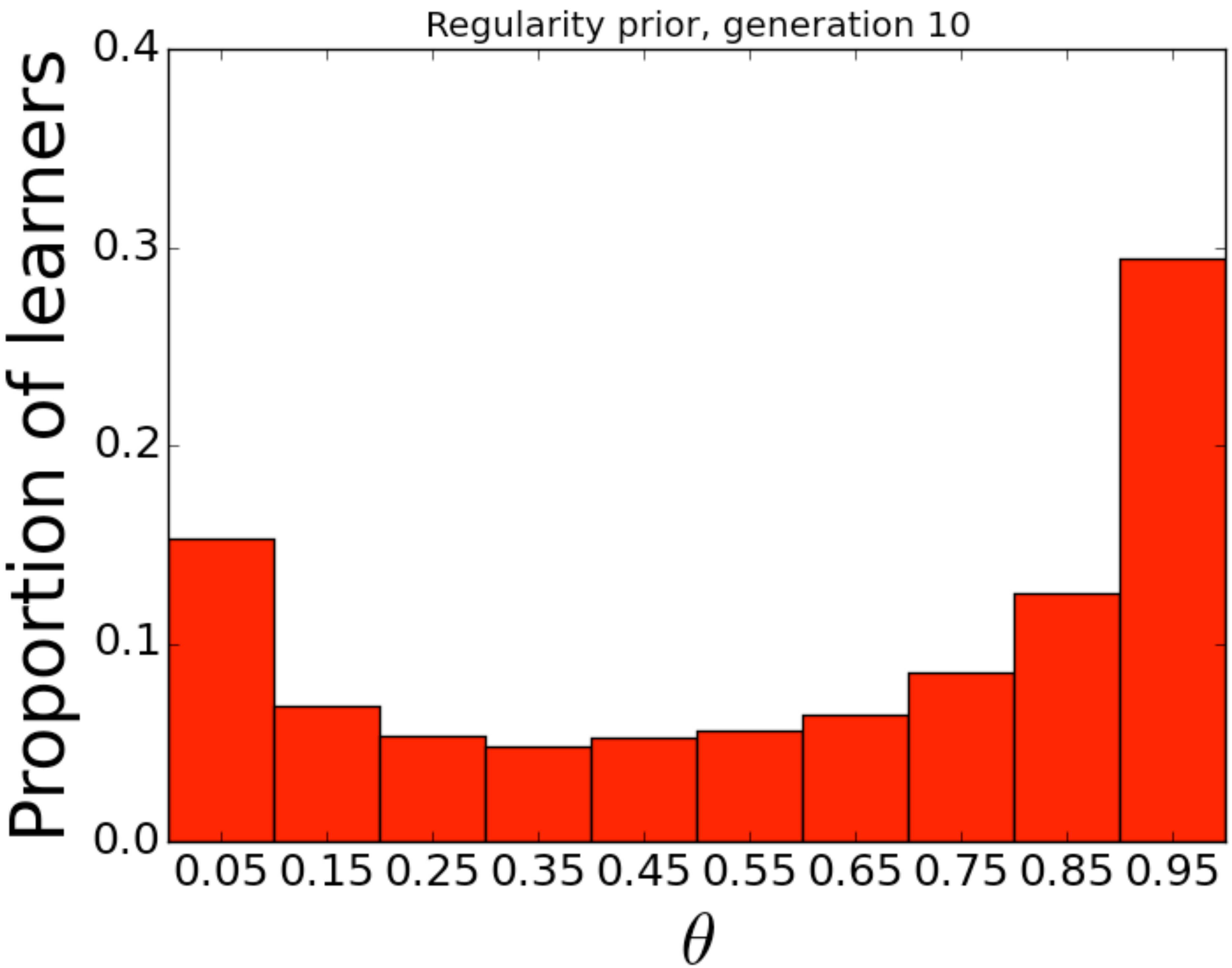
Watching the prior reveal itself



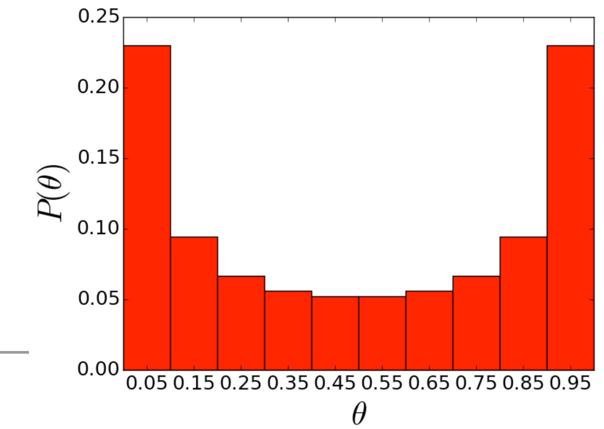


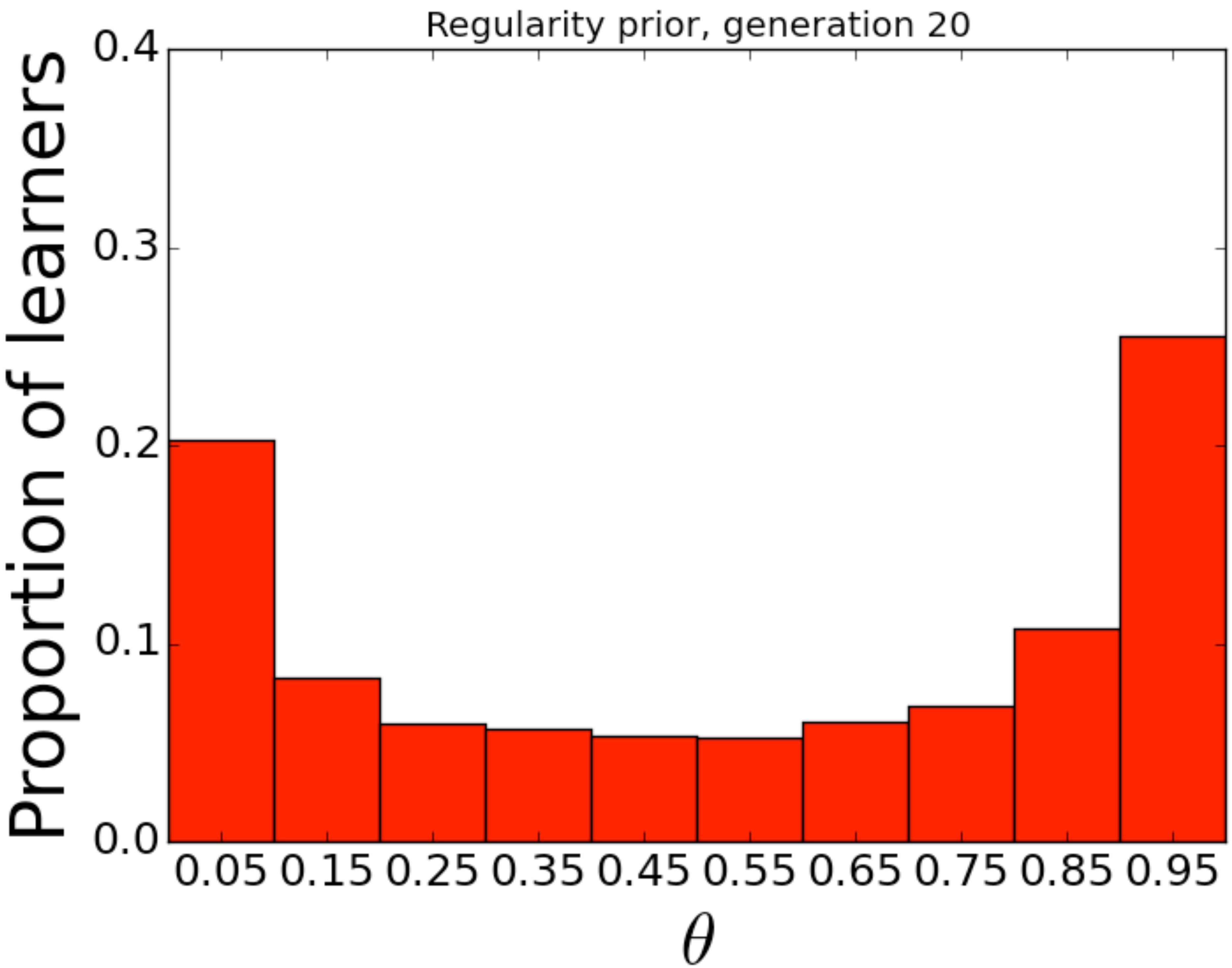
Watching the prior reveal itself



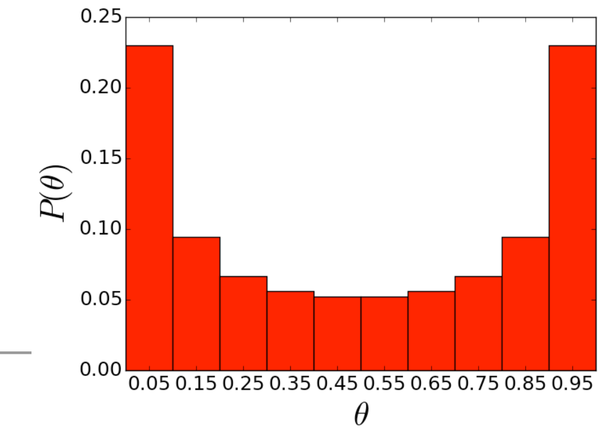


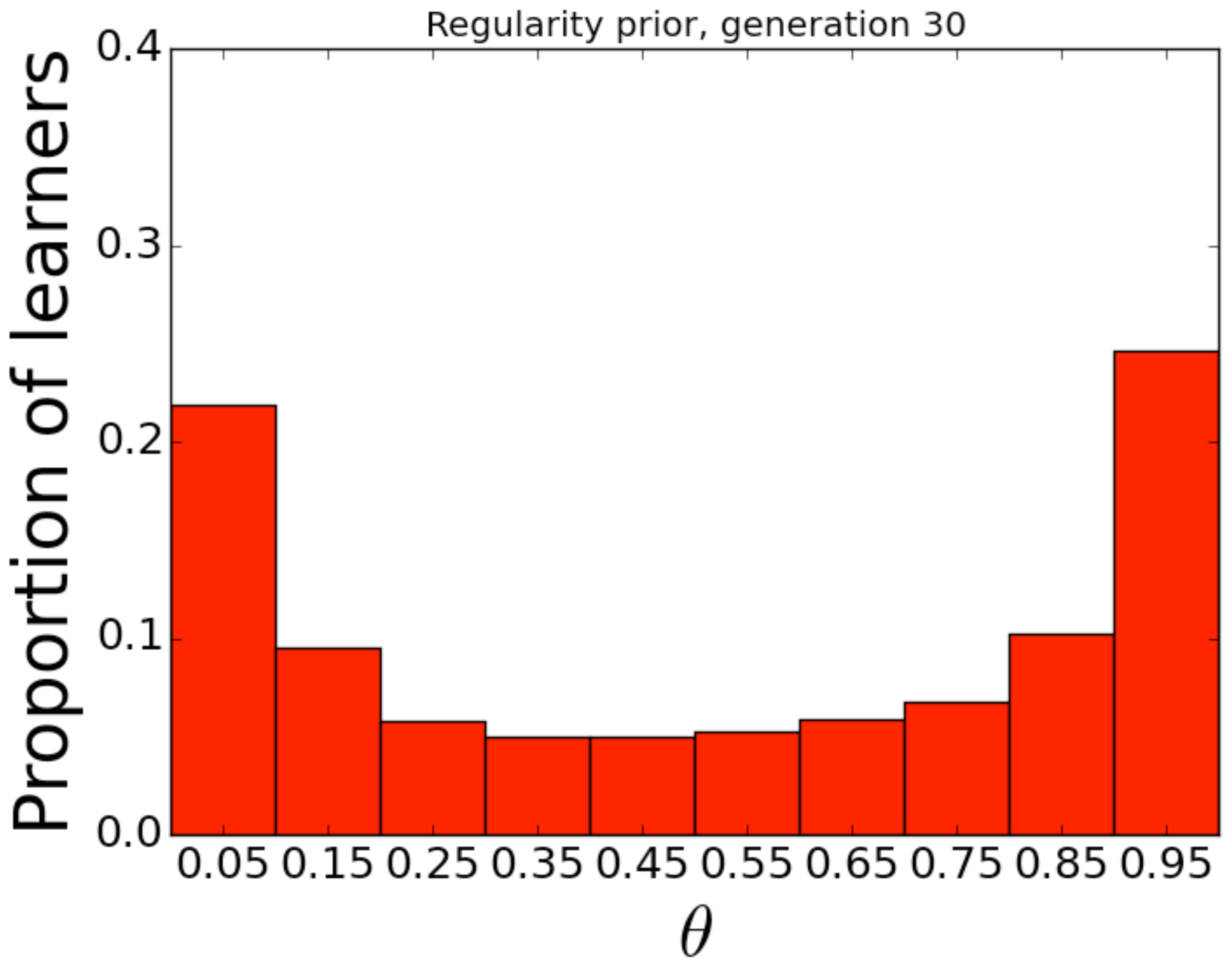
Watching the prior reveal itself



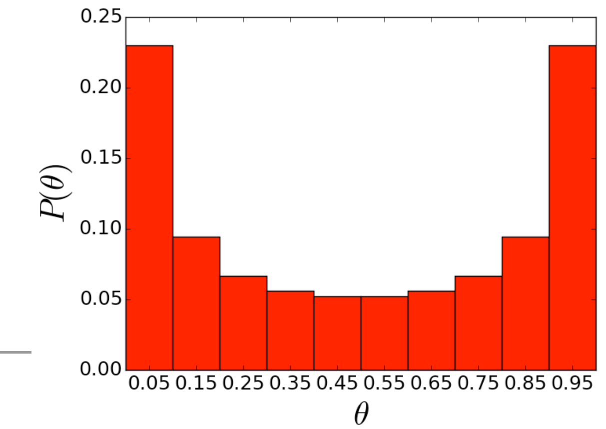


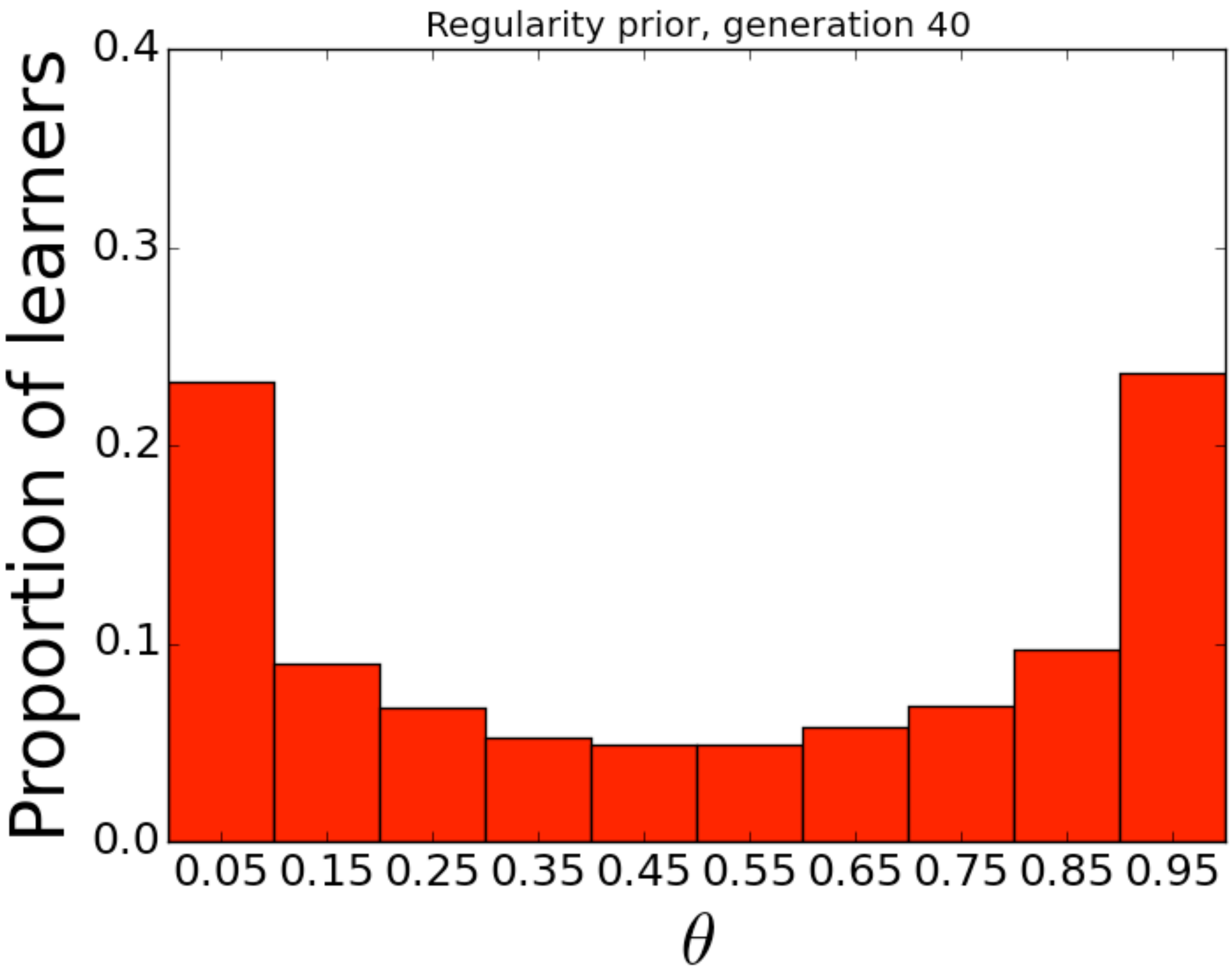
