Submodular Optimization and Approximation Algorithm

Satoru Iwata
岩田　覚
RIMS, Kyoto University
京都大学数理解析研究所

Outline

- Submodular Functions
  - Examples
  - Discrete Convexity
- Submodular Function Minimization
- Approximation Algorithms
- Submodular Function Maximization
- Approximating Submodular Functions

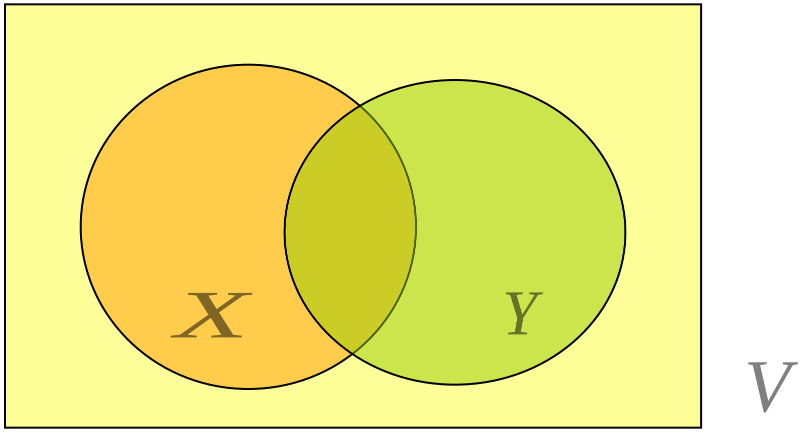
Submodular Functions

V : Finite Set

$f : 2^V \to \mathrm{R}$ $\quad\quad \forall X, Y \subseteq V$

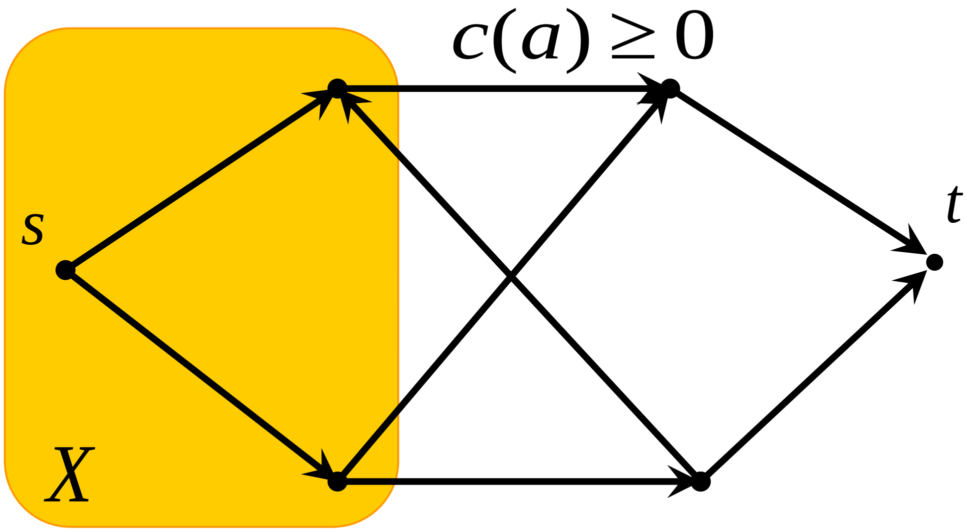
$$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$$

- Cut Capacity Functions
- Matroid Rank Functions
- Entropy Functions



Cut Capacity Function

Cut Capacity $\kappa(X) = \sum\{c(a) \mid a : \text{leaving } X\}$



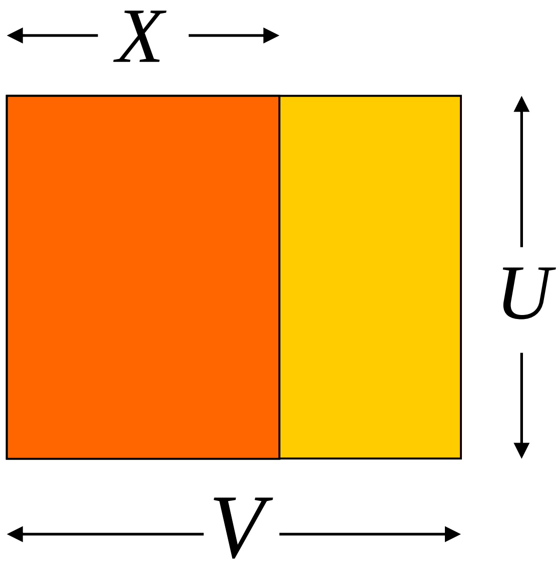
Max Flow Value＝Min Cut Capacity

Matroid Rank Functions

Matrix Rank Function

Whitney (1935)

