

Original citation:

Adamczyk, Marek, Sviridenko, Maxim and Ward, Justin (2014) Submodular stochastic probing on matroids. In: STACS '14: 31st International Symposium on Theoretical Aspects of Computer Science, Lyon, France, 5-8 Mar 2014. Published in: STACS '14: 31st International Symposium on Theoretical Aspects of Computer Science pp. 29-40.

Permanent WRAP url:

<http://wrap.warwick.ac.uk/64300>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution- 3.0 Unported (CC BY 3.0) license and may be reused according to the conditions of the license. For more details see <http://creativecommons.org/licenses/by/3.0/>

A note on versions:

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: publications@warwick.ac.uk

warwick**publications**wrap

highlight your research

<http://wrap.warwick.ac.uk/>

Submodular Stochastic Probing on Matroids

Marek Adamczyk^{*1}, Maxim Sviridenko², and Justin Ward³

1 Department of Computer, Control, and Management Engineering, Sapienza University of Rome, Italy, adamczyk@dis.uniroma1.it

2 Department of Computer Science, University of Warwick, United Kingdom, M.I.Sviridenko@warwick.ac.uk

3 Department of Computer Science, University of Warwick, United Kingdom, J.D.Ward@warwick.ac.uk

Abstract

In a *stochastic probing* problem we are given a universe E , where each element $e \in E$ is *active* independently with probability $p_e \in [0, 1]$, and only a *probe* of e can tell us whether it is active or not. On this universe we execute a process that one by one probes elements — if a probed element is active, then we have to include it in the solution, which we gradually construct. Throughout the process we need to obey *inner* constraints on the set of elements taken into the solution, and *outer* constraints on the set of all probed elements. This abstract model was presented by Gupta and Nagarajan [18], and provides a unified view of a number of problems. Thus far all the results in this general framework pertain only to the case in which we are maximizing a linear objective function of the successfully probed elements. In this paper we generalize the stochastic probing problem by considering a monotone submodular objective function. We give a $(1 - 1/e)/(k^{in} + k^{out} + 1)$ -approximation algorithm for the case in which we are given $k^{in} \geq 0$ matroids as inner constraints and $k^{out} \geq 1$ matroids as outer constraints. There are two main ingredients behind this result. First is a previously unpublished stronger bound on the continuous greedy algorithm due to Vondrak [22]. Second is a rounding procedure that also allows us to obtain an improved $1/(k^{in} + k^{out})$ -approximation for linear objective functions.

1998 ACM Subject Classification F. Theory of Computation

Keywords and phrases approximation algorithms, stochastic optimization, submodular optimization, matroids, iterative rounding

Digital Object Identifier 10.4230/LIPIcs.STACS.2014.29

1 Introduction

Uncertainty in input data is a common feature of most practical problems, and research in finding good solutions (both experimental and theoretical) for such problems has a long history dating back to 1950 [6, 11]. We consider adaptive stochastic optimization problems in the framework of Dean et al. [13]. Here the solution is in fact a process, and the optimal one might even require larger than polynomial space to describe. Since the work of Dean et al. a number of such problems were introduced [10, 14, 15, 16, 4, 17, 12]. Gupta and Nagarajan [18] present an abstract framework for a subclass of adaptive stochastic problems giving a unified view for Stochastic Matching [10] and Sequential Posted Pricing [9].

We describe the framework following [18]. We are given a universe E , where each element $e \in E$ is *active* with probability $p_e \in [0, 1]$ independently. The only way to find out if an

* Supported by the ERC StG project PAA1 no. 259515, and by NCN grant N N206 567940.



element is active, is to *probe* it. We call a probe *successful* if an element turns out to be active. On universe E we execute an algorithm that probes the elements one-by-one. If an element is active, the algorithm must add it to the current solution. In this way, the algorithm gradually constructs a solution consisting of active elements.

Here, we consider the case in which we are given constraints on both the elements probed and the elements included in the solution. Formally, suppose that we are given two independence systems of downward-closed sets: an *outer* independence system (E, \mathcal{I}^{out}) restricting the set of elements probed by the algorithm, and an *inner* independence system (E, \mathcal{I}^{in}) , restricting the set of elements taken by the algorithm. We denote by Q^t the set of elements probed in the first t steps of the algorithm, and by S^t the subset of active elements from Q^t . Then, S^t is the partial solution constructed by the first t steps of the algorithm. We require that, at each time t , $Q^t \in \mathcal{I}^{out}$ and $S^t \in \mathcal{I}^{in}$. Thus, at each time t , the element e that we probe must satisfy both $Q^{t-1} \cup \{e\} \in \mathcal{I}^{out}$ and $S^{t-1} \cup \{e\} \in \mathcal{I}^{in}$. Gupta and Nagarajan [18] considered many types of systems \mathcal{I}^{in} and \mathcal{I}^{out} , but we focus only on matroid intersections, i.e. on the special case in which \mathcal{I}^{in} is an intersection of k^{in} matroids $\mathcal{M}_1^{in}, \dots, \mathcal{M}_{k^{in}}^{in}$, and \mathcal{I}^{out} is an intersection of k^{out} matroids $\mathcal{M}_1^{out}, \dots, \mathcal{M}_{k^{out}}^{out}$. We always assume that $k^{out} \geq 1$ and $k^{in} \geq 0$. We assume familiarity with matroid algorithmics (see [20], for example) and, above all, with principles of approximation algorithms (see [21], for example).

Considering submodular objective functions is a common practice in combinatorial optimization as it extends the range of applicability of many methods. So far, the framework of stochastic probing has been used to maximize the expected weight of the solution found by the process. We were given weights $w_e \geq 0$ for $e \in E$ and, if S denotes the solution at the end of a process, the goal was to maximize $\mathbb{E}_S [\sum_{e \in S} w_e]$. We generalize the framework as we consider a monotone submodular function $f : 2^E \mapsto \mathbb{R}_{\geq 0}$, and objective of maximizing $\mathbb{E}_S [f(S)]$.

1.1 Our results

Our result is a new algorithm for stochastic probing problem based on iterative randomized rounding of linear programs and the *continuous greedy process* introduced by Calinescu et al. [8].

► **Theorem 1.** *An algorithm based on the continuous greedy process and iterative randomized rounding is a $\frac{(1-e^{-1})}{k^{in}+k^{out}+1}$ -approximation for stochastic probing problem with monotone submodular objective function.*

Additionally, we improve the bound of $\frac{1}{4(k^{in}+k^{out})}$ given by Gupta and Nagarajan [18] in the case of a linear objective.