Watching the prior reveal itself



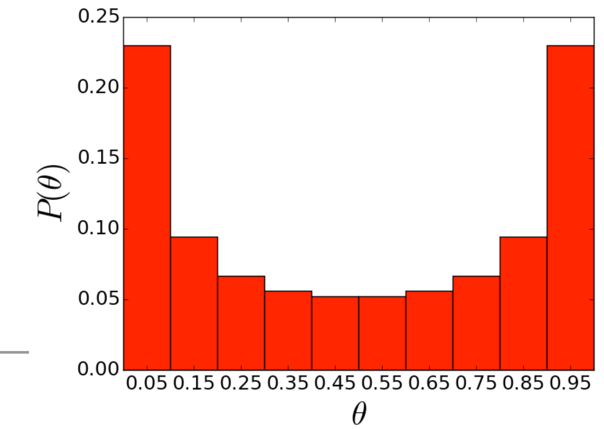


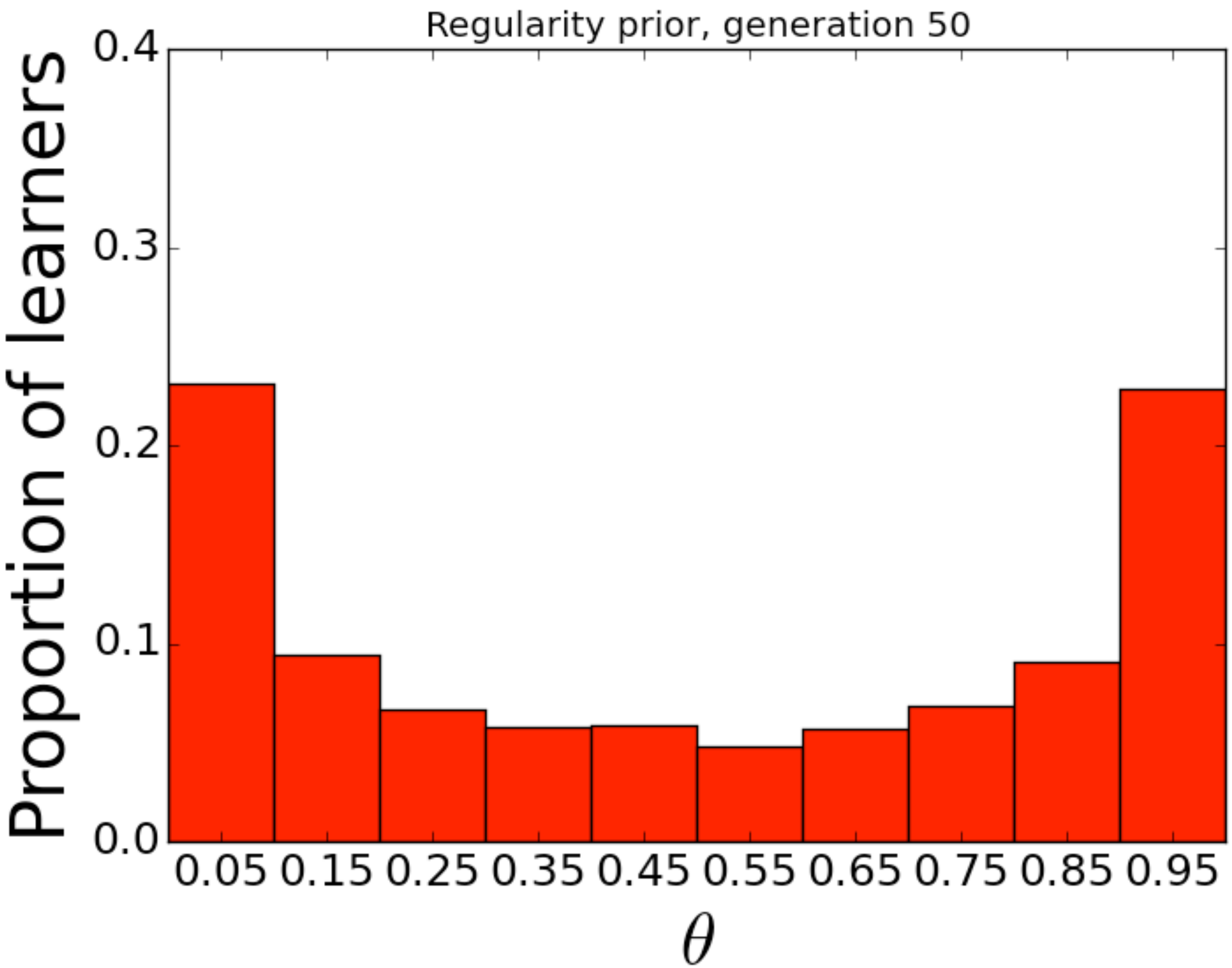
Watching the prior reveal itself



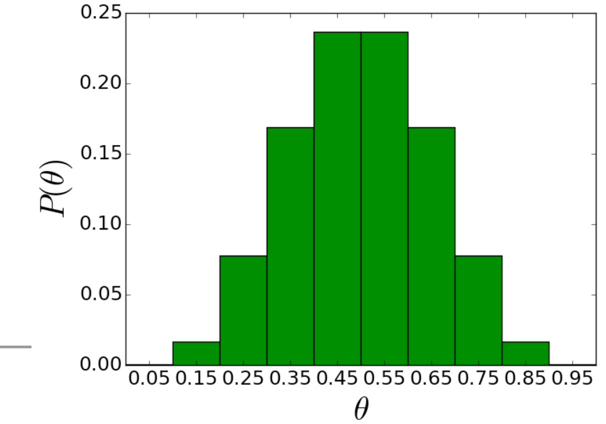


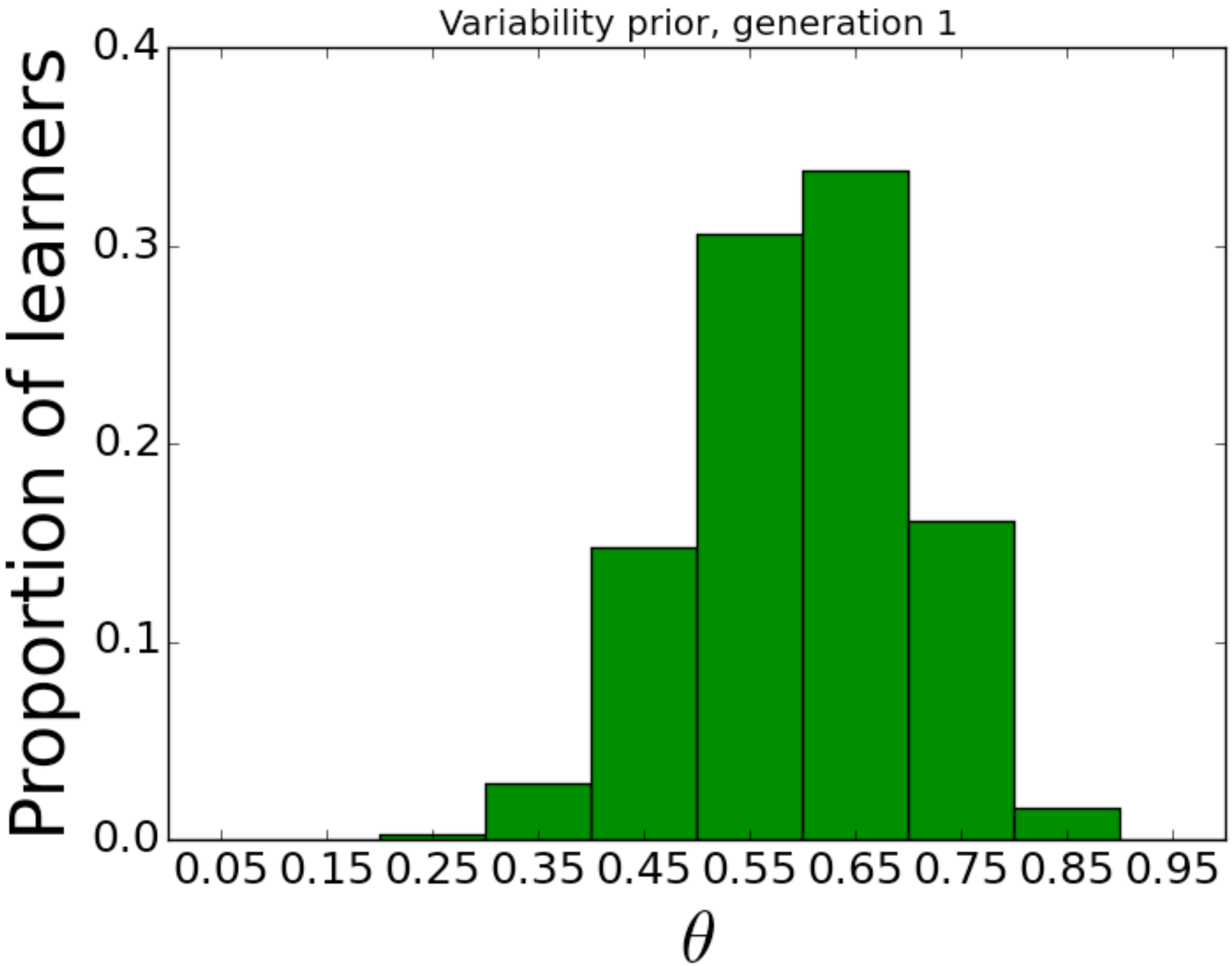
Watching the prior reveal itself



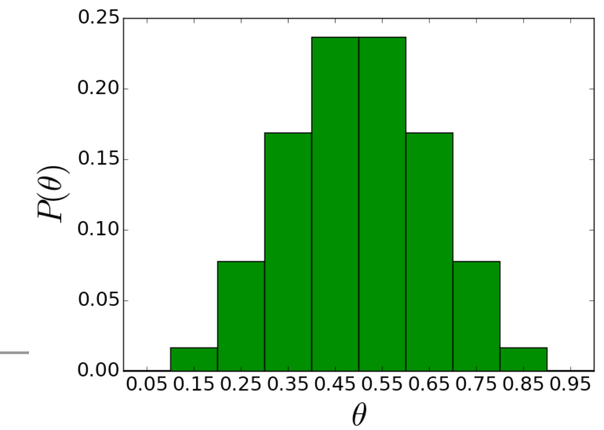


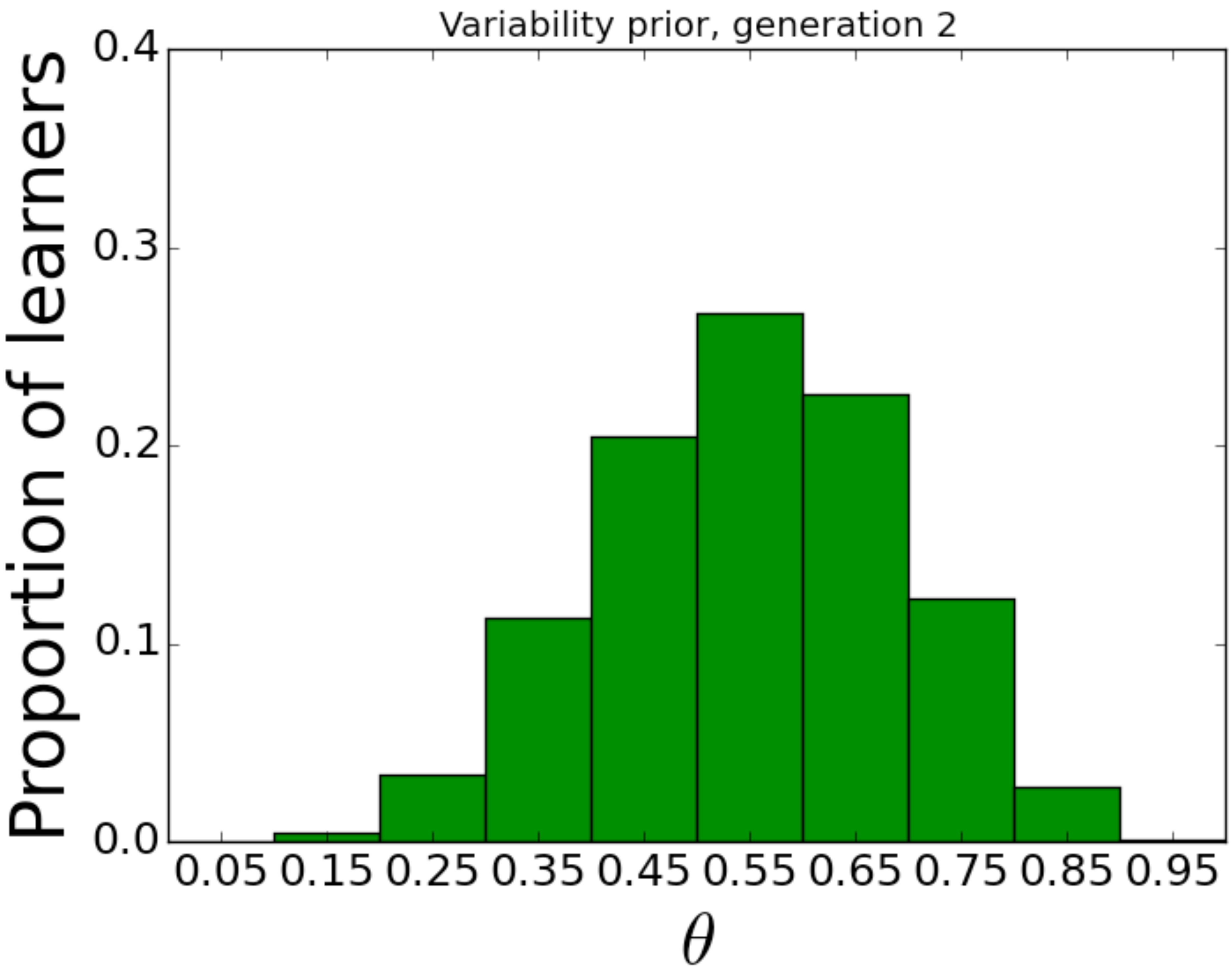
