Simple Bayesian Algorithm for Top-m Arm Identification

Master Thesis

Kevin Klein

November 16, 2019

Advisors: Johannes Kirschner, Mojmír Mutný, Prof. Dr. Andreas Krause

Department of Computer Science, ETH Zürich

Abstract

This example thesis briefly shows the main features of our thesis style,
and how to use it for your purposes.

Contents

Chapter 1

Introduction

What is the relationship between the optimal allocation and the algorithm? How does this statement translate to real world applications? What is particular about the algorithm being Bayesian? What are constaints on prior and posterior distributions? Are we in a frequentist or Bayesian setting?

Chapter 2

Background

2.1 Chernoff's 2-player game

Tbd whether relevant.

2.2 Bandits

The so-called stochastic multi-armed bandits is a general model to simulate decision-making in uncertain environments. In particular, one assumes a set of options or arms to choose from, each choice leading to an outcome, also referred to as reward. In general, one assumes many sequential selections among the set of arms. It is typically assumed that the outcome of the selection of an individual arm follows a fixed but unknown probability distribution.

Within the bandits model, the concern revolves around which arm to choose next. Allocation strategies tackling this question address either of two problems: explore-exploit or pure-explore.

The former concerns itself with maximizing the *cummulative reward*. This means that one attempts to maximize the *sum of the rewards obtained* through all arm selections. Naturally, starting off without any knowledge about the underlying distributions, this involves both exploration of arm qualities as well as exploitation of gathered knowledge.

The latter revolves around *simple reward*. This problem consists of seeking to maximize the reward one obtains if one leverages the current knowledge for *another* draw. This implies that the focus lies on *identification* of high-quality candidates, where quality is usually assessed by high means. Hence it is part of the task to make estimated best and true best equal as well as obtaining a high *confidence* on this statement.

The bandits model can be used for all processes involving sequential decision making. Yet, an important simplification is the assumption of the distributions being fixed over time. This assumption might be more or less representative of the true, to-be-modelled processes.

The explore-exploit problem is applied in Recommender Systems, ad campaigns, in the context of Reinforcement Learning and Robotics.

For real-world applications of the pure-exploration bandit please refer to 2.5.1.

There exist many specialized versions of the Bandits model, such as contextual bandits or adversarial bandits.

2.3 Thompson Sampling

Thompson sampling is a sampling strategy often employed for choosing arms in the bandits scenario. By default, it particularly lends itself to the cummulative reward setting. The technique is based on having and updating prior distributions over all arms. In contrast to greedily selecting the arm that empirically maximizes the metric of desire, Thompson sampling suggests to randomly draw beliefs on arms from the prior distributions. Subsequently, it selects the arm which maximizes the metric of desire on the sampled belief. With many repetitions, this leads to each arm being sampled proportionally to its likelihood of maximizing the metric of desire according to current belief. Intuitively, this is very appealing for the cummulative reward scenario as it creates a natural balance between exploration and exploitation.

2.4 Notation

We assume k arms to choose from and denote the set of possible arms as $[k]$. Subsets $S \subset [k]$ are assumed to be of size $m < k$.

A possible set of means for the arms is denoted as the k-dimensional vector θ, where every mean can lie between 0 and 1. We also refer to such a θ as parameter.

Finding ourselves in a frequentist setting, we assume an underlying true mean of the arms refer to this as θ^*. Moreover, this ground truth implies a true best arm l^* and true top-m arms S^*. When referring to an individual arm in the top-m case, we proceed to use j for arms in S^*, i for arms not in S^* and l for arms of which this knowledge does not exist.

Given the constraint that the means lie between 0 and 1, there are infinitely many possible parameter vectors which make arm $l \in [k]$ the best one under

θ or $S \subset [k]$ top-m under θ. Hence we group such θ in the following way:

$$\Theta := [0,1]^k \tag{2.1}$$

$$\Theta_{1,l} := \{\theta \in \Theta | l = \arg\max_{l' \in [k]} \theta_{l'}\} \tag{2.2}$$

$$\Theta_{m,S} := \{\theta \in \Theta | S = \text{top-}m(\theta)\} \tag{2.3}$$

$$= \{\theta \in \Theta | \min_{j_1 \in S} \theta_{j_1} > \max_{j_2 \notin S} \theta_{j_2}\} \Theta_{m,l} = \{\theta \in \Theta | l \text{ is in top-}m \text{ under } \theta\}$$

$$\tag{2.4}$$

$$\tag{2.5}$$

where top-m returns the m highest values, i.e. means, from its argument θ.