► **Theorem 2.** *The iterative randomized rounding algorithm is a $\frac{1}{k^{in}+k^{out}}$ -approximation for the stochastic probing problem with a linear objective function.*

1.2 Applications

On-line dating and kidney exchange [10]

Consider an online dating service. For each pair of users, machine learning algorithms estimate the probability that they will form a happy couple. However, only after a pair meets do we know for sure if they were successfully matched (and together leave the dating service). Users have individual patience numbers that bound how many unsuccessful dates

they are willing to go on until they will leave the dating service forever. The objective of the service is to maximize the number of successfully matched couples.

To model this as a stochastic probing problem, users are represented as vertices V of a graph $G = (V, E)$, where edges represent matched couples. Set E of edges is our universe on which we make probes, with p_e being the probability that a couple $e = (u_1, u_2)$ forms a happy couple after a date. The inner constraints are matching constraints — a user can be in at most one couple —, and outer constraints are b -matching — we can probe at most $t(u)$ edges adjacent to user u , where $t(u)$ denotes the patience of u . Both inner and outer constraints are intersections of two matroids for bipartite graphs. In similar way we can model kidney exchanges.

In weighted bipartite case Theorem 2 gives a $1/4$ -approximation. Even though b -matchings in general graphs are not intersections of two matroids, we are able to exploit the matching structure to give the same factor- $1/4$ approximation. Since the technique is very similar to the case of intersection of two matroids, we omit the proof. This matches the current-best bound for general graphs of Bansal et al. [5], who also give a $1/3$ -approximation in the bipartite case.

Bayesian mechanism design [18]

Consider the following mechanism design problem. There are n agents and a single seller providing a certain service. Agent's i value for receiving service is v_i , drawn independently from a distribution D_i over set $\{0, 1, \dots, B\}$. The valuation v_i is private, but the distribution D_i is known. The seller can provide service only for a subset of agents that belongs to system $\mathcal{I} \in 2^{[n]}$, which specifies feasibility constraints. A mechanism accepts bids of agents, decides on subset of agents to serve, and sets individual prices for the service. A mechanism is called *truthful* if agents bid their true valuations. Myerson's theory of virtual valuations yields *truthful* mechanisms that maximize the expected revenue of a seller, although they sometimes might be impractical. On the other hand, practical mechanisms are often non-truthful. The Sequential Posted Pricing Mechanism (SPM) introduced by Chawla et al. [9] gives a nice trade-off — it is truthful, simple to implement, and gives near-optimal revenue. An SPM offers each agent a “take-it-or-leave-it” price for the service. Since after a refusal a service won't be provided, it is easy to see that an SPM is a truthful mechanism.

To see an SPM as a stochastic probing problem, we consider a universe $E = [n] \times \{0, 1, \dots, B\}$, where element (i, c) represents an offer of price c to agent i . The probability that i accepts the offer is $\mathbb{P}[v_i \geq c]$, and seller earns c then. Obviously, we can make only one offer to an agent, so outer constraints are given by a partition matroid; making at most one probe per agent also overcomes the problem that probes of $(i, 1), \dots, (i, B)$ are not independent. The inner constraints on universe $[n] \times \{0, 1, \dots, B\}$ are simply induced by constraints \mathcal{I} on $[n]$.

Gupta and Nagarajan [18] give an LP relaxation for any single-seller Bayesian mechanism design problem. Provided that we can optimize over $\mathcal{P}(\mathcal{I})$, the LP can be used to construct an efficient SPM. Moreover, the approximation guarantee of the constructed SPM is with respect to the optimal mechanism, which need not be an SPM.

In the case constraints \mathcal{I} are an intersection of k matroids the resulting SPM is a $\frac{1}{4(k+1)}$ -approximation [18]. Here, we give an improved approximation algorithm with a factor- $\frac{1}{k+1}$ guarantee. In particular, when $k = 1$ we match [9, 19] with $1/2$ -approximation.

1.3 Related work

The stochastic matching problem with applications to online dating and kidney exchange was introduced by Chen et al. [10], where authors proved a $1/4$ -approximation of a greedy strategy for unweighted case. The authors also show that the simple greedy approach gives no constant approximation in the weighted case. Their bound was later improved to $1/2$ by Adamczyk [1]. As noted in our discussion of applications, Bansal et al. [5] gave $1/3$ and $1/4$ -approximations for weighted stochastic matching in bipartite and general graphs, respectively.

Sequential Posted Pricing mechanisms were investigated first by Chawla et al. [9], followed by Yan [24], and Kleinberg and Weinberg [19]. Gupta and Nagarajan [18] were first to propose looking at SPM from the point of view of stochastic adaptive problems.

Asadpour et al. [4] were first to consider a stochastic adaptive problem with submodular objective function. In our terms, they considered only a single outer matroid constraint.

Work of Calinescu et al. [8] provides the tools for submodular functions we use in this paper. The method of [24] was based on “correlation gap” [3], something we address implicitly in Subsection 2.2.2.

2 Preliminaries

For set $S \subseteq E$ and element $e \in E$ we use $S + e$ to denote $S \cup \{e\}$, and $S - e$ to denote $S \setminus \{e\}$. For set $S \subseteq E$ we shall denote by $\mathbf{1}_S$ a characteristic vector of set S , and for a single element e we shall write $\mathbf{1}_e$ instead of $\mathbf{1}_{\{e\}}$. For random event \mathcal{A} we shall denote by $\chi[\mathcal{A}]$ a 0-1 random variable that indicates whether \mathcal{A} occurred. The optimal strategy will be denoted by OPT , and we shall denote the expected objective value of its outcome as $\mathbb{E}[OPT]$.

2.1 Matroids and polytopes

Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid, where E is the universe of elements and $\mathcal{I} \subseteq 2^E$ is a family of independent sets. For element $e \in E$, we shall denote the matroid \mathcal{M} with e contracted by \mathcal{M}/e , i.e. $\mathcal{M}/e = (E - e, \{S \subseteq E - e \mid S + e \in \mathcal{I}\})$.

The following lemma is a slightly modified¹ basis exchange lemma, which can be found in [20].

► **Lemma 3.** *Let $A, B \in \mathcal{I}$ and $|A| = |B|$. There exists a bijection $\phi : A \mapsto B$ such that: 1) $\phi(e) = e$ for every $e \in A \cap B$, 2) $B - \phi(e) + e \in \mathcal{I}$.*

We shall use the following corollary, where we consider independent sets of possibly different sizes.

► **Corollary 4.** *Let $A, B \in \mathcal{I}$. We can find assignment $\phi_{A,B} : A \mapsto B \cup \{\perp\}$ such that:*

1. $\phi_{A,B}(e) = e$ for every $e \in A \cap B$,
2. for each $f \in B$ there exists at most one $e \in A$ for which $\phi_{A,B}(e) = f$,
3. for $e \in A \setminus B$, if $\phi_{A,B}(e) = \perp$ then $B + e \in \mathcal{I}$, otherwise $B - \phi_{A,B}(e) + e \in \mathcal{I}$.