Watching the prior reveal itself



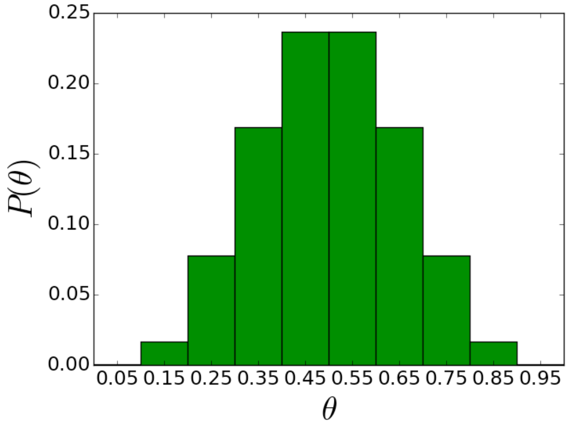


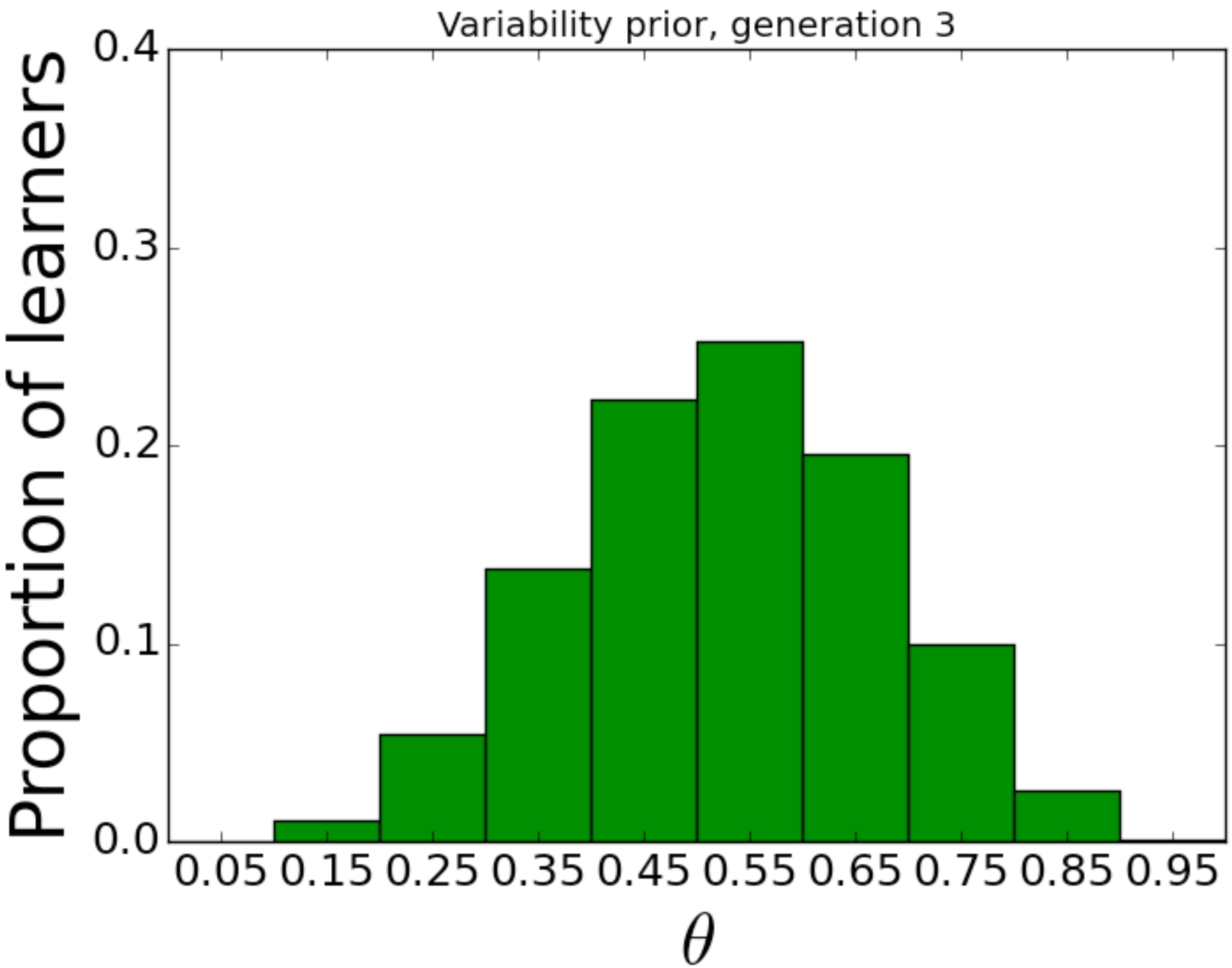
Watching the prior reveal itself



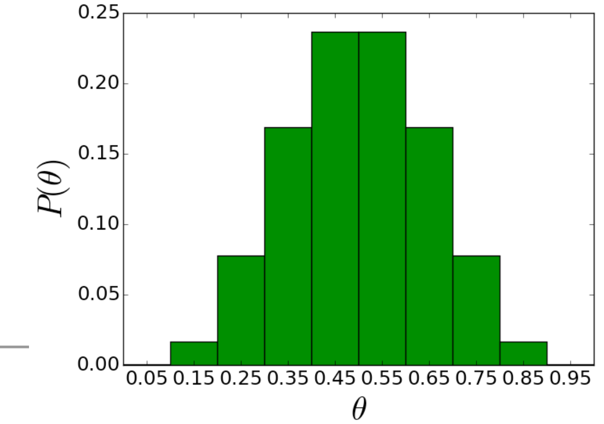


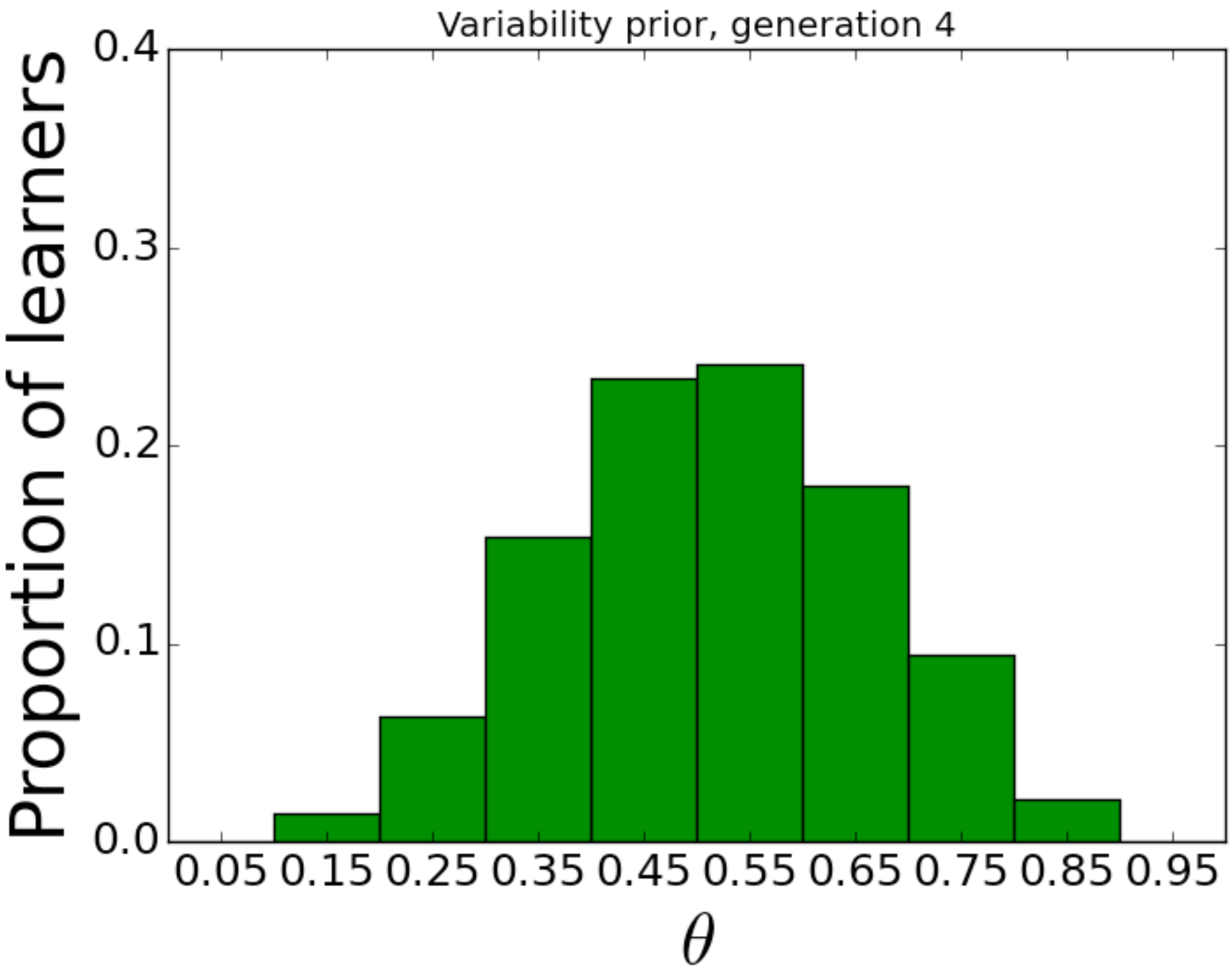
Watching the prior reveal itself



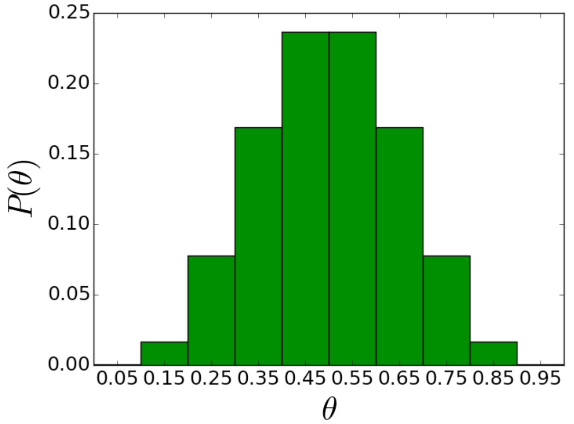


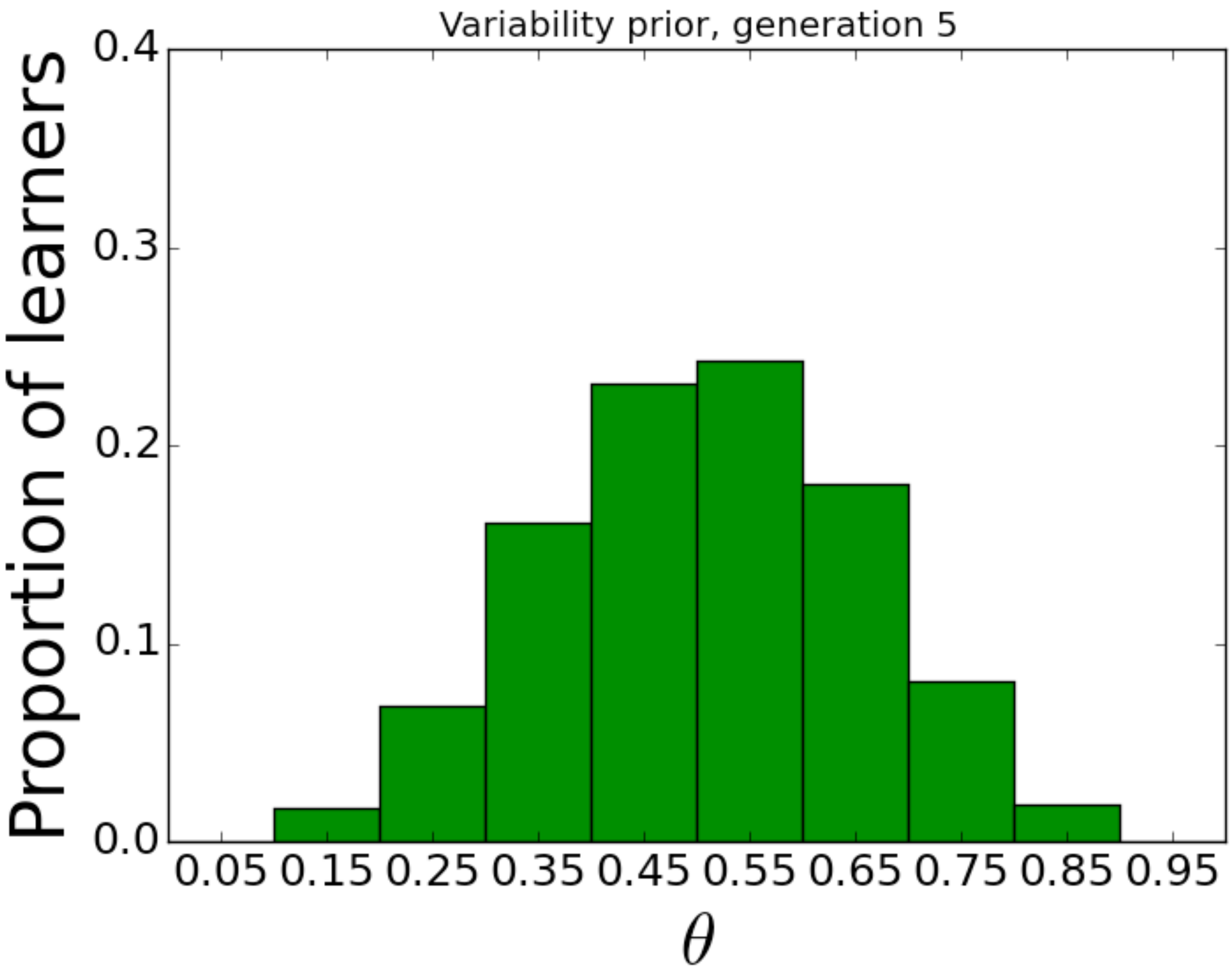
Watching the prior reveal itself