After having made n observations of rewards, bundled in D_n, Π_n expresses a posterior distribution with density π_n over elements of Θ. E.g. for sets $\Theta_{m,S}$, we have:

$$\Pi_n(\Theta_{m,S}) := \int_{\theta \in \Theta_{m,S}} \pi_n(\theta) d\theta \tag{2.6}$$

$$= \Pr[S \text{ is top-}m | D_n] \tag{2.7}$$

Clearly, it holds that $\Pi_n(\Theta) = 1$. As a shorthand for the top-m case, we will use:

$$\alpha_{n,l} := \Pi_n(\Theta_{m,l}) \tag{2.8}$$

$$\alpha_{n,S} := \Pi_n(\Theta_{m,S}) \tag{2.9}$$

We are mostly interested in strategies that allocate measurement effort over the arms in a randomized fashion, i.e. according to a probability distribution. We refer to such an allocation or distribution as ψ, defining the probability of each arm being sampled. Naturally, it holds that $\sum_{l \in [k]} \psi_l = 1$.

The n-th sampled arm is denoted as l_n. For adaptive strategies, the allocation can change after every sample. After n sample we refer to the allocation s as ψ_n and therefore after n samples we have $\psi_{n,l} = \Pr[l_n = l]$ for arm l. In the case of adaptive strategies, we might as well be interested in average allocation up to a certain sample n. We write:

$$\bar{\psi}_{n,l} := \frac{\sum_{n'=1}^{n} \psi_{n',l}}{n} \tag{2.10}$$

We use the notation $d(\theta_1 || \theta_2)$ to represent the KL-divergence between θ_1 and θ_2.

2.5 Best Arm Identification

2.5.1 Model

Best Arm identification implies disregarding the sum of the rewards encountered while sampling. Rather, the goal is to efficiently gather information maximizing the confidence of the suggestion made *after* the sampling phase. In other words, there is a purely explorative phase, also referred to as 'experiment' which is then typically followed by a purely exploitative phase, having committed to an option. For this reason, Best Arm Identification is also known as 'Experiment Design'.

Quite naturally, all other aspects being controlled for, requiring fewer samples is preferable. Hence one can formulate the goal in two ways: maximize confidence for a fix amount of samples or minimize amount of samples for a fixed confidence.

Confidence can mean and be evaluated in different ways. For Bayesian algorithms, it is possible to quantify the confidence the model has in the currently most favourable looking candidate. This quantity corresponds to the mass a posterior puts on parameters favouring said candidate. Another approach to measure confidence stems from the realm of Probably approximately correct learning, or PAC in short. In the PAC context, one desires to make a statement of the sort $\Pr[\text{output of algoritm is } \varepsilon\text{-correct}] \geq 1 - \delta$, where ε-correctness requires an explicit definition. In this case, under the toleration of ε, $1 - \delta$ can be thought of as confidence.

Real-world applications of Best Arm identification include:

- Physical simulations: Given a collection of designs, e.g. the bodywork of a car, determine e.g. aerodynamic properties of the cars. Elaborate simulations can come with significant resource demands, such as compute power and direct and indirect costs of losing time. Coming to the same conclusion, consisting of the superiority of one of the designs, with fewer simulations can be crucial.

- Crop selection: Experimenting different crop types in a given growing environment measuring yields, generate a recommendation for that same growing environment.

"ML: drive generation of own data instead of" -¿ it is about acquisition

2.5.2 Optimal allocation

We start off by describing what it means for an allocation to be optimal and by enumerating some of its properties. We do not present a constructive allocation, rather we assume knowledge about the underlying truth to provide a best possible bound. Intuitively it is not possible to match the

performance of that optimal allocation without the knowledge of the underlying truth. Hence the desired statement is to say that a concrete adative allocation *converges* to the optimal allocation.

As mentioned in 2.5.1, the goal consists of maximizing confidence, i.e. the mass the posterior lays onto the true best arm. Oberve that for any allocation sampling each arm infinitely many times, this quantity will tend towards 1. What makes the optimal allocation optimal is the rate at which the posterior of the true arm convergence towards 1.