$$\rho(X) = \operatorname{rank} A[U, X]$$

$A =$ 

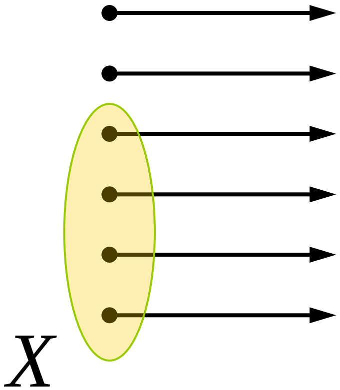
$$\forall X \subseteq V, \rho(X) \leq |X|$$

$$X \subseteq Y \Rightarrow \rho(X) \leq \rho(Y)$$

$\rho :$ Submodular

Entropy Functions

Information Sources



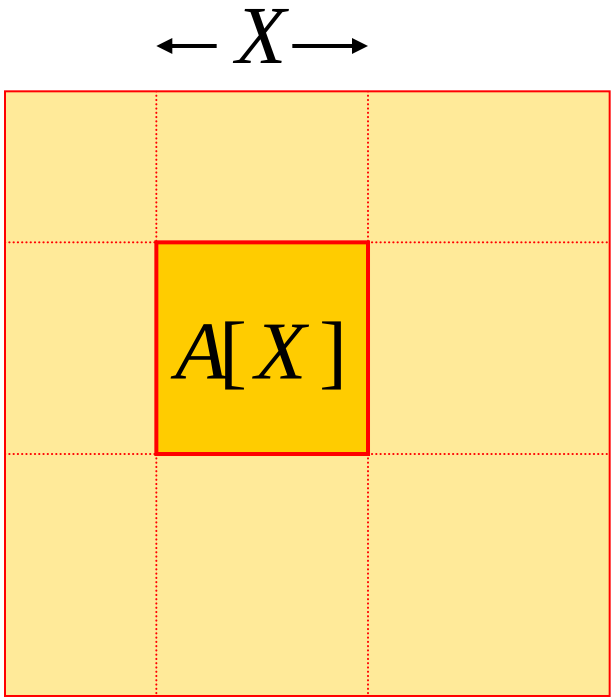
$$h(\phi) = 0$$

$h(X)$: Entropy of the Joint Distribution

$$h(X) + h(Y) \geq h(X \cap Y) + h(X \cup Y)$$

Conditional Mutual Information ≥ 0

Positive Definite Symmetric Matrices

$A =$ 

$$f(\phi) = 0$$

$$f(X) = \log\det A[X]$$

Ky Fan's Inequality

$$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$$

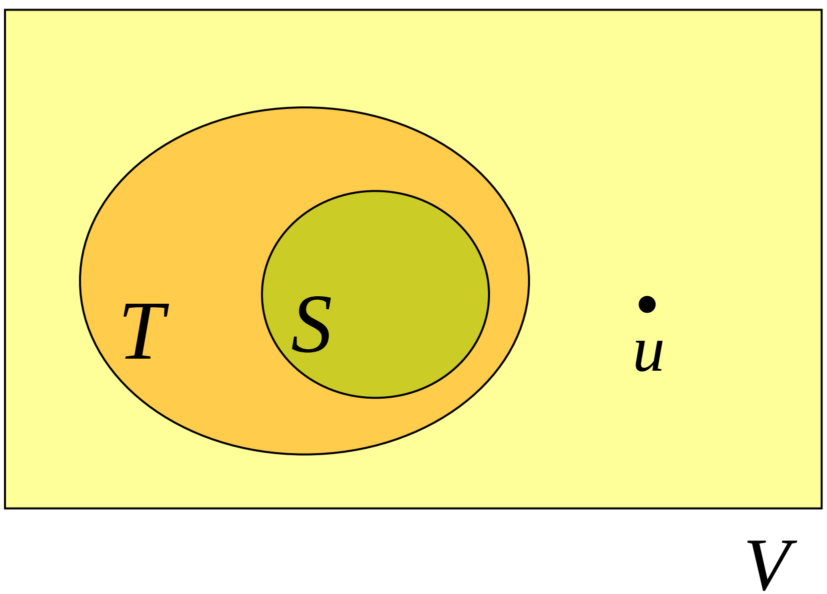
Extension of the Hadamard Inequality

$$\det A \leq \prod_{i \in V} A_{ii}$$

Discrete Concavity

$$S \subseteq T \Rightarrow$$

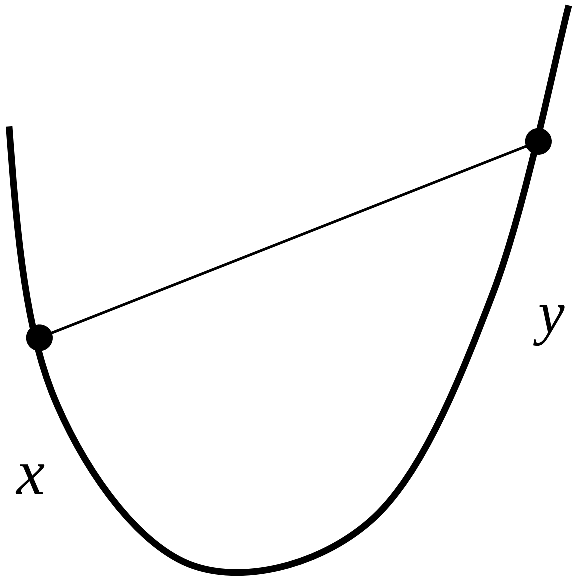
$$f(S \cup \{u\}) - f(S) \geq f(T \cup \{u\}) - f(T)$$