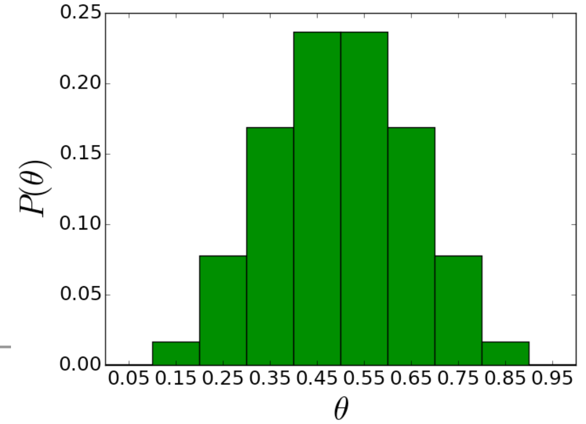


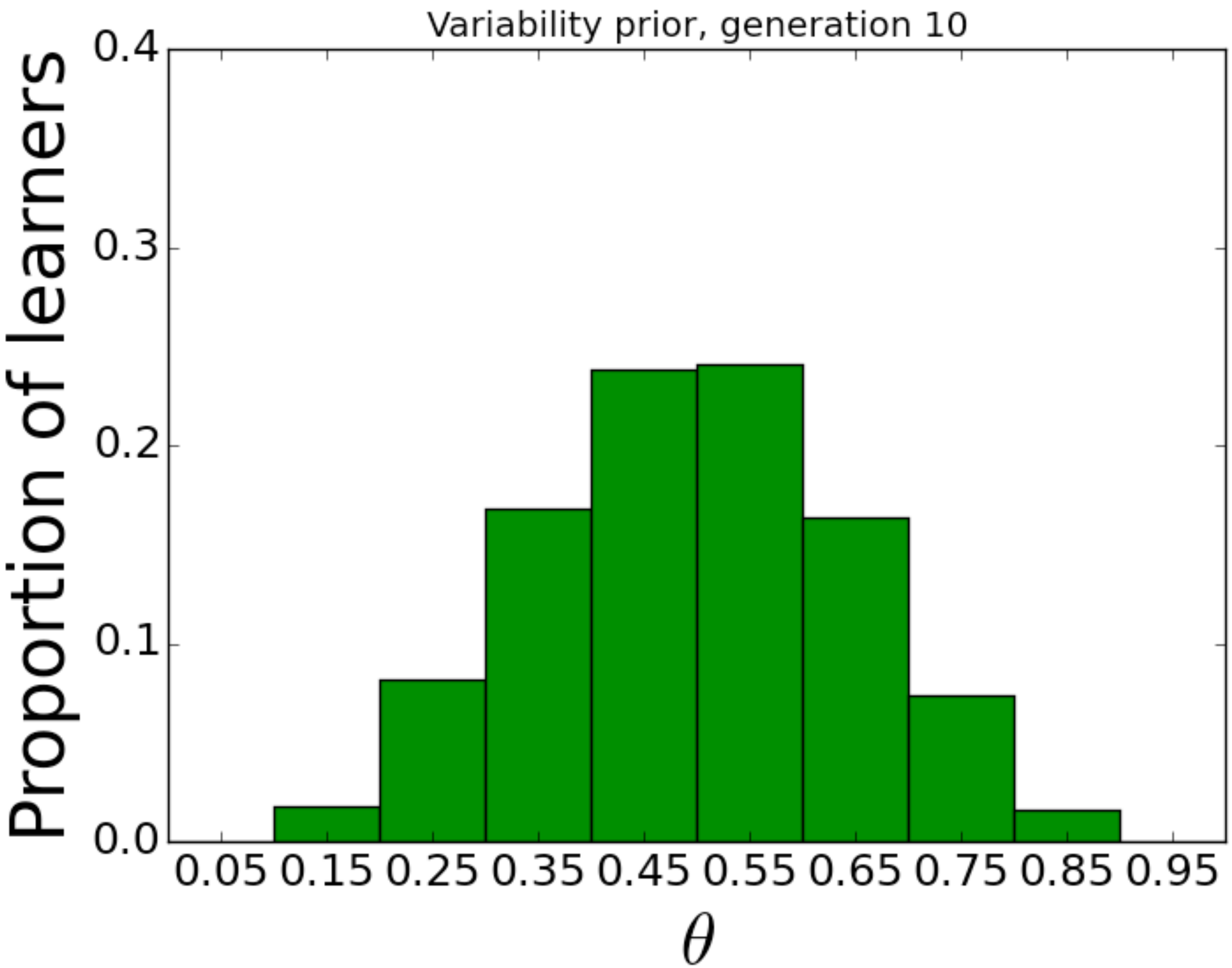
Watching the prior reveal itself





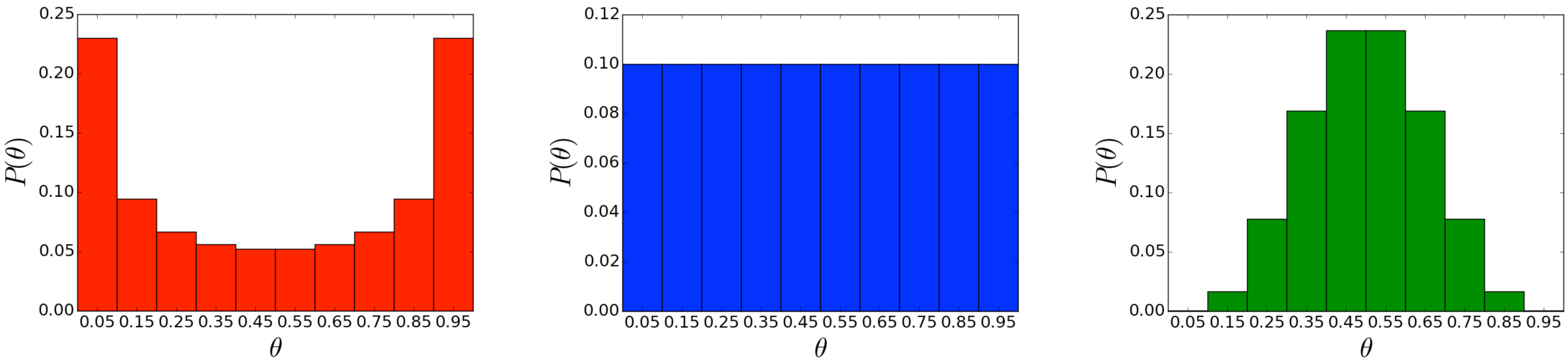
Watching the prior reveal itself



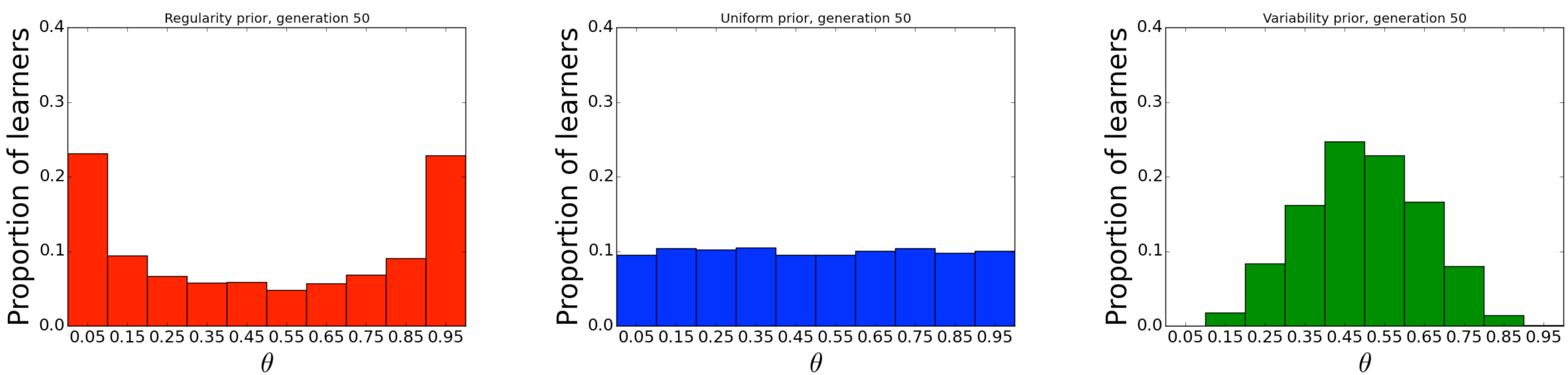


Culture converges to the prior

- Priors



- Distribution of languages after 50 generations



The “stationary distribution”