We consider optimization over *matroid polytopes* which have the general form $\mathcal{P}(\mathcal{M}) = \{x \in \mathbb{R}_{\geq 0}^E \mid \forall A \in \mathcal{I} \sum_{e \in A} x_e \leq r_{\mathcal{M}}(A)\}$, where $r_{\mathcal{M}}$ is the rank function of \mathcal{M} . We know [20]

¹ The difference is that we do not assume that A, B are bases, but independent sets of the same size.

that the matroid polytope $\mathcal{P}(\mathcal{M})$ is equivalent to the convex hull of $\{\mathbf{1}_A \mid A \in \mathcal{I}\}$, i.e. characteristic vectors of all independent sets of \mathcal{M} . Thus, we can represent any $x \in \mathcal{P}(\mathcal{M})$ as $x = \sum_{i=1}^m \beta_i \cdot \mathbf{1}_{B_i}$, where $B_1, \dots, B_m \in \mathcal{I}$ and β_1, \dots, β_m are non-negative weights such that $\sum_{i=1}^m \beta_i = 1$. We shall call sets B_1, \dots, B_m a *support* of x in $\mathcal{P}(\mathcal{M})$.

2.2 Submodular functions

2.2.1 Multilinear extension

A set function $f : 2^E \mapsto \mathbb{R}_{\geq 0}$ is *submodular*, if for any two subsets $S, T \subseteq E$ we have $f(S \cup T) + f(S \cap T) \leq f(S) + f(T)$. We call function f *monotone*, if for any two subsets $S \subseteq T \subseteq E : f(S) \leq f(T)$. For a set $S \subseteq E$, we let $f_S(A) = f(A \cup S) - f(S)$ denote the marginal increase in f when the set A is added to S . Note that if f is monotone submodular, then so is f_S for all $S \subseteq E$. Moreover, we have $f_S(\emptyset) = 0$ for all $S \subseteq E$, so f_S is normalized. Without loss of generality, we assume also that $f(\emptyset) = 0$.

We consider the *multilinear extension* $F : [0, 1]^E \mapsto \mathbb{R}_{\geq 0}$ of f , whose value at a point $y \in [0, 1]^E$ is given by

$$F(y) = \sum_{A \subseteq E} f(A) \prod_{e \in A} y_e \prod_{e \notin A} (1 - y_e).$$

Note that $F(\mathbf{1}_A) = f(A)$ for any set $A \subseteq E$, so F is an extension of f from discrete domain 2^E into a real domain $[0, 1]^E$. The value $F(y)$ can be interpreted as the expected value of f on a random subset $A \subseteq E$ that is constructed by taking each element $e \in E$ with probability y_e . Following this interpretation, Calinescu et al. [8] show that $F(y)$ can be estimated to any desired accuracy in polynomial time, using a sampling procedure.

Additionally, they show that F has the following properties, which we shall make use of in our analysis:

► **Lemma 5.** *The multilinear extension F is linear along the coordinates, i.e. for any point $x \in [0, 1]^E$, any element $e \in E$, and any $\xi \in [-1, 1]$ such that $x + \xi \cdot \mathbf{1}_e \in [0, 1]^E$, it holds that $F(x + \xi \cdot \mathbf{1}_e) - F(x) = \xi \cdot \frac{\partial F}{\partial y_e}(x)$, where $\frac{\partial F}{\partial y_e}(x)$ is the partial derivative of F in direction y_e at point x .*

► **Lemma 6.** *If $F : [0, 1]^E \mapsto \mathbb{R}$ is a multilinear extension of monotone submodular function $f : 2^E \mapsto \mathbb{R}$, then 1) function F has second partial derivatives everywhere; 2) for each $e \in E$, $\frac{\partial F}{\partial y_e} \geq 0$ everywhere; 3) for any $e_1, e_2 \in E$ (possibly equal), $\frac{\partial^2 F}{\partial y_{e_1} \partial y_{e_2}} \leq 0$, which means that $\frac{\partial F}{\partial y_{e_2}}$ is non-increasing with respect to y_{e_1} .*

2.2.2 Continuous greedy algorithm

In [8] the authors utilized the multilinear extension in order to maximize a submodular monotone function over a matroid constraint. They showed that a *continuous greedy algorithm* finds a $(1 - 1/e)$ -approximate maximum of the above extension F over any downward closed polytope. In the special case of the matroid polytope, they show how to employ the pipage rounding [2] technique to the fractional solution to obtain an integral solution.

Another extension of f studied in [7] is given by:

$$f^+(y) = \max \left\{ \sum_{A \subseteq E} \alpha_A f(A) \mid \sum_{A \subseteq E} \alpha_A \leq 1, \forall A \subseteq E : \alpha_A \geq 0, \forall j \in E : \sum_{A: j \in A} \alpha_A \leq y_j \right\}.$$

Intuitively, the solution $(\alpha_A)_{A \subseteq E}$ above represents the distribution over 2^E that maximizes the value $\mathbb{E}[f(A)]$ subject to the constraint that its marginal values satisfy $\mathbb{P}[i \in A] \leq y_i$. The value $f^+(y)$ is then the expected value of $\mathbb{E}[f(A)]$ under this distribution, while the value of $F(y)$ is the value of $\mathbb{E}[f(A)]$ under the particular distribution that places each element i in A independently. However, the following allows us to relate the value of F on the solution of the continuous greedy algorithm to the optimal value of the relaxation f^+ .

► **Lemma 7.** *Let f be a submodular function with multilinear extension F , and let \mathcal{P} be any downward closed polytope. Then, the solution $x \in \mathcal{P}$ produced by the continuous greedy algorithm satisfies $F(x) \geq (1 - 1/e) \max_{y \in \mathcal{P}} f^+(y)$.*

This follows from a simple modification of the continuous greedy analysis, given by Vondrák [22].

2.3 Overview of the iterative randomized rounding approach

We now give a description of the general rounding approach that we employ in both the linear and submodular case. In each case, we formulate a mathematical programming relaxation of the following general form

$$\max_{x \in [0,1]^E} \{g(x) \mid \forall j \in [k^{in}] : p \cdot x \in \mathcal{P}(\mathcal{M}_j^{in}); \forall j \in [k^{out}] : x \in \mathcal{P}(\mathcal{M}_j^{out})\} \quad (1)$$

with $p \in [0,1]^E$ being the vector of probabilities. Here $g : [0,1]^E \mapsto \mathbb{R}_{\geq 0}$ is an objective function chosen so that the optimal value of (1) can be used to bound the expected value of an optimal policy for the given instance using the following lemma. Note that our program will always have constraints as given in (1), only the objective function g changes between the linear and monotone submodular cases.