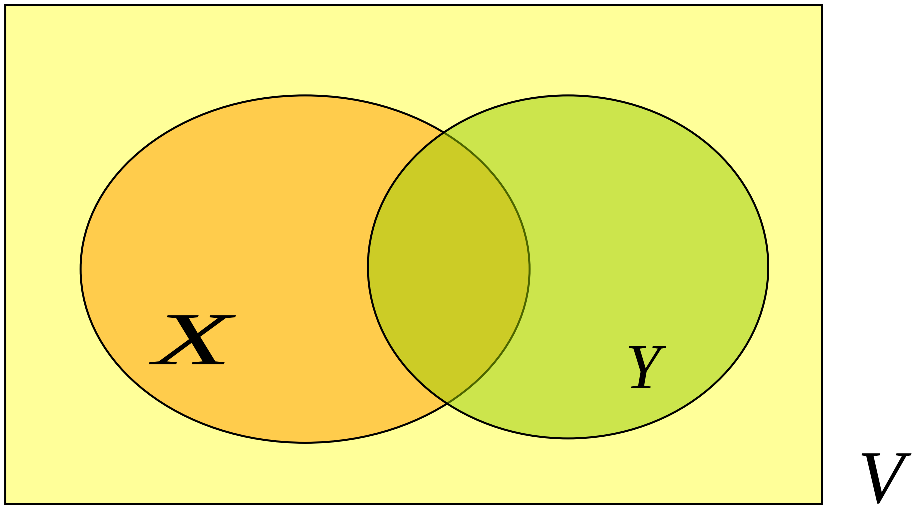
Diminishing Returns

Discrete Convexity

Convex Function

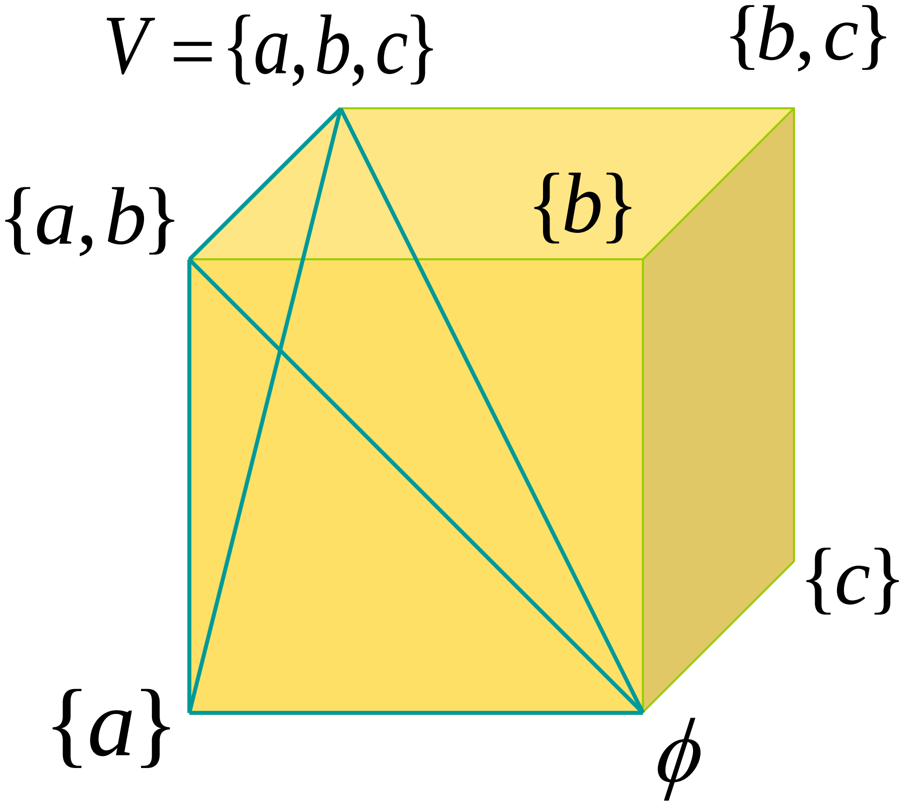


$$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$$



Discrete Convexity

Lovász (1983)