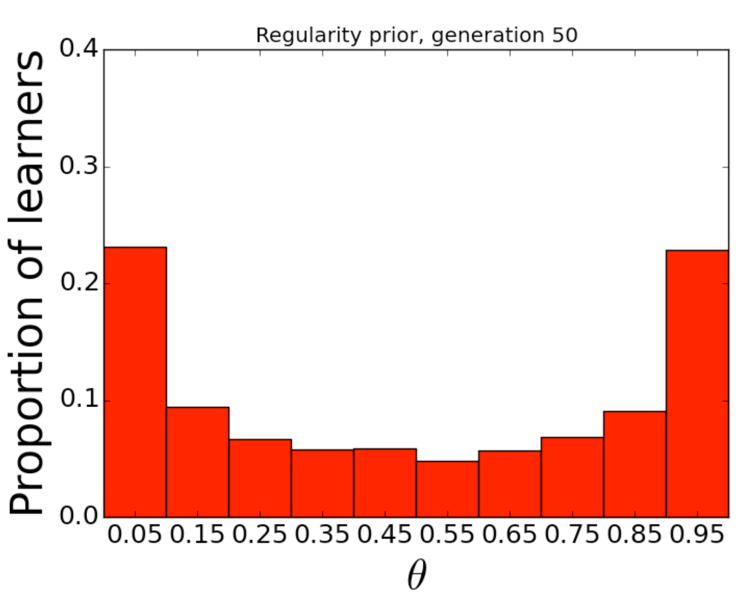
The **stationary distribution** is a probability distribution over languages. In other words, it states how likely it is you’ll see a particular language.

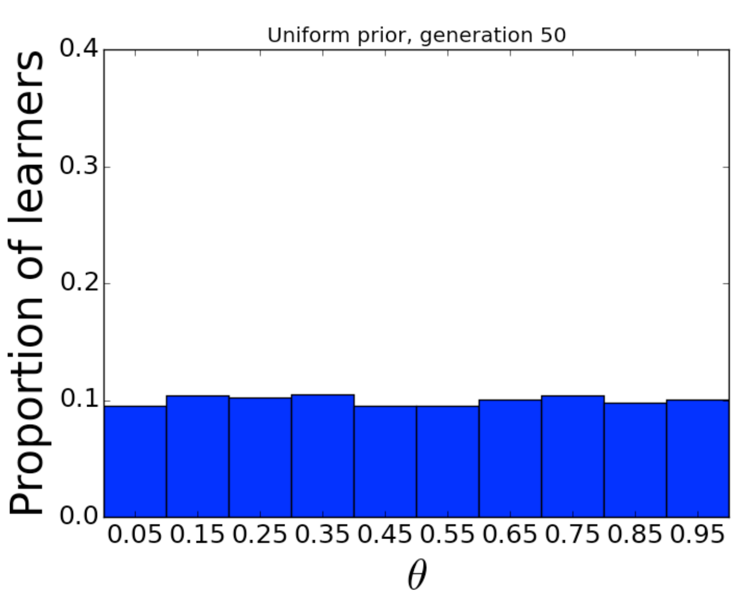
It’s **stationary** in the sense that it is not expected to be different in the future.

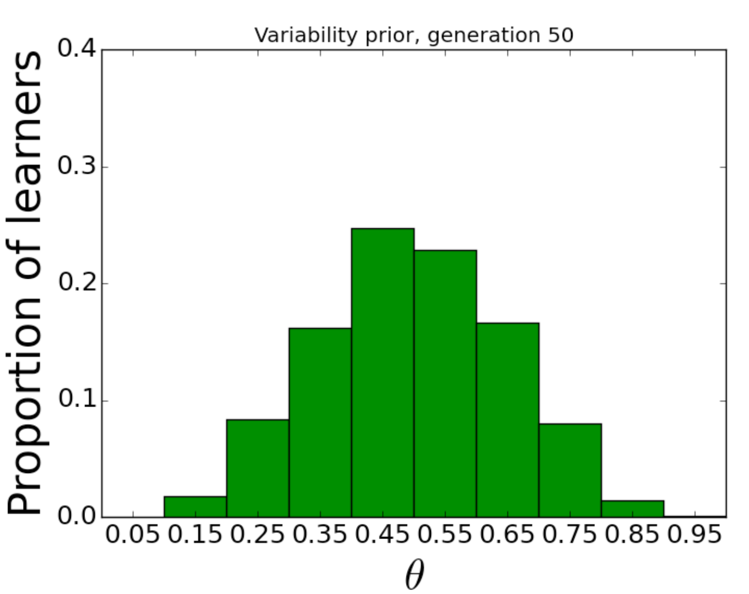
That doesn’t mean languages aren’t changing, just that the overall expectation of particular languages appearing is stable.

It **emerges** (eventually) from the process of cultural evolution, when the particular starting point of the simulation is eventually washed away.

When we talk about **language universals**, we’re essentially talking about a stationary distribution.







"Convergence to the prior"
The relationship between learners and language

Prior bias → Stationary distribution

Nature of learners → Linguistic universals

Innateness → Language

These are all essentially equivalent…

Convergence to the prior suggests that the languages we see are simply a transparent reflection of what is innate.

Hang on a minute...

- So iterated bayesian learning appears to provide strong support for a nativist view of the language faculty. Right?
- This runs counter to the results we'd been working on here in Edinburgh
    - We argued that it was features of the bottleneck that was driving adaptation of the language
    - To put it another way, we're saying that cultural evolution matters beyond merely delivering up the prior
- Hmmm…

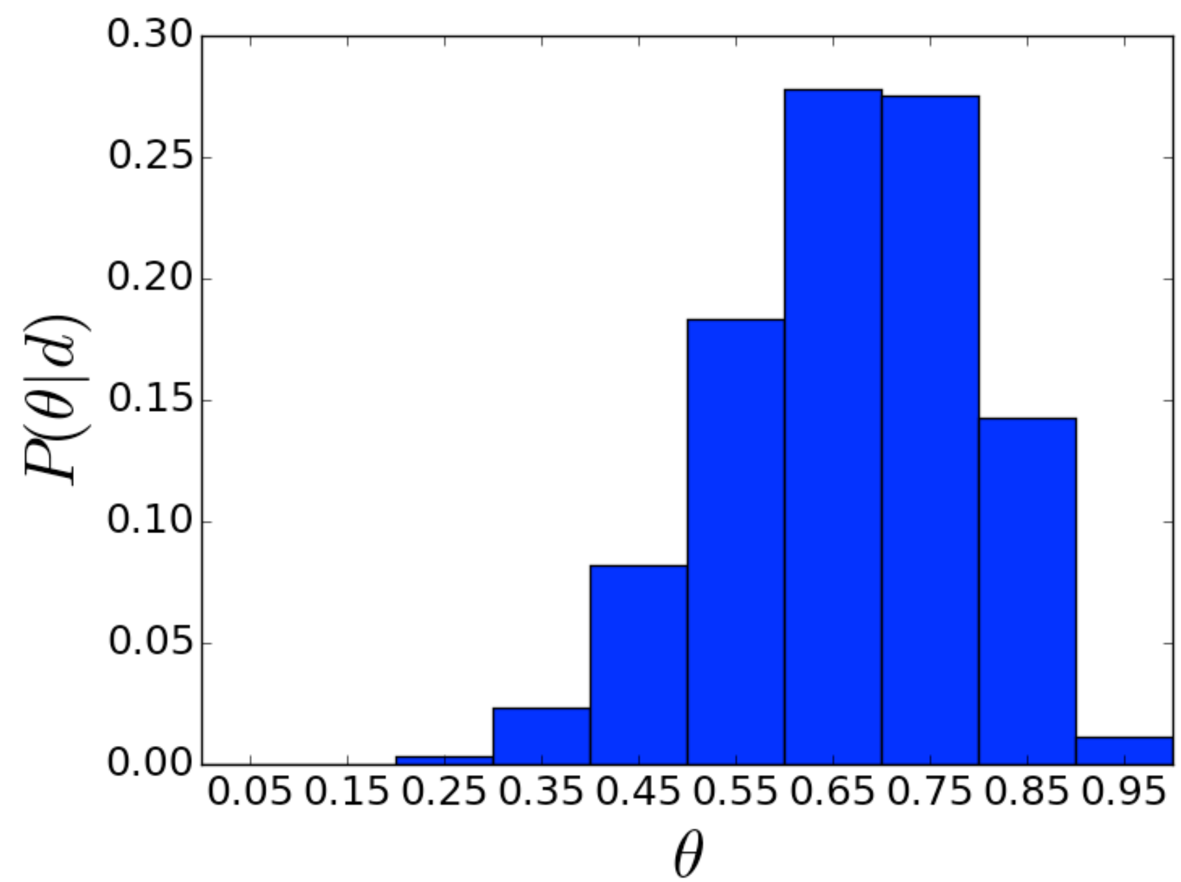
Some subtleties in the model

- Kirby, Dowman & Griffiths (2007): tried to square the Bayesian model with what we **thought** we knew about cultural evolution of language

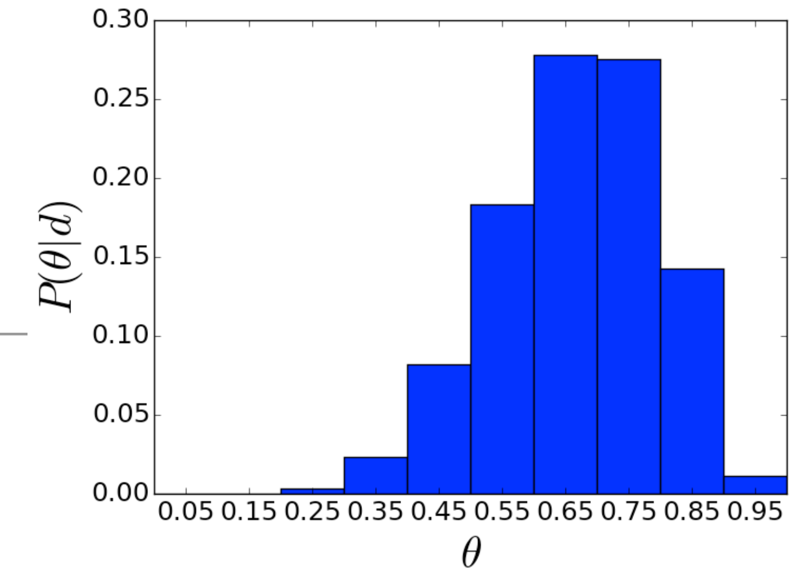
- Whole thing revolves around a very subtle point

  - How do you decide, given the posterior, which language to select?

$$P(h|d) = \frac{P(d|h)P(h)}{P(d)}$$



Sampling vs. MAP



- There are (at least) two sensible choices:

  - Sampling: given a particular distribution of probabilities, pick your hypothesis from the distribution proportionally.