► **Lemma 8.** *Let OPT be the optimal feasible strategy for some stochastic probing problem in our general setting, and define $x_e = \mathbb{P}[OPT \text{ probes } e]$. Then, $x = (x_e)_{e \in E}$ is a feasible solution to the related relaxation of the form (1).*

Proof. Since OPT is a feasible strategy, the set of elements Q probed by any execution of OPT is always an independent set of each outer matroid $\mathcal{M} = (E, \mathcal{I}_j^{out})$, i.e. $\forall j \in [k^{out}] Q \in \mathcal{I}_j^{out}$. Thus, for any $j \in [k^{out}]$, the vector $\mathbb{E}[\mathbf{1}_Q] = x$ may be represented as a convex combination of vectors from $\{\mathbf{1}_A \mid A \in \mathcal{I}_j^{out}\}$, and hence $x \in \mathcal{P}(\mathcal{M}_j^{out})$. Analogously, the set of elements S that were successfully probed by OPT satisfy $\forall j \in [k^{in}] S \in \mathcal{I}_j^{in}$ for every possible execution of OPT . Hence, for any $j \in [k^{in}]$ the vector $\mathbb{E}[\mathbf{1}_S] = p \cdot x$ may be represented as a convex combination of vectors from $\{\mathbf{1}_A \mid A \in \mathcal{I}_j^{in}\}$, and hence $x \in \mathcal{P}(\mathcal{M}_j^{in})$. ◀

Suppose that f is the objective function for a given instance of stochastic probing over a universe E of elements. Our algorithm first obtains a solution x^0 to a relaxation of the form (1) using either linear programming or the continuous greedy algorithm. Our algorithm proceeds iteratively, maintaining a current set of constraints, a current fractional solution x , and a current set S of elements that have been successfully probed. Initially, the constraints are as given in (1), $x = x^0$, and $S = \emptyset$. At each step, the algorithm selects single element \bar{e} to probe, then permanently sets $x_{\bar{e}}$ to 0. It then updates the outer constraints, replacing \mathcal{M}_j^{out} with $\mathcal{M}_j^{out}/\bar{e}$ for each $j \in [k^{out}]$. If the probe succeeds, the algorithm adds \bar{e} to S and then updates the inner constraints, replacing \mathcal{M}_j^{in} with $\mathcal{M}_j^{in}/\bar{e}$ for each $j \in [k^{in}]$. Finally, we modify our fractional solution x so that it is feasible for the updated constraints. The algorithm terminates when the current solution $x = 0^E$.

In order to analyze the approximation performance of our algorithm, we keep track of a current potential value z , related to the value of the remaining fractional solution x . Let x^t , z^t , and S^t be the current value of x , z , and S at the beginning of step $t + 1$. We show that initially we have $z^0 = g(x^0) \geq \beta \cdot \mathbb{E}[OPT]$ for some constant $\beta \in [0, 1]$, and then analyze the expected decrease $z^t - z^{t+1}$ at an arbitrary step $t + 1$. We show that for each step we have $\alpha \cdot \mathbb{E}[z^t - z^{t+1}] \leq \mathbb{E}[f(S^{t+1}) - f(S^t)]$, for some $\alpha < 1$. That is, the expected increase in the value of the current solution is at least α times the expected decrease in z . Then, we employ the following Lemma to conclude that the algorithm is an $\alpha\beta$ -approximation in expectation. The proof is based on Doob's optional stopping theorem for martingales. Hence, we need to deploy language from martingale theory, such as stopping time and filtration. See [23] for extended background on martingale theory.

► **Lemma 9.** *Suppose the algorithm runs for τ steps and that $z^0 = g(x^0) \geq \beta \cdot \mathbb{E}[OPT]$, $z^\tau = 0$. Let $(\mathcal{F}_t)_{t \geq 0}$ be the filtration associated with our iterative algorithm, where \mathcal{F}_i represents all information available after the i th iteration. Finally, suppose that in each step in our iterative rounding procedure, $\mathbb{E}[f(S^{t+1}) - f(S^t) | \mathcal{F}_t] \geq \alpha \cdot \mathbb{E}[z^t - z^{t+1} | \mathcal{F}_t]$. Then, the final solution S^τ produced by the algorithm satisfies $\mathbb{E}[f(S^\tau)] \geq \alpha\beta \cdot \mathbb{E}[OPT]$.*

Proof. Let G_{t+1} be the gain $f(S^{t+1}) - f(S^t)$ in f at step $t + 1$, and let L_{t+1} be the corresponding loss $z^t - z^{t+1}$ in z at time $t + 1$. We set $G_0 = L_0 = 0$. Define variable $D_t = G_t - \alpha \cdot L_t$. The sequence of random variables $(D_0 + D_1 + \dots + D_t)_{t \geq 0}$ forms a sub-martingale, i.e.

$$\mathbb{E} \left[\sum_{i=0}^{t+1} D_i \middle| \mathcal{F}_t \right] = \sum_{i=0}^t D_i + \mathbb{E}[G_{t+1} - \alpha \cdot L_{t+1} | \mathcal{F}_t] \geq \sum_{i=0}^t D_i.$$

Let τ be the step in which the algorithm terminates, i.e. $\tau = \min \{t \mid x^t = 0^E\}$. Then, the event $\tau = t$ depends only on $\mathcal{F}_0, \dots, \mathcal{F}_t$, so τ is a stopping time. Also, by the definition of the algorithm $x^\tau = 0^E$. It is easy to verify that all the assumptions of Doob's optional stopping theorem are satisfied, and from this theorem we get that $\mathbb{E}[\sum_{i=0}^\tau D_i] \geq \mathbb{E}[D_0]$. Since $D_0 = 0$, we have

$$0 \leq \mathbb{E} \left[\sum_{i=0}^\tau D_i \right] = \mathbb{E} \left[\sum_{i=0}^\tau G_i - \alpha \cdot \sum_{i=0}^\tau L_i \right] = \mathbb{E} \left[\sum_{i=0}^\tau G_i \right] - \alpha \cdot \mathbb{E} \left[\sum_{i=0}^\tau L_i \right].$$

It remains to note that $\sum_{i=0}^\tau G_i = f(S^\tau)$ is the total gain of the algorithm, so $\mathbb{E}[\sum_{i=0}^\tau G_i] = \mathbb{E}[f(S^\tau)]$. On the other hand, $\sum_{i=0}^\tau L_i = g(x^0) - g(x^\tau) = g(x^0) \geq \beta \cdot \mathbb{E}[OPT]$. ◀

Henceforth, we will implicitly condition on all information \mathcal{F}_t available to the algorithm just before it makes step $t + 1$. That is, when discussing step $t + 1$ of the algorithm, we write shortly $\mathbb{E}[\cdot]$ instead of $\mathbb{E}[\cdot | \mathcal{F}_t]$.