Note that the optimal allocation assumes knowledge about the underlying true value and therefore does not need to be adaptive. It can be framed as the following thought experiment: Assume you know the true underlying means. Yet your adversary doesn't trust your 'knowledge'. He only trusts the sample rewards that he can observe for himself. Now it is your task to leverage your knowledge about the true means to sample most the arms in a way that convinces the adversary as quickly as possible.

Rate of convergence

Russo [2] shows that the rate at which the posterior Π_n of the set parameters under which the true best arm l^* is optimal Θ^* cannot be faster than the following:

$$\Pi_n(\Theta_{l^*}) = 1 - \Pi_n(\Theta_{l^*}^c) = 1 - \exp\{-n\Gamma^*\} \tag{2.11}$$

$$\Gamma^* = \max_{\psi} \min_{\theta \in \Theta_{l^*}^c} \sum_{l \in [k]} \psi_l d(\theta_l^* || \theta_l) \tag{2.12}$$

where n corresponds to the number of samples acquired.

The allocation ψ maximizing this quantity is what we refer to as optimal allocation.

Defining properties

Russo's underlying idea is that the optimal allocation gathers equal *evidence*, e.g. compared to having equal effort for a uniform distribution. This notion of evidence relies on the comparison between the true best against all other arms. It takes both their respective true means as well as sampling frequencies into consideration. For an allocation ψ, he defines

$$C_i(\psi_{l^*}, \psi_i) := \min_{x \in \mathbb{R}} \psi_{l^*} d(\theta_{l^*}^* || x) + \psi_i d(\theta_i^* || x) \tag{2.13}$$

and goes on to show that the optimal allocation ψ^* is identified by fulfilling the condition

$$\forall i_1, i_2 \neq l^* : C_{i_1}(\psi_{l^*}, \psi_{i_1}) = C_{i_2}(\psi_{l^*}, \psi_{i_2}) \tag{2.14}$$

Moreover, he shows that the optimal allocation is unique.

2.5.3 A Constrained Optimal Allocation

In order to bridge the gap between algorithm and optimal allocation, Russo introduces the concept of *constraining* the optimal allocation. His algorithm naturally implies the constraint that in the limit, β of the measurement effort is allocated to the true best arm.

He is able to show that his algorithm's allocation converges to the overall optimal allocation by showing that

- The algorithm's average allocation $\bar{\psi}_n$ converges to the optimal allocation under the constraint $\psi_{l*}^* = \beta$

- The hyperparameter β can be tuned to equal the overall optimal value.

The optimal convergence exponent under said constraint becomes:

$$\Gamma_\beta^* = \max_{\psi:\psi_{l*}=\beta} \min_{\theta \in \Theta_{l*}^c} \sum_{l \in [k]} \psi_l d(\theta_l^* || \theta_l) \tag{2.15}$$

2.5.4 Top-Two Thompson Sampling algorithm

Russo proposes different algorithms that satisfy aforementioned theoretical results, yet suggests that one of them outperfoms the other empirically: Top-Two Thompson Sampling (TTTS) algorithm.

The main idea behind TTTS 1 is to repeat drawing with Thompson sampling from the prior until two different candidates are obtained. This illustrates how Thompson sampling, usually only truly useful for exploit-explore settings can be used for pure-exploit: some of its focus is shifted towards inferior-looking candidates. Among those two candidates, the former is picked with probability β, explaining the provenance of the hyperparameter.

Algorithm 1 Given a prior Π_n at step n

$\hat{\theta} \sim \Pi_n$
$l_1 := \arg\max(\hat{\theta})$
repeat
  $\hat{\theta} \sim \Pi_n$
  $l_2 := \arg\max(\hat{\theta})$
until $l_1 \neq l_2$
$B \sim Bernoulli(\beta)$
if $B = 1$ **then**
  $l_n := l_1$
else
  $l_n := l_2$
Play l_n, observe reward and update priors

The attentive reader might wonder how exactly prior/posteriors are updated. This is addressed in the empirical section 4.2.

2.5.5 Alternative approaches

What are alternative approaches to Russo's? How do they compare against Russo's?

2.6 Top-*m* Arm Identification

Top-*m* Arm Identification is a generalization of Best Arm Identification. The objective is to identify the set of arms, with cardinality *m*, which contains the *m* best arms.

Applications of top-*m* arm identification are similary to those mentioned in 2.5.1. Natural explications for desiring the identification *m* instead of 1 high-quality option can be diversification or regulation.