    (If it's ten times more likely to be language A than language B, 10% of the time pick language B)

  - MAP: given a particular distribution of probabilities, pick the best. This is called the maximum a-posteriori (MAP) hypothesis

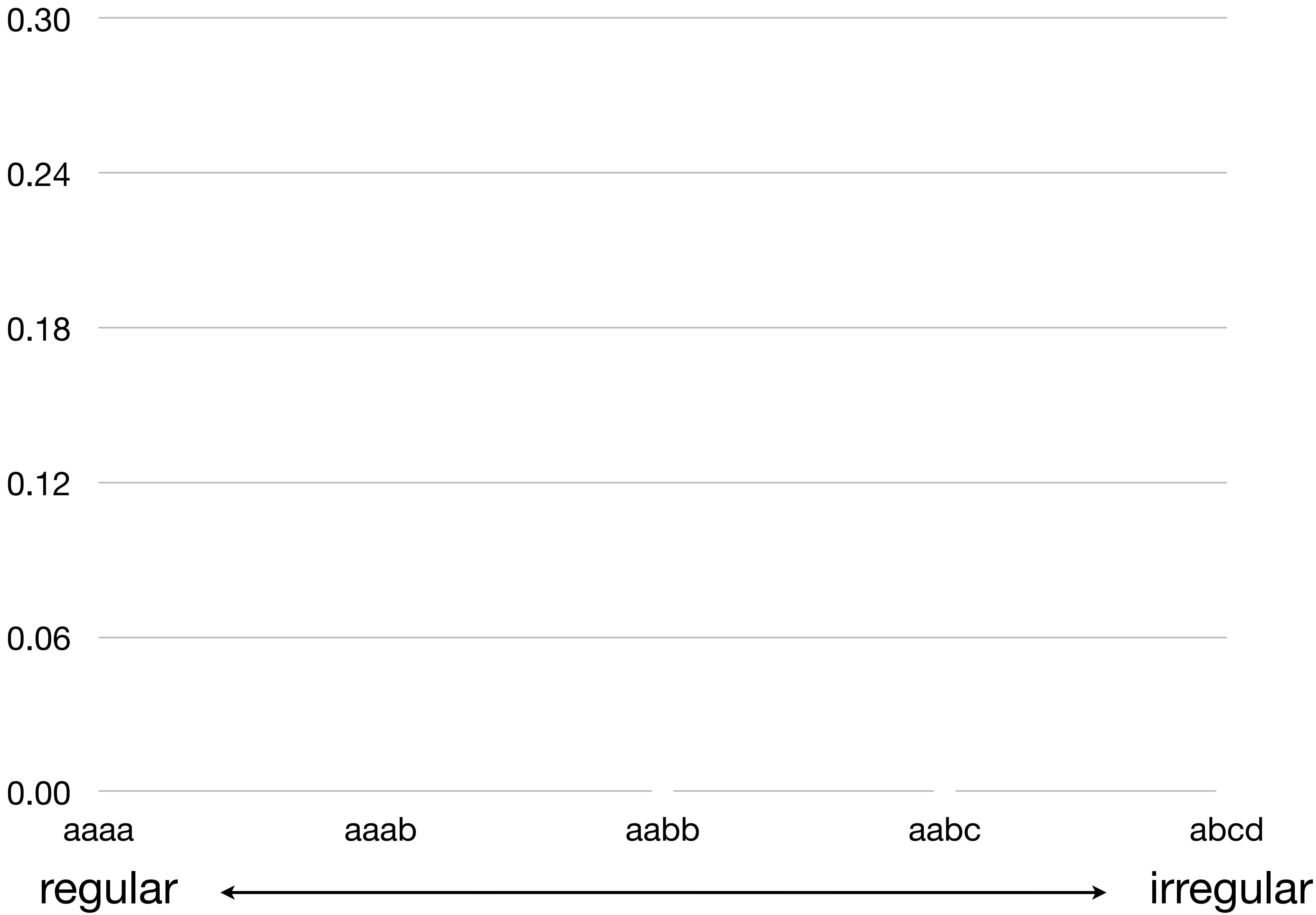
    (If it's more likely to be language A than language B, pick language A)

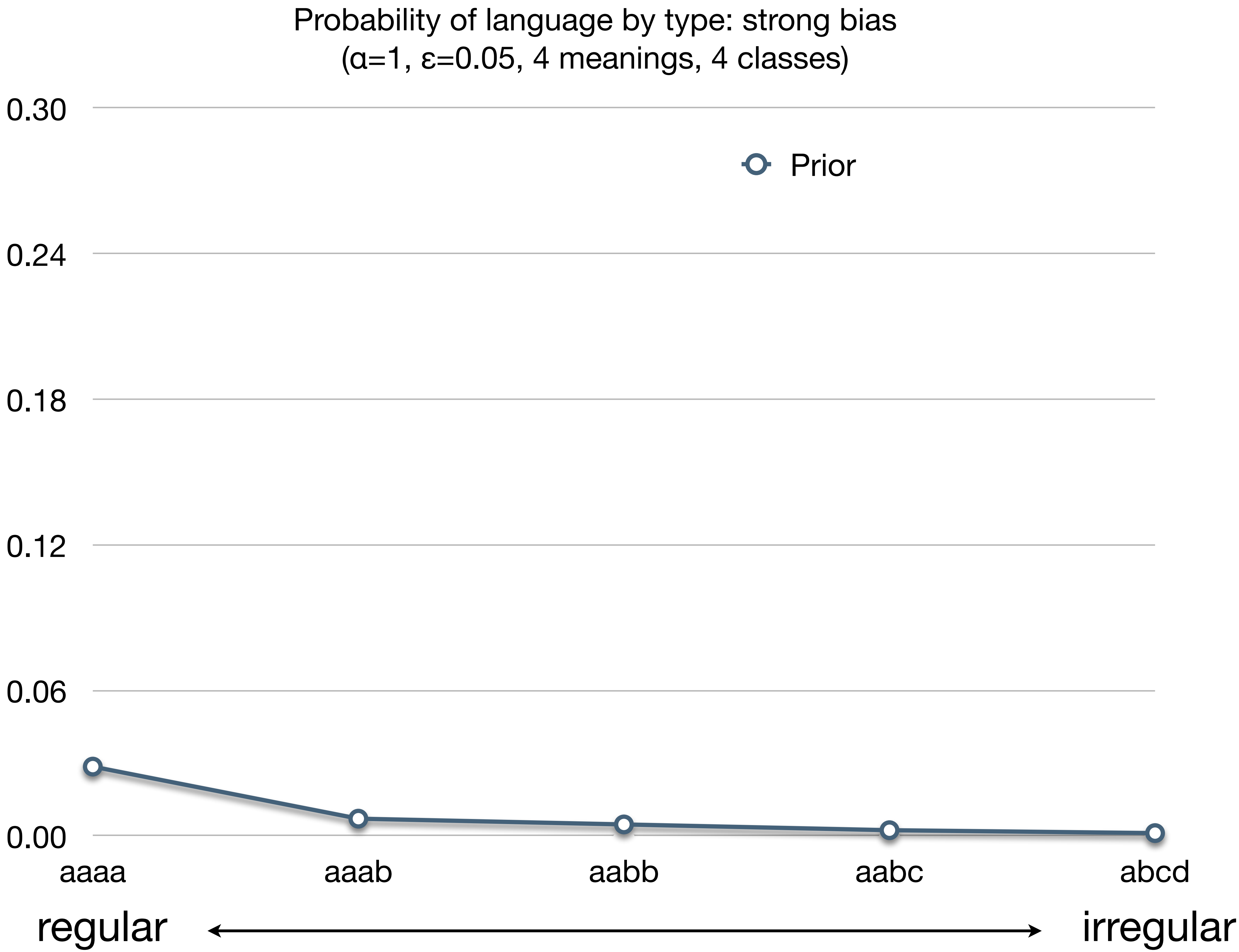
- Griffith & Kalish (2007) were using *sampling.* Kirby et al. (2007) tried MAP.

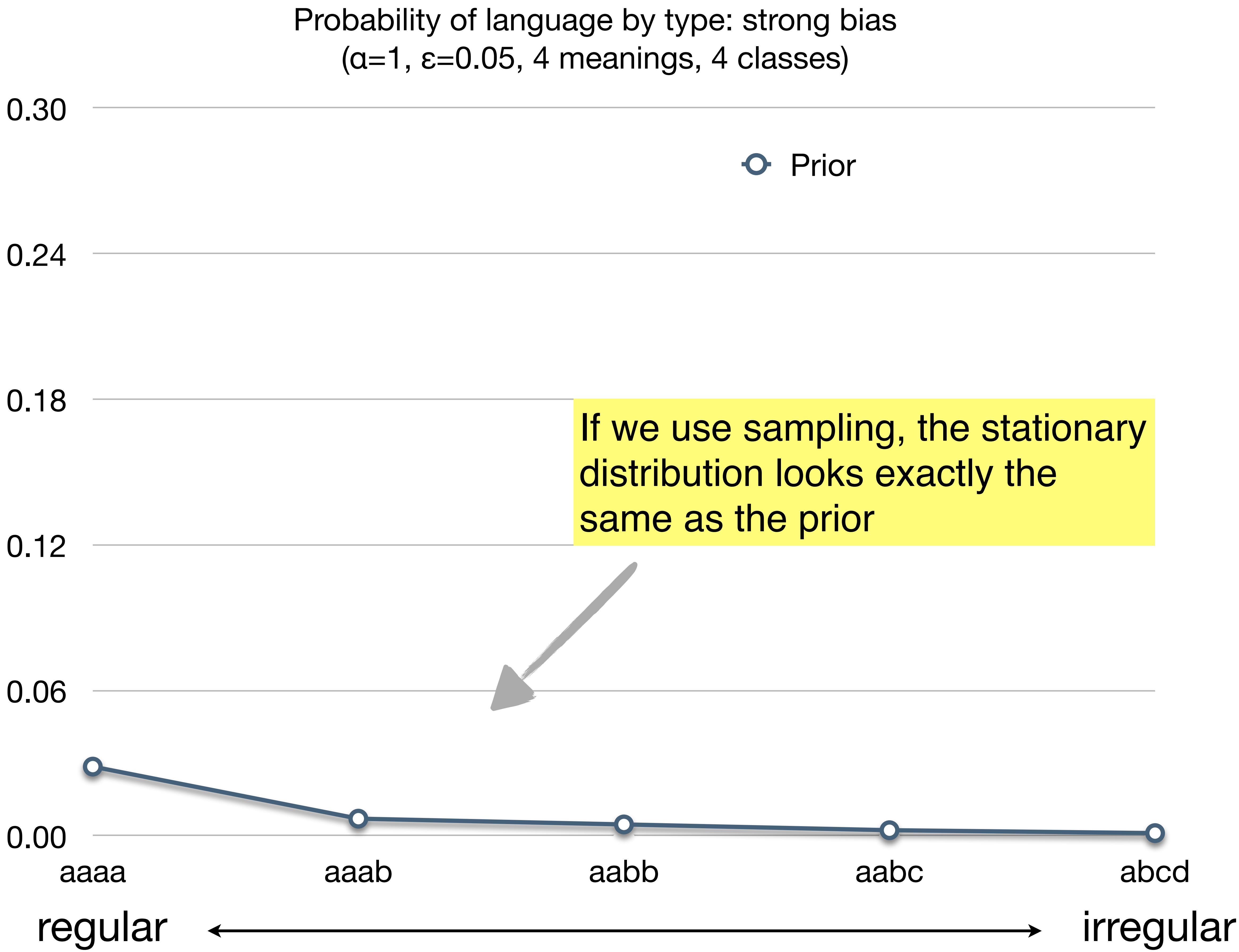
Another model: the evolution of regular paradigms

- Model language as a set of meanings
- These meanings can be expressed **regularly**, or **irregularly** (note: slightly confusingly, this is as different type of “regularity” than you’ve seen before!)
- Start with the assumption that there is a slight innate bias in favour of regularity (based on the simplicity bias)
  - We can vary the strength of this bias
- Assume learners pick the best (i.e. MAP) hypothesis. What happens?

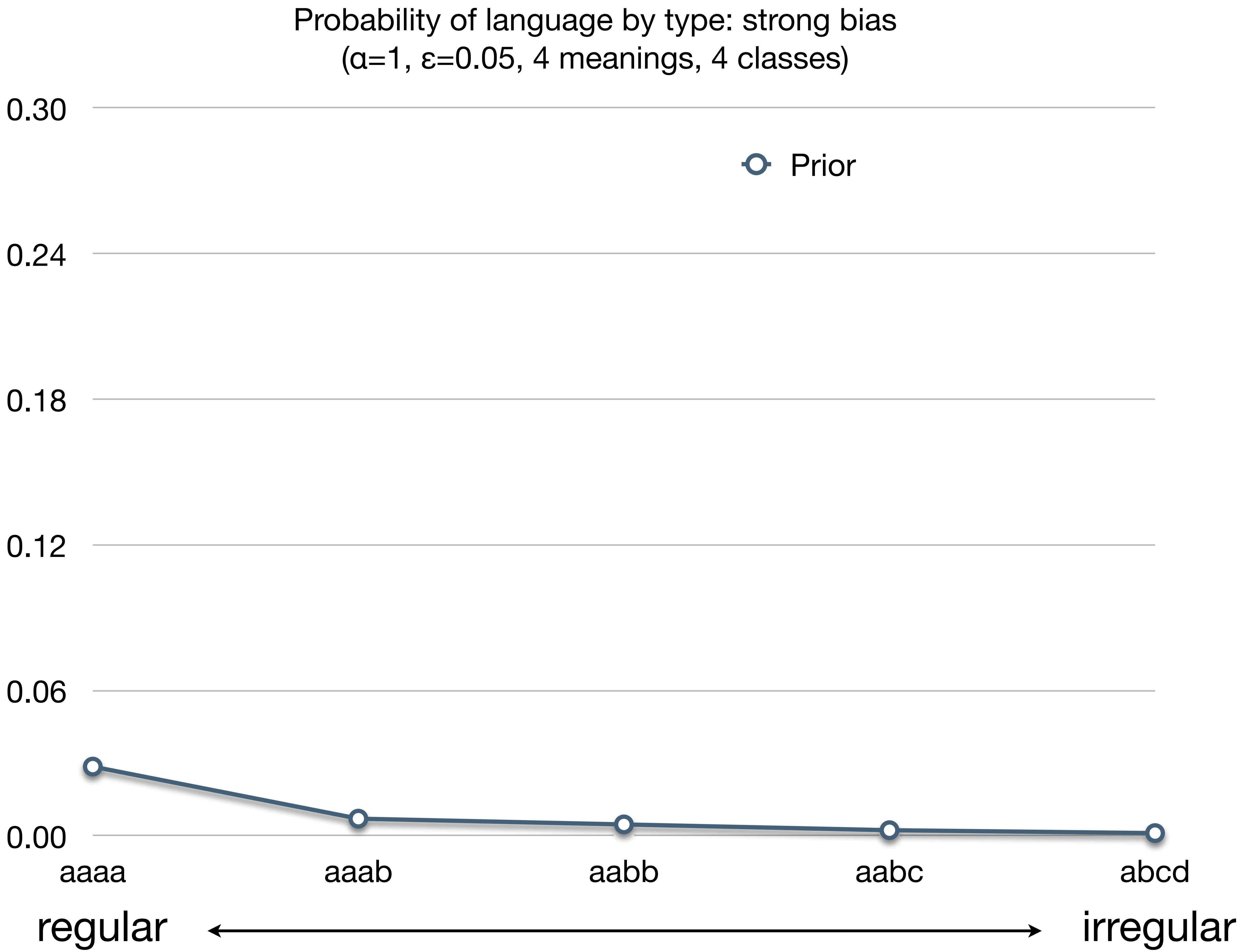
Probability of language by type: strong bias
(α=1, ε=0.05, 4 meanings, 4 classes)