3 Linear stochastic probing

In this setting, we are given a weight w_e and a probability p_e for each element $e \in E$ and $f(S)$ is simply $\sum_{e \in S} w_e$. We consider the relaxation (1) in which $g(x) = f(x)$. Then, Lemma 8 shows that the optimal policy OPT must correspond to some feasible solution x^* of (1). Moreover, because f is linear, $\mathbb{E}[OPT] = \sum_{e \in S} \mathbb{P}[OPT \text{ probes } e] p_e w_e = \sum_{e \in S} x_e p_e w_e = f(x^*)$.

At each step, our algorithm randomly selects an element \bar{e} to probe. Let $\Sigma = \sum_{e \in E} x_e$. Then, our algorithm chooses element e with probability x_e / Σ . As discussed in the previous

overview, it then sets $x_{\bar{e}} = 0$ and carries out the probe, updating the matroid constraints to reflect both the choice of \bar{e} and the probe. Finally, it updates x to obtain a new fractional solution that is feasible in the updated constraints. Note that because x_e is set to 0 after probing e , we will never probe an element e twice.

Let us now describe how to update the current solution x to ensure feasibility in each of the updated matroid constraints. Let \bar{e} be the element that we probed and let $\mathcal{M}_j^{\text{out}}$ be some outer matroid. Currently we have $x \in \mathcal{P}(\mathcal{M}_j^{\text{out}})$ and we must obtain a solution x' so that $x' \in \mathcal{P}(\mathcal{M}_j^{\text{out}}/\bar{e})$. We represent the vector x as a convex combination of independent sets $x = \sum_{i=1}^m \beta_i^{\text{out}} \mathbf{1}_{B_i^{\text{out}}}$, where $B_1^{\text{out}}, \dots, B_m^{\text{out}}$ is the support of x with respect to matroid $\mathcal{M}_j^{\text{out}}$. We obtain $x' \in \mathcal{P}(\mathcal{M}_j^{\text{out}}/\bar{e})$ by replacing each independent set B_b^{out} for which $B_b^{\text{out}} + \bar{e} \notin \mathcal{M}_j^{\text{out}}$ with some other set B_c^{out} such that $B_c^{\text{out}} + \bar{e} \in \mathcal{M}_j^{\text{out}}$. We pick one set B_a^{out} with $\bar{e} \in B_a^{\text{out}}$ to guide the update process. We pick the set $B_a^{\text{out}} \ni \bar{e}$ at random with probability $\beta_a^{\text{out}}/x_{\bar{e}}$ (note that for any element e , $\sum_{a:e \in B_a^{\text{out}}} \beta_a^{\text{out}} = x_e$). For any set $B_b^{\text{out}} : \bar{e} \notin B_b^{\text{out}}$, let $\phi_{a,b}$ be the mapping from B_a^{out} into B_b^{out} from Corollary 4. If $\phi_{a,b}(\bar{e}) = \perp$, or $\phi_{a,b}(\bar{e}) = \bar{e}$, then in fact $B_b^{\text{out}} + \bar{e} \in \mathcal{M}_j^{\text{out}}$, and we can just include B_b^{out} in the support of $\mathcal{M}_j^{\text{out}}/\bar{e}$. Otherwise, we substitute B_b^{out} with $B_b^{\text{out}} - \phi_{a,b}(\bar{e})$ in the support of $(x_e)_{e \in E}$ in $\mathcal{P}(\mathcal{M}_j^{\text{out}}/\bar{e})$, since we know that $B_b^{\text{out}} - \phi_{a,b}(\bar{e}) + \bar{e} \in \mathcal{M}_j^{\text{out}}$.

Similarly, if \bar{e} is successfully probed we must perform a support update for each inner matroid. Here, we proceed as in the case of the outer matroids, except we have $p \cdot x \in \mathcal{M}_j^{\text{in}}$ and must obtain x' such that $p \cdot x' \in \mathcal{M}_j^{\text{in}}/e$. We write $p \cdot x$ as a combination independent sets $p \cdot x = \sum_{i=1}^m \beta_i^{\text{in}} \mathbf{1}_{B_i^{\text{in}}}$, and now choose a random set $B_a^{\text{in}} \ni \bar{e}$ to guide the support update with probability $\beta_a^{\text{in}}/p_{\bar{e}}x_{\bar{e}}$. (note that for any element e , we have $\sum_{a:e \in B_a^{\text{in}}} \beta_a^{\text{in}} = p_e x_e$). As in the previous case, we replace B_b^{in} with $B_b^{\text{in}} - \phi_{a,b}(\bar{e})$ for each base B_b^{in} such that $B_b^{\text{in}} + \bar{e} \notin \mathcal{M}_j^{\text{in}}$.

We now turn to the analysis of the probing algorithm. Suppose that the algorithm runs for τ steps and consider the quantity $z^t = f(x^t)$. Then, $z^0 = f(x^0) \geq \mathbb{E}[OPT]$ and $z^\tau = f(0^E) = 0$, so the conditions of Lemma 9 are satisfied with $\beta = 1$. It remains to bound the expected loss $\mathbb{E}[z^t - z^{t+1}]$ in step $t + 1$. In order to do this, we consider the value $\delta_i = p_i(x_i^t - x_i^{t+1})$ for each $i \in E$. We consider arbitrary step $t + 1$, but we are going to denote x^t by x and x^{t+1} by x' . The decrease δ_i may be caused both by the probing step, in which we set $x'_{\bar{e}}$ to 0, or by the matroid update step, in which we decrease several coordinates of x . Let us first consider the losses due to each matroid update.

► **Lemma 10.** *Let x and x' be the current fractional solution before and after one update for a given outer matroid $\mathcal{M}_j^{\text{out}}$. Then, for each $i \in E$, we have $\mathbb{E}[\delta_i^{\text{out}}] \triangleq \mathbb{E}[p_i(x_i - x'_i)] \leq \frac{1}{\Sigma} (1 - x_i) p_i x_i$.*

Proof. The expectation $\mathbb{E}[\delta_i^{\text{out}}]$ is over the random choice of an element \bar{e} to probe and the random choice of an independent set to guide the update. Let $\mathcal{E}_a^{\text{out}}$ denote the event that some set B_a^{out} is chosen to guide a support update for $\mathcal{M}_j^{\text{out}}$.