2.6.1 Current approaches: LUCB

How are those methods evaluated? What are possible qualitative shortcomings of those methods? What are possible quantitative shortcomings of those methods?

Confidence estimation 1

Confidence estimation 2

Confidence estimation 3

Chapter 3

Charactarizing the Constrained Optimal Top-m Allocation

Analogously to 2.5.2, we define the optimal top-m allocation by its convergence rate of parameters reflecting the true mean. Moreover, just as in Russo's top-1 case, the algorithm we will propose introduces a constraint. We will also put the optimal allocation under that constraint with the idea being that this is but a hyperparameter that can be optimized over.

First, we introduce some of Russo's results that also apply in our scenario. Then we will introduce some general properties, followed by a complete characterization of the optimal allocation. The proofs of the statements are provided in a separate section. Our main focus lies again on the posterior mass put of the true top-m arms.

How can C be interpreted? How does this tie in with Chernoff's statements? How does this compare to the top-1 case? What's an example of an optimal allocation? How would those statements look like without the constraint?

Russo proves a proposition about the posterior convergence rate about general parameter sets $\tilde{\Theta}$.

Proposition 3.1 (Russo: Proposition 5) *For any open set $\tilde{\Theta} \subset \Theta$ and average allocation $\bar{\psi}_n$*

$$\Pi_n(\tilde{\Theta}) \doteq \exp\{-n \inf_{\theta \in \tilde{\Theta}} \sum_{l \in [k]} \bar{\psi}_{n,l} d(\theta_l^* || \theta_l)\}$$

Instead of approaching $\Pi_n(\Theta_{m,S^*}^c)$ directly, we express it with via $\Theta_{m,l}$ and its relaxation $\tilde{\Theta}_l$.

$$\bar{\Theta}_i = \{\theta \in \Theta | \text{top-}m(\theta, S^* \cup \{i\}) \neq S^*\} \tag{3.1}$$

Oberve that we have $\bar{\Theta}_i \supsetneq \Theta_i$. Moreover we have a direct relationship between Θ_{m,S^*}^c and $\bar{\Theta}_i$:

3. Characterizing the Constrained Optimal Top-m Allocation

Lemma 3.2

$$\Theta_{m,S^*}^c = \bigcup_{i \notin S^*} \bar{\Theta}_i$$

Leveraging this relationship of the sets allows us to bridge the gap between the posterior of Θ_{m,S^*}^c and the posterior of individual sets $\bar{\Theta}_i$. Note the transition from a union of sets to a minimum over sets permitted by the usage of the $\doteq$ equality.

Lemma 3.3 *If $\alpha_{n,S^*} \to 1$, then*

$$\Pi_n(\Theta_{m,S^*}^c) \doteq \max_{i \notin S^*} \Pi_n(\bar{\Theta}_i) \doteq 1 - \min_{j \in S^*} \Pi_n(\Theta_j) \tag{3.2}$$

Plugging $\bar{\Theta}_i$ into 3.1 leaves us with a sum of KL divergences. We seek to simplify this sum to individual terms just after defining:

$$C_{j,i}(\psi_j, \psi_i) = \min_{x \in \mathbb{R}} \psi_j d(\theta_j^* || x) + \psi_i d(\theta_i^* || x) \tag{3.3}$$

In particular, $C_{j,i}$ can be thought of as evidence that a j is superior to a i.

Lemma 3.4 *For any $i \notin S^*$ and any allocation ψ,*

$$\min_{\theta \in \bar{\Theta}_i} \sum_{l \in [k]} \psi_l d(\theta_l^* || \theta_l) = \min_{j \in S^*} C_{j,i}(\psi_j, \psi_i) \tag{3.4}$$

Having established this relationship to a minimum over $C_{j,i}$, we state two properties of the latter function.