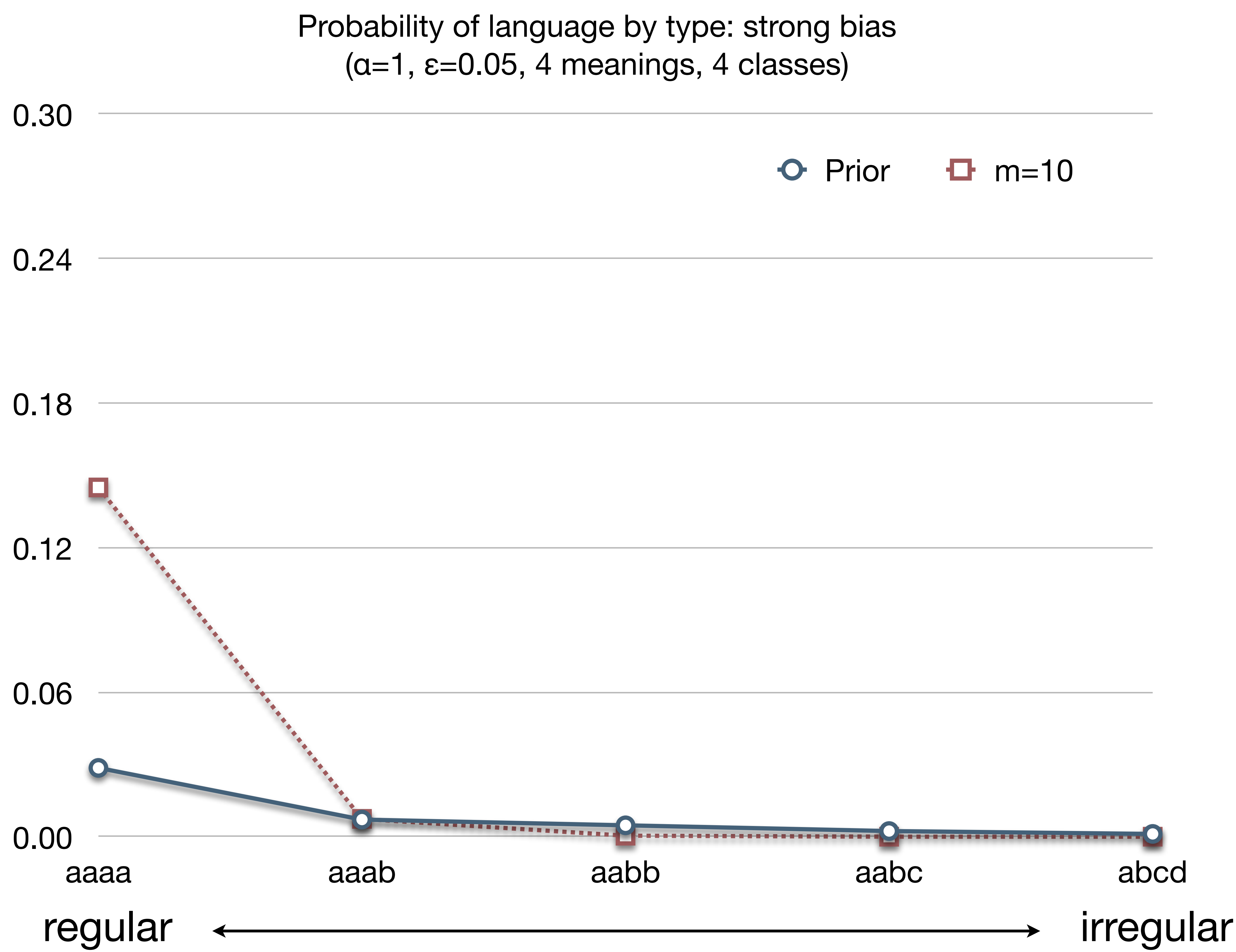


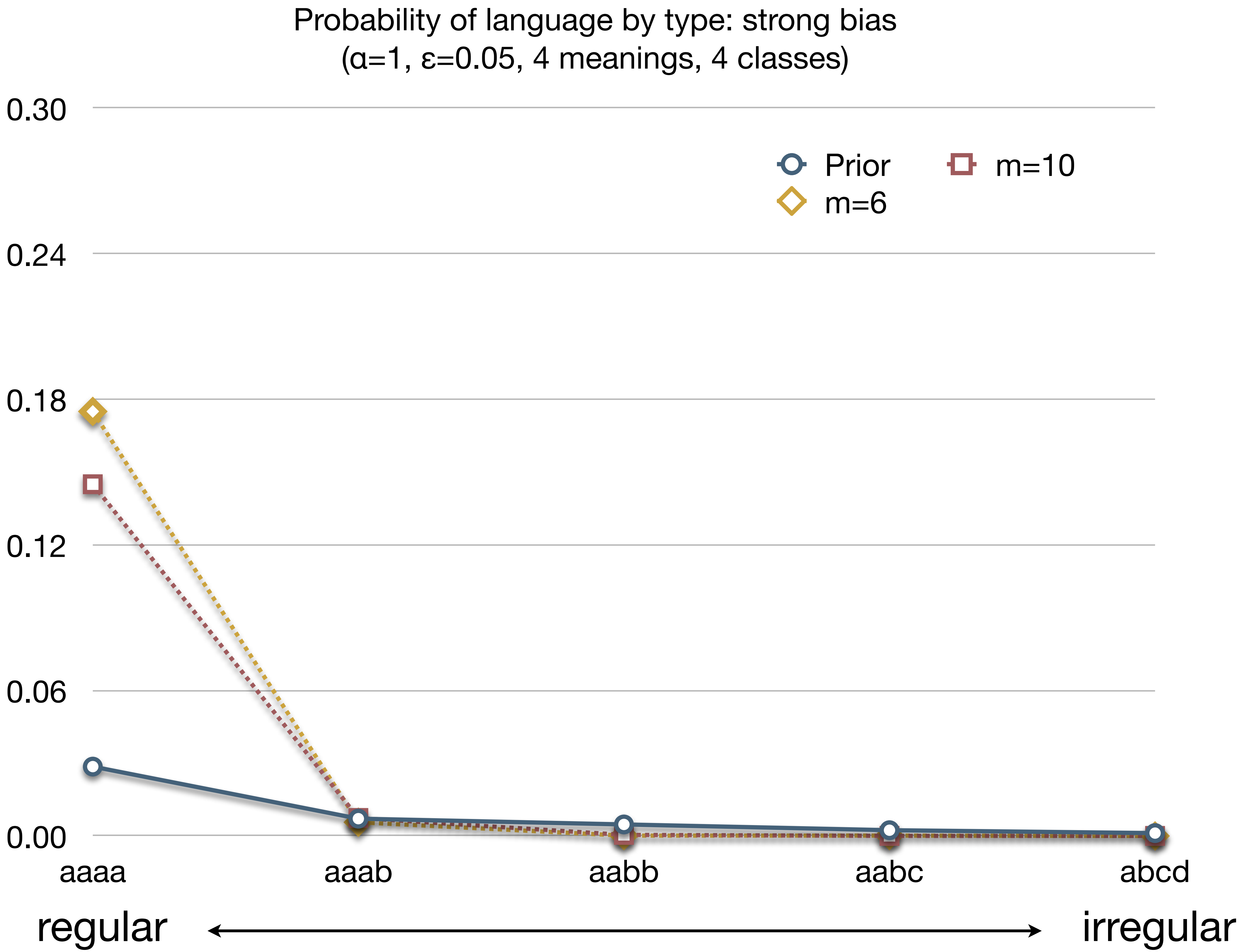

Probability of language by type: strong bias
(α=1, ε=0.05, 4 meanings, 4 classes)
0.30
0.24
0.18
0.12
0.06
0.00
Prior
aaaa
aaab
aabb
aabc
abcd
regular
irregular

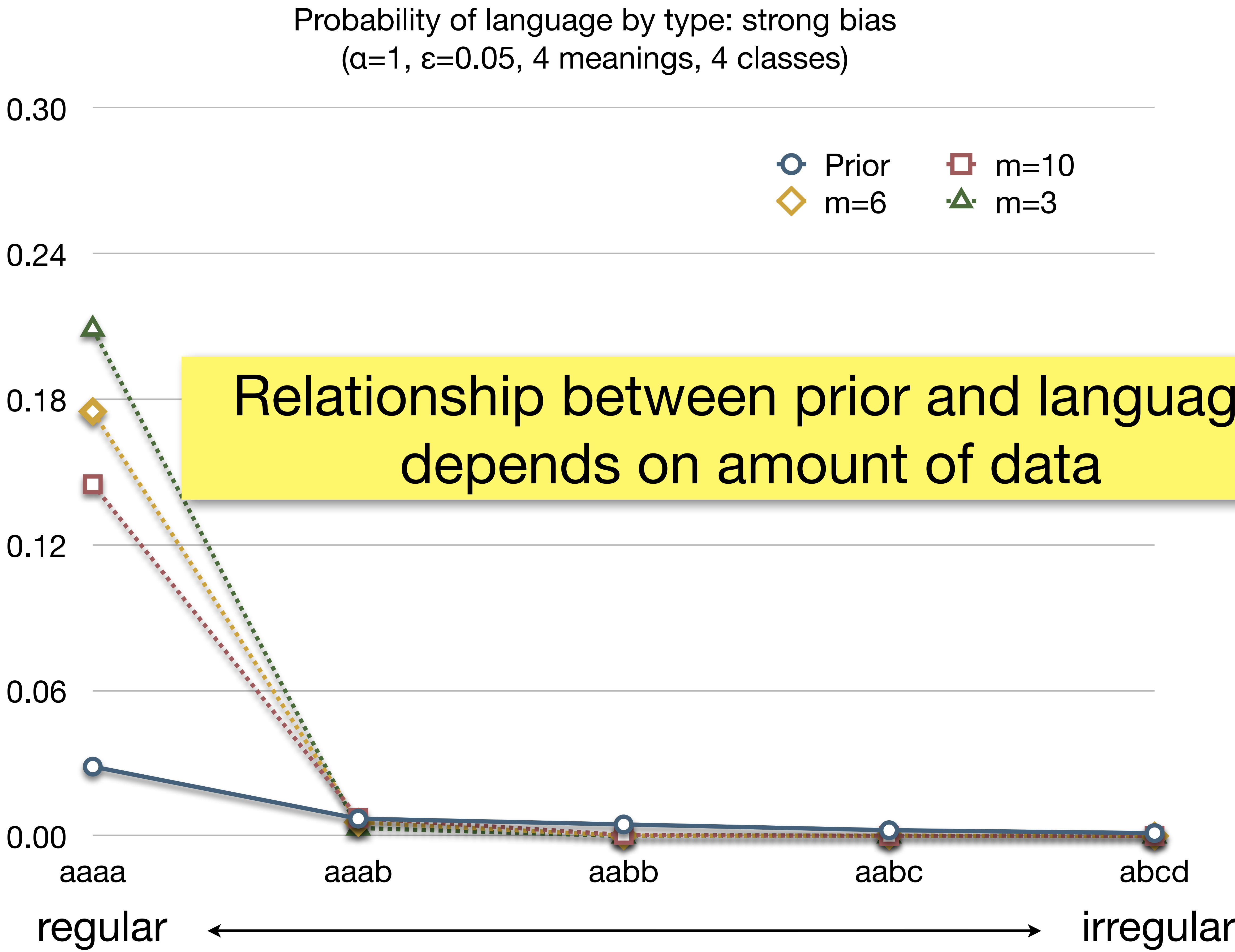


Probability of language by type: strong bias
(α=1, ε=0.05, 4 meanings, 4 classes)
0.30
0.24
0.18
0.12
0.06
0.00
Prior
If we use sampling, the stationary distribution looks exactly the same as the prior
aaaa
aaab
aabb
aabc
abcd
regular
irregular