In a given step the probability that the set B_a^{out} is chosen to guide the support update is equal to

$$\mathbb{P}[\mathcal{E}_a^{\text{out}}] = \sum_{e \in B_a^{\text{out}}} \frac{x_e \beta_a^{\text{out}}}{\Sigma x_e} = \sum_{e \in B_a^{\text{out}}} \frac{\beta_a^{\text{out}}}{\Sigma} = |B_a^{\text{out}}| \frac{\beta_a^{\text{out}}}{\Sigma}. \quad (2)$$

Moreover, conditioned on the fact B_a^{out} was chosen, the probability that an element $e \in B_a^{\text{out}}$ was probed is uniform over the elements of B_a^{out} :

$$\mathbb{P}[e \text{ probed} \mid \mathcal{E}_a^{\text{out}}] = \mathbb{P}[e \text{ probed} \wedge \mathcal{E}_a^{\text{out}}] / \mathbb{P}[\mathcal{E}_a^{\text{out}}] = \frac{x_e \beta_a^{\text{out}}}{\Sigma x_e} \Big/ |B_a^{\text{out}}| \frac{\beta_a^{\text{out}}}{\Sigma} = \frac{1}{|B_a^{\text{out}}|}. \quad (3)$$

We can write the expected decrease as $\mathbb{E}[\delta_i^{out}] = \sum_{a=1}^m \mathbb{P}[\mathcal{E}_a^{out}] \cdot \mathbb{E}[\delta_i^{out} | B_a^{out}]$. Note that for all $i \in B_a^{out}$, we have $\phi_{a,b}(i) = i$ for every set $B_b^{out} \ni i$. Thus, the support update will not change x_i for any $i \in B_a^{out}$, and so $\sum_{a=1}^m \mathbb{P}[\mathcal{E}_a^{out}] \cdot \mathbb{E}[\delta_i^{out} | \mathcal{E}_a^{out}] = \sum_{a:i \notin B_a^{out}} \mathbb{P}[\mathcal{E}_a^{out}] \cdot \mathbb{E}[\delta_i^{out} | \mathcal{E}_a^{out}]$.

Now let us condition on taking B_a^{out} to guide the support update. Consider a set $B_b^{out} \ni e$. If we remove i from B_b^{out} , and hence decrease $p_i x_i$ by $p_i \beta_b^{out}$, it must be the case that we have chosen to probe the single element $\phi_{a,b}^{-1}(i) \in B_a^{out}$. The probability that we probe this element is $\frac{1}{|B_b^{out}|}$. Hence

$$\begin{aligned} & \sum_{a:i \notin B_a^{out}} \mathbb{P}[\mathcal{E}_a^{out}] \cdot \mathbb{E}[\delta_i^{out} | B_a^{out}] \\ &= \sum_{a:i \notin B_a^{out}} \mathbb{P}[\mathcal{E}_a^{out}] \cdot \left(\sum_{b:i \in B_b^{out}} p_i \beta_b^{out} \cdot \mathbb{P}[\phi_{a,b}^{-1}(i) \text{ is probed} | \mathcal{E}_a^{out}] \right) \\ &\leq \sum_{a:i \notin B_a^{out}} \mathbb{P}[\mathcal{E}_a^{out}] \cdot \left(\sum_{b:i \in B_b^{out}} p_i \beta_b^{out} \cdot \frac{1}{|B_a^{out}|} \right) \\ &= \sum_{a:i \notin B_a^{out}} \mathbb{P}[\mathcal{E}_a^{out}] \cdot \frac{p_i x_i}{|B_a^{out}|} \\ &= \sum_{a:i \notin B_a^{out}} |B_a^{out}| \frac{1}{\Sigma} \beta_a^{out} \cdot \frac{p_i x_i}{|B_a^{out}|} = \frac{1}{\Sigma} \sum_{a:i \notin B_a^{out}} \beta_a^{out} p_i x_i = \frac{1}{\Sigma} (1 - x_i) p_i x_i. \quad \blacktriangleleft \end{aligned}$$

► **Lemma 11.** *Let x be the current fractional solution before and after one update for a given inner matroid \mathcal{M}_j^{in} . Then, for each $i \in E$, we have $\mathbb{E}[\delta_i^{in}] \triangleq \mathbb{E}[p_i(x_i - x'_i)] \leq \frac{1}{\Sigma} (1 - p_i x_i) p_i x_i$.*

Proof. Because we only perform a support update when the probe of a chosen element is successful, the expectation $\mathbb{E}[\delta_i^{in}]$ is over the random result of the probe, as well as the random choice of element \bar{e} to probe and the random choice of a base to guide the update. We proceed as in the case of Lemma 10, now letting \mathcal{E}_a^{in} denote the event that the probe was successful and B_a^{in} is chosen to guide the support update. We have:

$$\begin{aligned} \mathbb{P}[\mathcal{E}_a^{in}] &= \sum_{e \in B_a^{in}} p_e \frac{x_e}{\Sigma} \frac{\beta_a^{in}}{p_e x_e} = \sum_{e \in B_a^{in}} \frac{\beta_a^{in}}{\Sigma} = |B_a^{in}| \frac{\beta_a^{in}}{\Sigma}, \\ \mathbb{P}[e \text{ probed} | \mathcal{E}_a^{in}] &= \mathbb{P}[e \text{ probed} \wedge \mathcal{E}_a^{in}] / \mathbb{P}[\mathcal{E}_a^{in}] = p_e \frac{x_e}{\Sigma} \frac{\beta_a^{in}}{p_e x_e} \Big/ |B_a^{in}| \frac{\beta_a^{in}}{\Sigma} = \frac{1}{|B_a^{in}|}. \end{aligned}$$

By a similar argument as in Lemma 10 we then have that $\mathbb{E}[\delta_i^{in}]$ is at most:

$$\begin{aligned} & \sum_{a:i \notin B_a^{in}} \mathbb{P}[\mathcal{E}_a^{in}] \cdot \left(\sum_{b:i \in B_b^{in}} \beta_b^{in} \cdot \frac{1}{|B_a^{in}|} \right) = \sum_{a:i \notin B_a^{in}} \mathbb{P}[\mathcal{E}_a^{in}] \cdot \frac{p_i x_i}{|B_a^{in}|} \\ &= \sum_{a:i \notin B_a^{in}} |B_a^{in}| \frac{1}{\Sigma} \beta_a^{in} \cdot \frac{p_i x_i}{|B_a^{in}|} = \frac{1}{\Sigma} \sum_{a:i \notin B_a^{in}} \beta_a^{in} p_i x_i = \frac{1}{\Sigma} (1 - p_i x_i) p_i x_i. \quad \blacktriangleleft \end{aligned}$$