Lemma 3.5 *Each $\min_{j \in S^*} C_{j,i}(\psi_j, \psi_i)$ is a concave function.*

Lemma 3.6 *Given a $j \in S^*$, the solution to the minimization problem (3.22) is $\bar{\theta} \in \mathbb{R}$, satisfying:*

$$A'(\bar{\theta}) = \frac{\psi_j A'(\theta_j^*) + \psi_i A'(\theta_i^*)}{\psi_j + \psi_i}$$

where $A'(\theta)$ is the mean observation under θ. Therefore

$$\min_{j \in S^*} C_{j,i}(\psi_j, \psi_i) = \min_{j \in S^*} \psi_j d(\theta_j^* || \bar{\theta}) + \psi_i d(\theta_i^* || \bar{\theta})$$

3.1 Proofs

Proof (3.2)

$$\Theta^c_{m,S^*} = \{\theta \in \Theta \mid \min_{j \in S^*} \theta_j > \max_{i \notin S^*} \theta_i\}^c \tag{3.5}$$

$$= \{\theta \in \Theta \mid \max_{i \notin S^*} \theta_i \geq \min_{j \in S^*} \theta_j\} \tag{3.6}$$

$$= \{\theta \in \Theta \mid \exists i \notin S^* : \theta_i \geq \min_{j \in S^*} \theta_j\} \tag{3.7}$$

$$= \{\theta \in \Theta \mid \exists i \notin S^* : \text{top-}m(\theta, S^* \cup \{i\}) \neq S^*\} \tag{3.8}$$

$$= \bigcup_{i \notin S^*} \{\theta \in \Theta \mid \text{top-}m(\theta, S^* \cup \{i\}) \neq S^*\} \tag{3.9}$$

$$= \bigcup_{i \notin S^*} \bar{\Theta}_i \tag{3.10}$$

$\square$

Proof (3.3)

$$\max_{i \notin S^*} \Pi_n(\bar{\Theta}_i) \leq \Pi_n(\Theta^c_{m,S^*}) \leq (k-m) \max_{i \notin S^*} \Pi_n(\bar{\Theta}_i) \tag{3.11}$$

This equation follows from (3.2) and the fact that the union has an additive effect with respect to the density Π_n. There are $k - m$ possible is and each single one leads to a set, whose density is bounded by the maximal density of all such sets. Asymptotically this gives

$$\Pi_n(\Theta^c_{m,S^*}) \doteq \max_{i \notin S^*} \Pi_n(\bar{\Theta}_i) \tag{3.12}$$

For S^*, we have:

$$\Theta^c_{m,S^*} = \Theta - \bigcap_{j \in S^*} \Theta_j \tag{3.13}$$

$$1 - \min_{j \in S^*} \Pi(\Theta_j) \leq \Pi(\Theta^c_{m,S^*}) \leq 1 - \min_{j \in S^*}(\Pi(\Theta_j))^m \tag{3.14}$$

$$\lim_{n \to \infty} \frac{1}{n} \log\left(\frac{\min_{j \in S^*} \Pi(\Theta_j)}{\min_{j \in S^*}(\Pi(\Theta_j))^m}\right) = \lim_{n \to \infty} \frac{-(m-1)}{n} \log(\min_{j \in S^*} \Pi(\Theta_j)) \tag{3.15}$$

Observe that $\min_{j \in S^*} \Pi(\Theta_j) \geq \alpha_{n,S^*} \to 1$. Hence the limit of the fraction goes to 0 and we have $\min_{j \in S^*} \Pi(\Theta_j) \doteq \min_{j \in S^*}(\Pi(\Theta_j))^m$. Again, by the Squeeze theorem it follows that

$$\Pi_n(\Theta^c_{m,S^*}) \doteq 1 - \min_{j \in S^*} \Pi_n(\Theta_j) \tag{3.16}$$

$\square$

Proof (3.4)

$$\min_{\theta \in \bar{\Theta}_i} D_\psi(\theta^* || \theta) = \min_{\theta \in \bar{\Theta}_i} \sum_{j=1}^{k} \psi_j d(\theta_j^* || \theta_j) \tag{3.17}$$

$$= \min_{\theta \in \bar{\Theta}_i} \sum_{j \in S^*} \psi_j d(\theta_j^* || \theta_j) + \psi_i d(\theta_i^* || \theta_i) + \sum_{j \notin S^* \cup \{i\}} \psi_j d(\theta_j^* || \theta_j)$$
$$\tag{3.18}$$

$$= \min_{\theta \in \bar{\Theta}_i} \sum_{j \in S^*} \psi_j d(\theta_j^* || \theta_j) + \psi_i d(\theta_i^* || \theta_i) \tag{3.19}$$

$$= \min_{j \in S^*} \min_{\theta \in \bar{\Theta}_i} \psi_j d(\theta_j^* || \theta_j) + \psi_i d(\theta_i^* || \theta_i) \tag{3.20}$$