$\hat{f}$: Linear Interpolation

$\hat{f}$: Convex

⇕

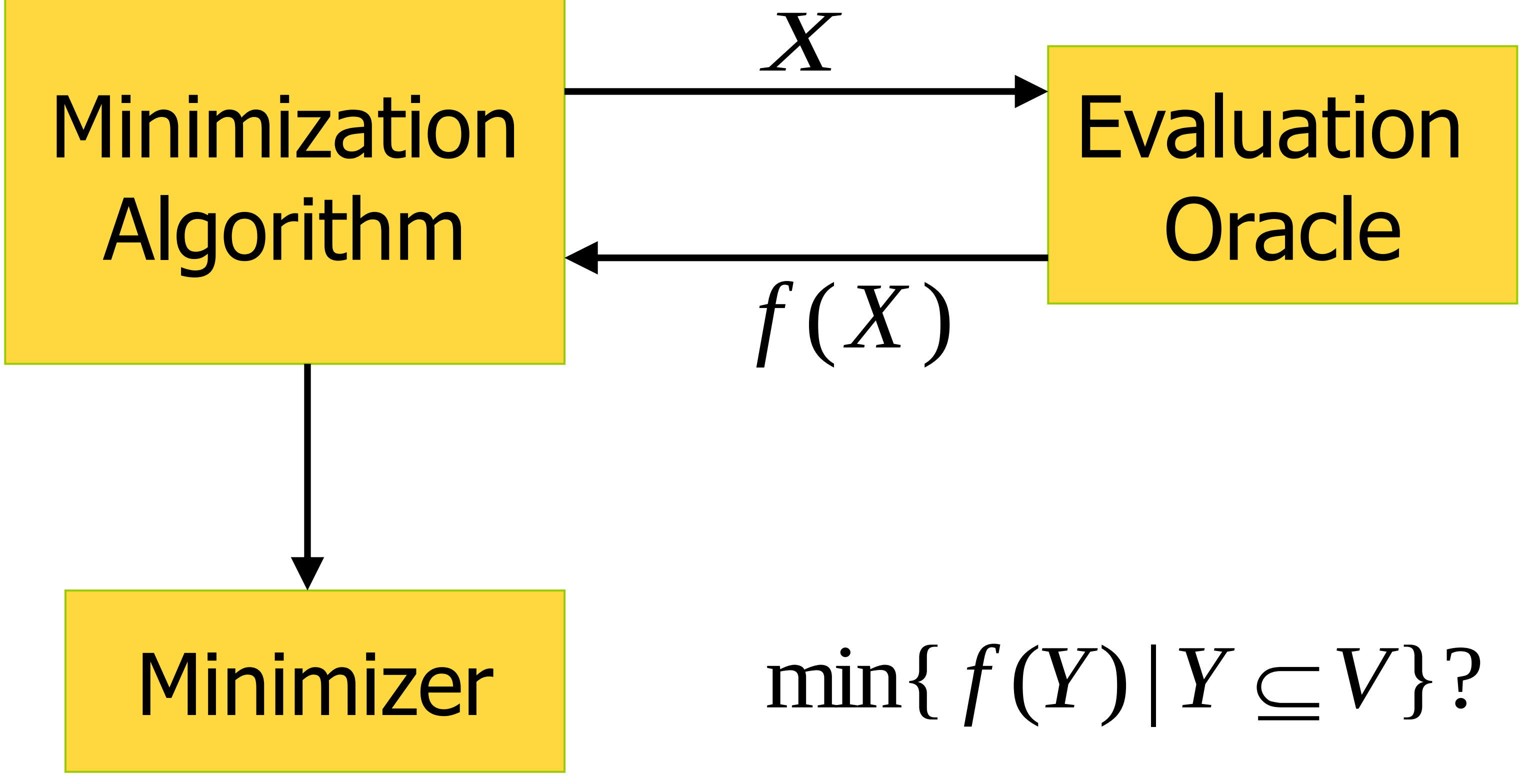
f : Submodular

$\theta \in [0,1]$: Chosen Uniformly at Random

$X := \{v \mid x(v) \geq \theta\}$, $\quad \hat{f}(x) = \mathrm{E}[f(X)]$

Submodular Function Minimization

Assumption: $f(\phi) = 0$



$\min\{ f(Y) \mid Y \subseteq V \}?$

Submodular Function Minimization

Grötschel, Lovász, Schrijver (1981, 1988)

Ellipsoid Method

Cunningham (1985)

$O(n^5 \gamma \log M)$

$O(n^7 \gamma \log n)$

Iwata, Fleischer, Fujishige (2000)

$O(n^7 \gamma + n^8)$