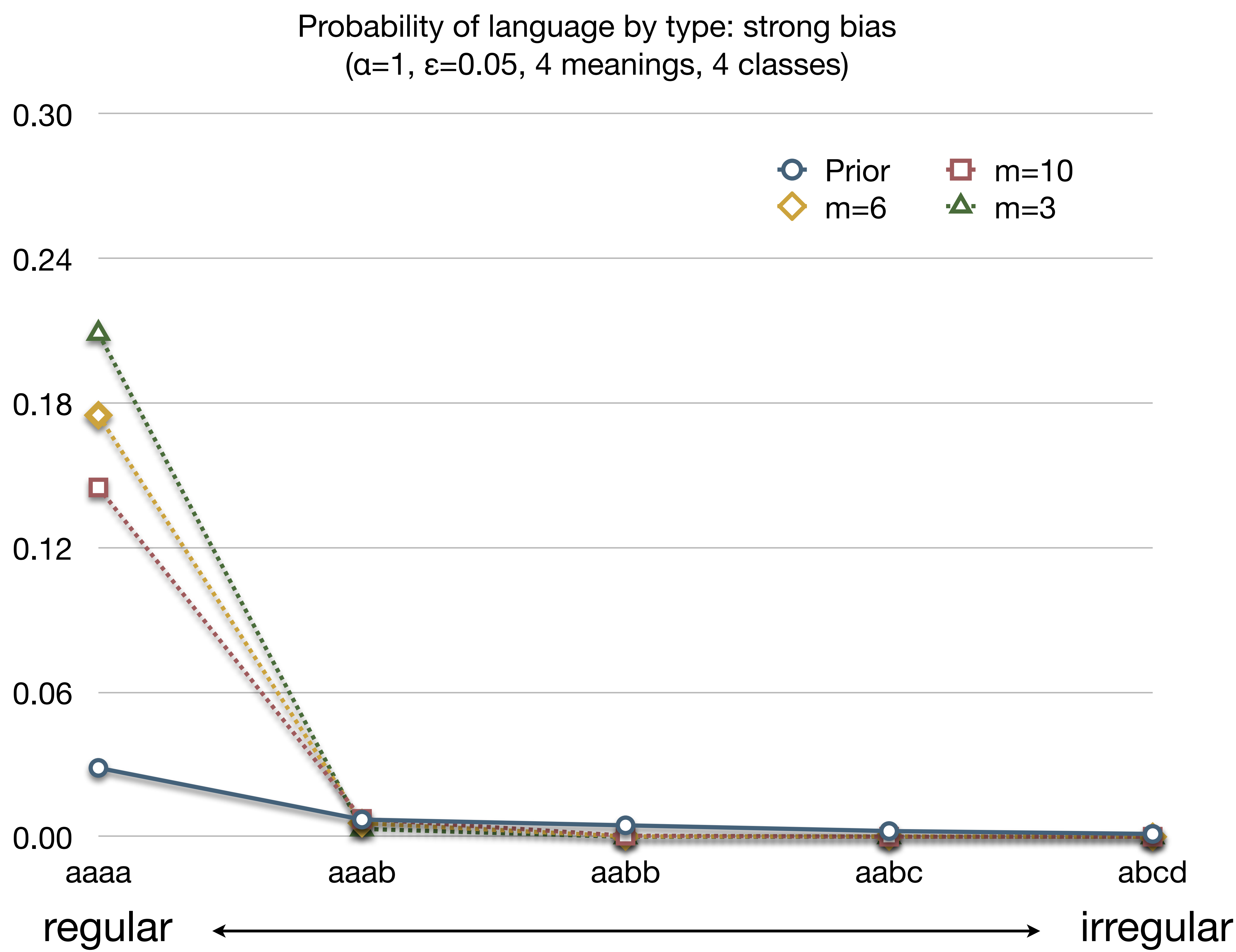

Probability of language by type: strong bias
(α=1, ε=0.05, 4 meanings, 4 classes)
0.30
0.24
0.18
0.12
0.06
0.00
Prior
aaaa
aaab
aabb
aabc
abcd
regular
irregular

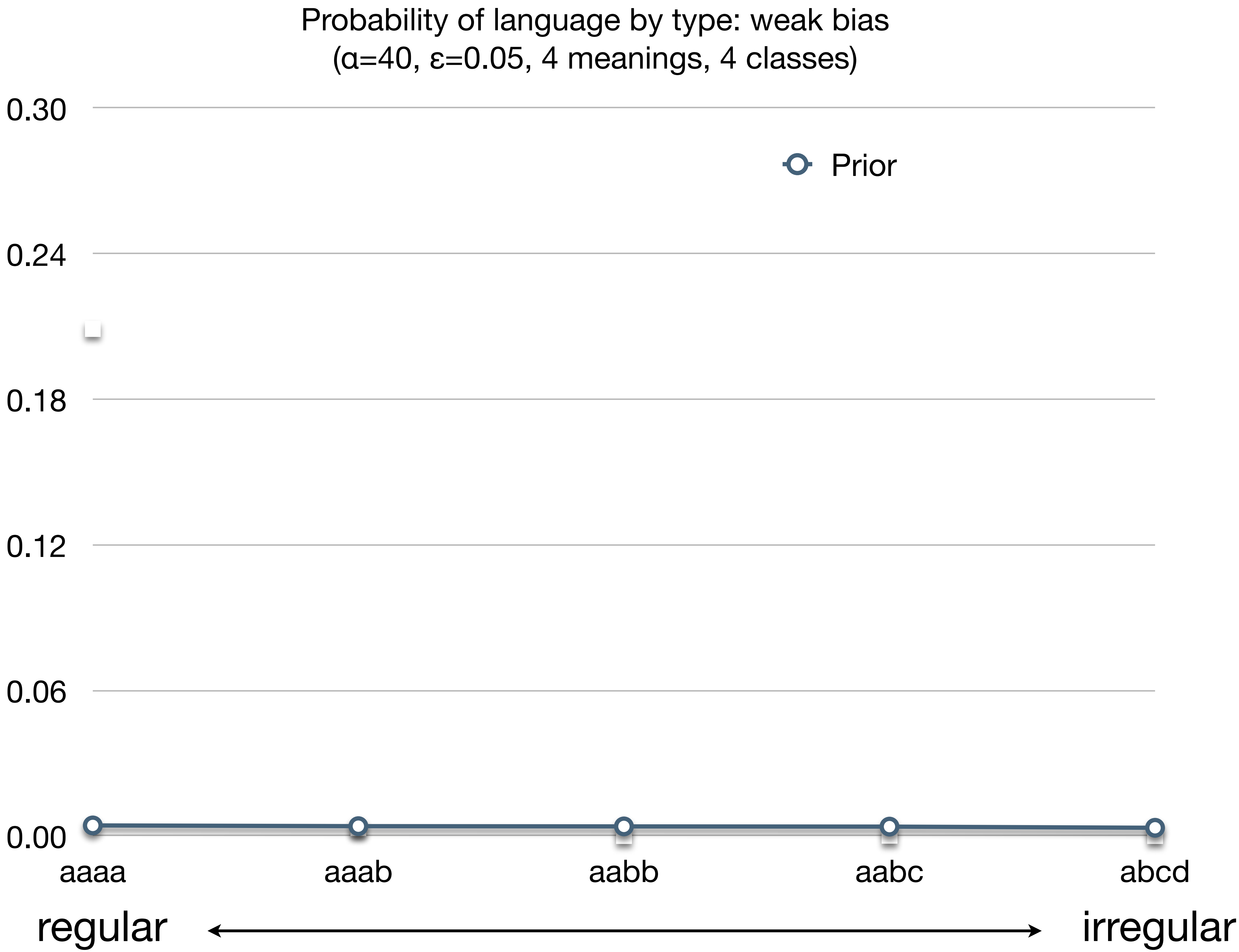

Probability of language by type: strong bias
(α=1, ε=0.05, 4 meanings, 4 classes)
0.30
0.24
0.18
0.12
0.06
0.00
Prior
m=10
aaaa
aaab
aabb
aabc
abcd
regular
irregular

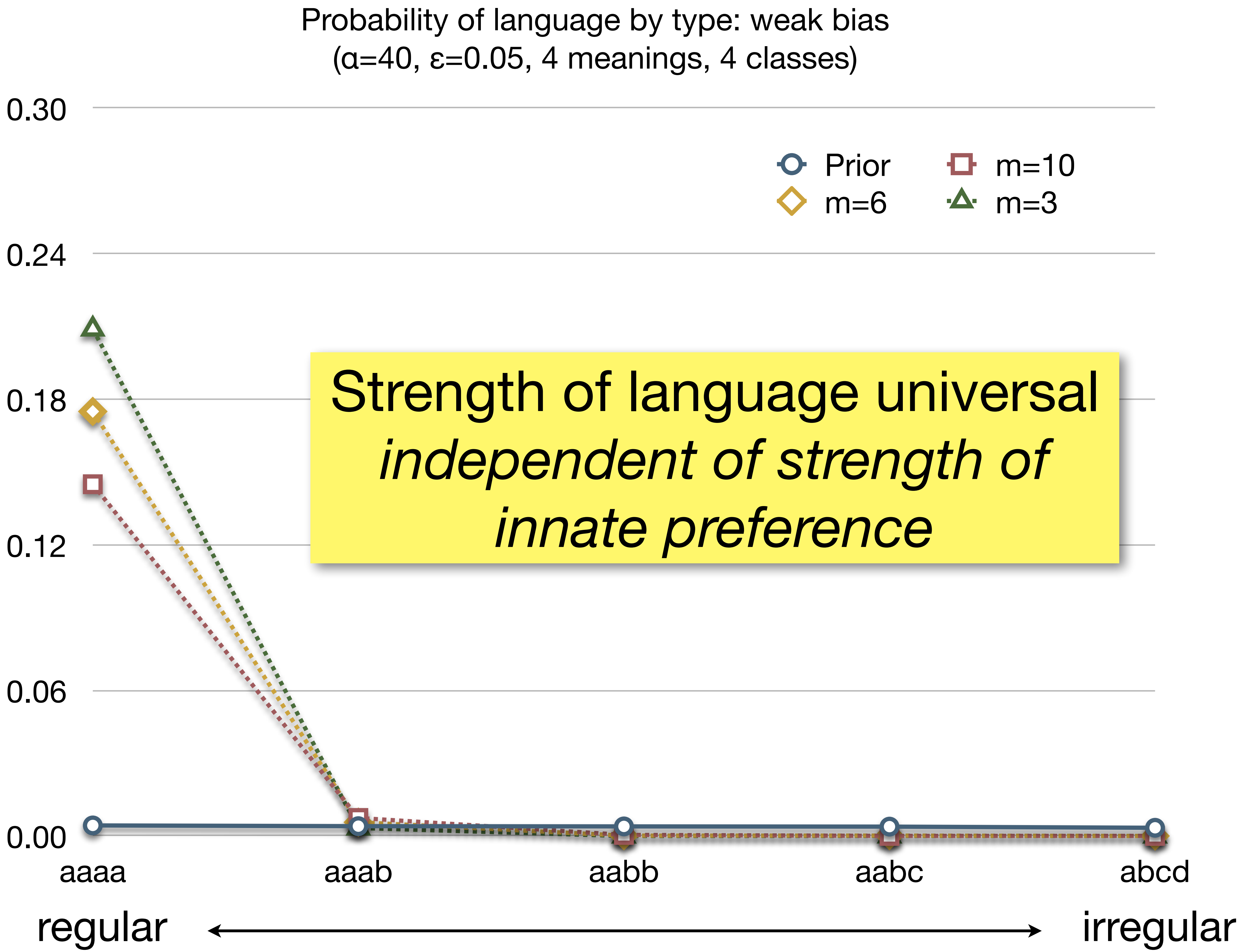

Probability of language by type: strong bias
(α=1, ε=0.05, 4 meanings, 4 classes)
0.30
0.24
0.18
0.12
0.06
0.00
Prior
m=10
m=6
aaaa
aaab
aabb
aabc
abcd
regular
irregular



Relationship between prior and language depends on amount of data


Probability of language by type: strong bias
(α=1, ε=0.05, 4 meanings, 4 classes)
0.30
0.24
0.18
0.12
0.06
0.00
Prior
m=10
m=6
m=3
aaaa
aaab
aabb
aabc
abcd
regular
irregular


Probability of language by type: weak bias
(α=40, ε=0.05, 4 meanings, 4 classes)
0.30
0.24
0.18
0.12
0.06
0.00
Prior
aaaa
aaab
aabb
aabc
abcd
regular
irregular



Probability of language by type: weak bias
(α=40, ε=0.05, 4 meanings, 4 classes)
0.30
0.24
0.18
0.12
0.06
0.00
Prior
m=6
m=10
m=3
Strength of language universal
independent of strength of innate preference
aaaa
aaab
aabb
aabc
abcd
regular
irregular

Conclusions

- Iterated Bayesian Learning allows us to more precisely understand the relationship between learning bias and eventual language structure
- If (and only if) you assume learners pick the best hypothesis, then cultural evolution does a lot of work for you
  - Very weak innate biases are all that’s needed to explain strong linguistic universals
  - If we see strong universals in language, then **we can’t necessarily assume that these arise from strong innate constraints!**
- So which is right, sampling or MAP? Strong constraints or weak biases? Next week, we’ll turn to biological evolution to provide an answer!

References

- Griffiths, T. L. and Kalish, M. L. (2007) Language evolution by iterated learning with Bayesian agents. Cognitive Science, 31, 441-480.

- Kirby, S., Dowman, M. and Griffiths, T. (2007) Innateness and culture in the evolution of language. Proceedings of the National Academy of Sciences, 104, 5241-5245.