$$= \min_{j \in S^*} \min_{\theta \in \bar{\Theta}_i} \psi_j d(\theta_j^* || \theta_j) + \psi_i d(\theta_i^* || \theta_j) \tag{3.21}$$

$$= \min_{j \in S^*} \min_{x \in \mathbb{R}} \psi_j d(\theta_j^* || x) + \psi_i d(\theta_i^* || x) \tag{3.22}$$

$$= \min_{j \in S^*} C_{j,i}(\psi_j, \psi_i) \tag{3.23}$$

(3.19) follows from noting that for any feasible θ, we can define an alternative θ' s.t. $\theta_i' = \theta_i$, $\theta_j' = \theta_j$ for all $j \in S^*$ and $\theta_j' = \theta_j^*$ for all $j \notin S^* \cup \{i\}$. For such a θ', all terms with $j \notin S^* \cup \{i\}$ are zero while all others terms remain unchanged. Hence the minimum occurs with such a θ'. Importantly, θ' remains feasible according to current definitions of $\bar{\Theta}_i$, i.e. $\theta' \in \bar{\Theta}_i$.

(3.20) follows from a similar observation: only a single arm from S^* needs to be inferior to arm i under θ. Recall that the terms of the individual arms do not influence each other. This implies that the minimization will gravitate towards setting all but one arm from S^* in θ to the their true value - as the KL-divergence is minimized for those values. Hence the terms of all but one arm from S^* will be cancelled out by the minimization. As i remains superior to one arm in S^*, we have top-$m(\theta, S^* \cup \{i\}) \neq S^*$. Thereby such a θ is feasible according to $\bar{\Theta}_i$.

(3.21) follows from the same argument Russo employs. The monotonicity of the KL-Divergence combined with the possibility of $\theta_i = \theta_j$ tells us that the minimum will be reached in the case of equality.

(3.22) follows from observing that our minimization over θ has reduced to a minimization over θ_j. The latter is a one-dimensional real. □

Proof (??) By Russo's Lemma 2 (c) we know that $g(x, (\psi_{S^*}, \psi_i)) = \min_{x \in \mathbb{R}} \psi_j d(\theta_j^* || x) + \psi_i d(\theta_i^* || x)$ is concave in $(x, (\psi_{S^*}, \psi_i))$. As S^* is a concave, non-empty set (note that we can always permute arm indices such that it becomes concave) and $g(x, (\psi_{S^*}, \psi_j)) \geq 0 > -\infty$, Boyd and Vandenberghe's [1] (3.2.5) condition for conservation of concavity are met. □

Proof (3.6) Russo's equation (16) in Section D tells us that:

$$d(\theta||\theta') = (\theta - \theta')A'(\theta) - A(\theta) + A(\theta')$$

Appyling this identity to (3.22) for a given j gives us:

$$\psi_j d(\theta_j^*||x) + \psi_i d(\theta_i^*||x) \tag{3.24}$$
$$= \psi_j((\theta_j^* - x)A'(\theta_j^*) - A(\theta_j^*) + A(x)) + \psi_i((\theta_i^* - x)A'(\theta_i^*) - A(\theta_i^*) + A(x)) \tag{3.25}$$
$$= -x(\psi_j A'(\theta_j^*) + \psi_i A'(\theta_i^*)) + A(x)(\psi_j + \psi_i) + c \tag{3.26}$$

Where c is independent of x. As we seeks to minimize this quantity wrt. x, we take the derivative of it wrt x and set it to 0. This gives us:

$$A'(x) = \frac{\psi_j A'(\theta_j^*) + \psi_i A'(\theta_i^*)}{\psi_j + \psi_i} \qquad \square$$

Algorithm

4.1 Analysis/Properties

4.2 Empirical behaviour

What true distributions are assumed? What prior and posterior distributions are assumed? How is C computed? How is alpha computed, as it is defined via a huge integral? How is psi computed, as there is no closed form?

4.3 Proofs

Chapter 5

Conclusion

Dummy Appendix

You can defer lengthy calculations that would otherwise only interrupt the flow of your thesis to an appendix.

Bibliography

[1] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

[2] Daniel Russo. Simple bayesian algorithms for best arm identification. *CoRR*, abs/1602.08448, 2016.

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**

With my signature I confirm that
− I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
− I have documented all methods, data and processes truthfully.
− I have not manipulated any data.
− I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.