We perform the matroid updates sequentially for each of the k^{in} and k^{out} matroids. Note that once we decrease a coordinate x_i to 0, it cannot be altered in any further updates, so no coordinate is ever decreased below 0. Now, we consider the expected decrease

$\mathbb{E}[\delta_i] = \mathbb{E}[p_i(x_i - x'_i)]$ due to both the initial probing step, in which we decrease the probed element's coordinate to 0, and the following matroid updates. We have:

$$\begin{aligned}
\mathbb{E}[\delta_i] &\leq \mathbb{P}[i \text{ probed}]p_ix_i + k^{out}\mathbb{E}[\delta_i^{out}] + k^{in}\mathbb{E}[\delta_i^{in}] \\
&= \frac{1}{\Sigma}p_ix_i^2 + k^{out}\frac{1}{\Sigma}(1-x_i)p_ix_i + k^{in}\frac{1}{\Sigma}(1-p_ix_i)p_ix_i \\
&= \frac{1}{\Sigma}k^{out}p_ix_i - \frac{1}{\Sigma}(k^{out}-1)p_ix_i^2 + \frac{1}{\Sigma}k^{in}p_ix_i - \frac{1}{\Sigma}k^{in}p_i^2x_i^2 \\
&\leq \frac{k^{out} + k^{in}}{\Sigma}p_ix_i.
\end{aligned} \tag{4}$$

Because z^t is a linear function of x^t , the expected total decrease of z in this step is then

$$\mathbb{E}[z^t - z^{t+1}] = \sum_i \mathbb{E}[\delta_i] w_i \leq \frac{k^{out} + k^{in}}{\Sigma} \sum_i p_ix_i w_i.$$

On the other hand, the expected gain in $f(S)$ is $\sum_{e \in E} \mathbb{P}[e \text{ probed}] p_e w_e = \frac{1}{\Sigma} \sum_{e \in E} w_e p_e x_e$. Thus, by Lemma 9 the final solution S^τ produced by the algorithm satisfies $\mathbb{E}[f(S^\tau)] \geq \frac{1}{k^{out} + k^{in}} \mathbb{E}[OPT]$.

4 Submodular stochastic probing

We now consider the case in which we are given a set of elements E each becoming active with probability p_e , and we seek to maximize a given submodular function $f: 2^E \mapsto \mathbb{R}_{\geq 0}$. In this case, we consider the relaxation (1) in which $g(x) = f^+(p \cdot x)$. Then, Lemma 8 shows that the optimal policy OPT must correspond to some feasible solution x^* of (1), where $x_e^* = \mathbb{P}[OPT \text{ probes } e]$, and hence $\mathbb{P}[OPT \text{ takes } e] = p_e x_e^*$. The function $f^+(p \cdot x^*)$ gives the maximum value of $\mathbb{E}_{S \sim \mathcal{D}}[f(S)]$ over all distributions \mathcal{D} satisfying $\mathbb{P}_{S \sim \mathcal{D}}[e \in S] = x_e^* p_e$. Thus, $f^+(p \cdot x^*) \geq \mathbb{E}[OPT]$.

In general, we cannot obtain an optimal solution to this relaxation. Instead, we apply the continuous greedy algorithm to a variant of (1) in which $g(x)$ is given by $F(p \cdot x)$ to obtain an initial solution x^0 . From Lemma 7 we then have $F(p \cdot x^0) \geq (1 - 1/e)f^+(p \cdot x^*) \geq (1 - 1/e)\mathbb{E}[f(OPT)]$.

Given x^0 , our algorithm is exactly the same as in the linear case. However, we must be more careful in our analysis. We define the quantity

$$z^t = F(\mathbf{1}_{S^t} + p \cdot x^t) - F(\mathbf{1}_{S^t})$$

where S^t and x^t are, respectively, the set of successfully probed elements and the current fractional solution at time t . Note that because after probing an element we set its variable to zero, for all elements $i \in S$ we have $x_i = 0$, and so indeed $\mathbf{1}_{S^t} + p \cdot x^t \in [0, 1]^E$. Suppose that the algorithm runs for τ iterations, and note that $z^0 = F(p \cdot x^0) \geq (1 - 1/e)\mathbb{E}[f(OPT)]$ and $z^\tau = F(\mathbf{1}_{S^\tau + p \cdot 0^E}) - F(S^\tau) = 0$, so the conditions of Lemma 9 are satisfied with $\beta = (1 - 1/e)$.

We now analyze the expected decrease $z^t - z^{t+1}$ due to step $t + 1$ of the algorithm. Suppose that the algorithm selects element i to probe. Then, we have $S^{t+1} = S^t + i$ with probability p_i and $S^{t+1} = S^t$ otherwise. Thus, we have

$$\begin{aligned}
\mathbb{E}[z^t - z^{t+1}] &= \mathbb{E}[F(\mathbf{1}_{S^t} + x^t \cdot p) - F(\mathbf{1}_{S^t})] - \mathbb{E}[F(\mathbf{1}_{S^{t+1}} + x^{t+1} \cdot p) - F(\mathbf{1}_{S^{t+1}})] \\
&= \mathbb{E}[F(\mathbf{1}_{S^{t+1}}) - F(\mathbf{1}_{S^t})] + \mathbb{E}[F(\mathbf{1}_{S^t} + x^t \cdot p) - F(\mathbf{1}_{S^{t+1}} + x^{t+1} \cdot p)] \\
&\leq \mathbb{E}[F(\mathbf{1}_{S^{t+1}}) - F(\mathbf{1}_{S^t})] + \mathbb{E}[F(\mathbf{1}_{S^t} + x^t \cdot p) - F(\mathbf{1}_{S^t} + x^{t+1} \cdot p)]
\end{aligned} \tag{5}$$

where in the last line, we have used the fact that $S^{t+1} \geq S^t$ and F is increasing in all directions (Lemma 6). We shall first bound the second expectation in (5). We consider the vector δ of decreases in x , given by $\delta = (x^t - x^{t+1}) \cdot p$. For each $i \in E$, let $w_i = \frac{\partial F}{\partial x_i}(\mathbf{1}_{S^t}) = F(\mathbf{1}_{S^t+i}) - F(\mathbf{1}_{S^t})$. Let $y = \mathbf{1}_{S^t} + x^t \cdot p$, and suppose that we decrease the coordinates of y one at a time to obtain $y - \delta = \mathbf{1}_{S^t} + x^{t+1} \cdot p$, letting y^i be the value of y after the first $i - 1$ coordinates have been decreased.² We then have:

$$\begin{aligned} F(y) - F(y - \delta) &= \sum_i F(y^i) - F(y^{i+1}) = \sum_i F(y^i) - F(y^i - \delta_i \mathbf{1}_i) \\ &= \sum_i \delta_i \frac{\partial F}{\partial x_i}(y^i - \delta_i \mathbf{1}_i) \leq \sum_i \delta_i \frac{\partial F}{\partial x_i}(\mathbf{1}_{S^t}) = \sum_i \delta_i w_i, \end{aligned}$$

where the third equality follows from the fact that F is linear when one coordinate is changed (Lemma 5), while the inequality follows from the fact that the partial derivatives of F are coordinate-wise non-increasing (Lemma 6) and $y^i - \delta_i \mathbf{1}_i \geq \mathbf{1}_{S^t}$ for all i . Thus, we have:

$$\mathbb{E}[F(y) - F(y - \delta)] \leq \mathbb{E}\left[\sum_i \delta_i w_i\right] = \sum_i \mathbb{E}[\delta_i] \cdot w_i \leq \frac{1}{\Sigma} (k^{out} + k^{in}) \sum_i p_i x_i^t w_i,$$

where the last inequality follows, as in the linear case, from inequality (4).

Returning to the first expectation in (5), we note that:

$$\mathbb{E}[F(\mathbf{1}_{S^{t+1}}) - F(\mathbf{1}_{S^t})] = \sum_i \mathbb{P}[i \text{ probed}] p_i (F(S^t + i) - F(S^t)) = \frac{1}{\Sigma} \sum_i p_i x_i^t w_i.$$

Thus, the total expected decrease $\mathbb{E}[z^t - z^{t+1}]$ from one step of our rounding procedure is at most:

$$\frac{1}{\Sigma} \sum_i p_i x_i^t w_i + \frac{1}{\Sigma} \sum_i (k^{out} + k^{in}) p_i x_i^t w_i = (k^{out} + k^{in} + 1) \frac{1}{\Sigma} \sum_i p_i x_i^t w_i.$$

On the other hand, the expected increase of $f(S^{t+1}) - f(S^t)$ in this step is:

$$\frac{1}{\Sigma} \sum_{e \in E} p_e x_e^t (f(S^t + e) - f(S^t)) = \frac{1}{\Sigma} \sum_{e \in E} p_e x_e^t (F(\mathbf{1}_{S^t+e}) - F(\mathbf{1}_{S^t})) = \frac{1}{\Sigma} \sum_{e \in E} p_e x_e^t w_e.$$

Thus, by Lemma 9, the final solution S^τ produced by the algorithm satisfies

$$\mathbb{E}[f(S^\tau)] \geq \left(1 - \frac{1}{e}\right) \left(\frac{1}{k^{out} + k^{in} + 1}\right) \mathbb{E}[OPT].$$

References

- 1 Marek Adamczyk. Improved analysis of the greedy algorithm for stochastic matching. *Inf. Process. Lett.*, 111:731–737, August 2011.
- 2 Alexander A. Ageev and Maxim Sviridenko. Pipage rounding: A new method of constructing algorithms with proven performance guarantee. *J. Comb. Optim.*, 8(3):307–328, 2004.

² With slight abuse of notation, we write x_i for the value of the i th decreased coordinate of x and $\mathbf{1}_i$ for the characteristic vector of this coordinate. That is, we identify an element with its index

- 3 Shipra Agrawal, Yichuan Ding, Amin Saberi, and Yinyu Ye. Correlation robust stochastic optimization. In *SODA*, pages 1087–1096, 2010.
- 4 Arash Asadpour, Hamid Nazerzadeh, and Amin Saberi. Stochastic submodular maximization. In *WINE*, pages 477–489, 2008.
- 5 Nikhil Bansal, Anupam Gupta, Jian Li, Julián Mestre, Viswanath Nagarajan, and Atri Rudra. When LP is the cure for your matching woes: Improved bounds for stochastic matchings. *Algorithmica*, 63(4):733–762, 2012.
- 6 E. M. L. Beale. Linear programming under uncertainty. *Journal of the Royal Statistical Society. Series B*, 17:173–184, 1955.
- 7 Gruia Călinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a submodular set function subject to a matroid constraint (extended abstract). In *IPCO*, pages 182–196, 2007.
- 8 Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a Submodular Set Function Subject to a Matroid Constraint. *SIAM Journal on Computing*, 40(6):1740–1766, 2011.
- 9 Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In *STOC*, pages 311–320, 2010.
- 10 Ning Chen, Nicole Immorlica, Anna R. Karlin, Mohammad Mahdian, and Atri Rudra. Approximating matches made in heaven. In *ICALP*, pages 266–278, 2009.
- 11 G.B. Dantzig. Linear programming under uncertainty. *Management Science*, 1:197–206, 1955.
- 12 Brian C. Dean, Michel X. Goemans, and Jan Vondrák. Adaptivity and approximation for stochastic packing problems. In *SODA*, pages 395–404, 2005.
- 13 Brian C. Dean, Michel X. Goemans, and Jan Vondrák. Approximating the stochastic knapsack problem: The benefit of adaptivity. *Math. Oper. Res.*, 33(4):945–964, 2008.
- 14 Michel X. Goemans and Jan Vondrák. Stochastic covering and adaptivity. In *LATIN*, pages 532–543, 2006.
- 15 Sudipto Guha and Kamesh Munagala. Approximation algorithms for budgeted learning problems. In *STOC*, pages 104–113, 2007.
- 16 Sudipto Guha and Kamesh Munagala. Model-driven optimization using adaptive probes. In *SODA*, pages 308–317, 2007.
- 17 Anupam Gupta, Ravishankar Krishnaswamy, Marco Molinaro, and R. Ravi. Approximation algorithms for correlated knapsacks and non-martingale bandits. In *FOCS*, pages 827–836, 2011.
- 18 Anupam Gupta and Viswanath Nagarajan. A stochastic probing problem with applications. In *IPCO*, pages 205–216, 2013.
- 19 Robert Kleinberg and S. Matthew Weinberg. Matroid prophet inequalities. In *STOC*, pages 123–136, 2012.
- 20 A. Schrijver. *Combinatorial Optimization - Polyhedra and Efficiency*. Springer, 2003.
- 21 Vijay V. Vazirani. *Approximation algorithms*. Springer, 2001.
- 22 Jan Vondrák. Personal correspondence.
- 23 David Williams. *Probability with Martingales*. Cambridge mathematical textbooks. Cambridge University Press, 1991.
- 24 Qiqi Yan. Mechanism design via correlation gap. In *SODA*, pages 710–719, 2011.