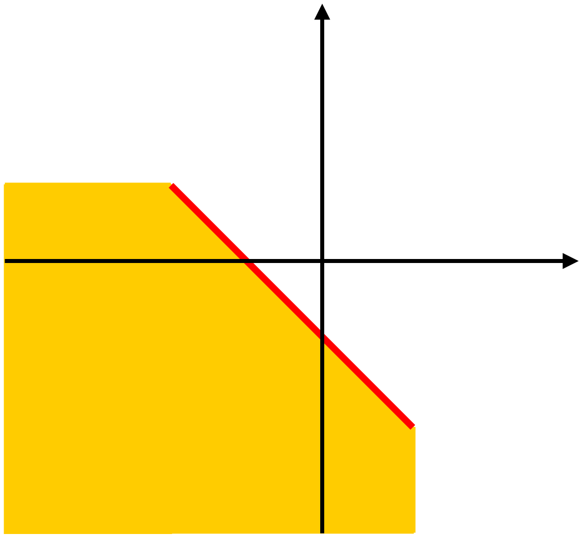
Schrijver (2000)

γ : Time for Function Evaluation

$$M = \max_{X \subseteq V} | f(X) |$$

Base Polyhedra

$$\mathbf{R}^V = \{x \mid V \to \mathbf{R}\}$$



$$x(Y) = \sum_{v \in Y} x(v)$$

Submodular Polyhedron

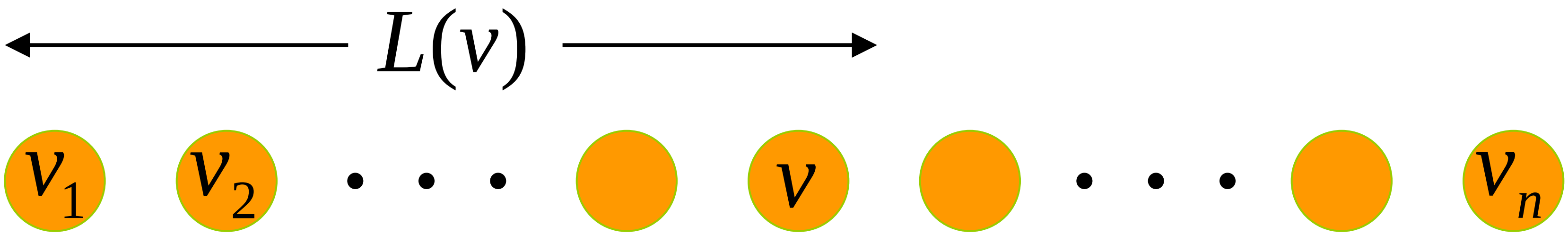
$$P(f) = \{x \mid x \in \mathbf{R}^V, \forall Y \subseteq V, x(Y) \leq f(Y)\}$$

Base Polyhedron

$$B(f) = \{x \mid x \in P(f), x(V) = f(V)\}$$

Greedy Algorithm

Edmonds (1970)
Shapley (1971)



$$y(v) = f(L(v)) - f(L(v) - \{v\}) \qquad (v \in V)$$

y : Extreme Base

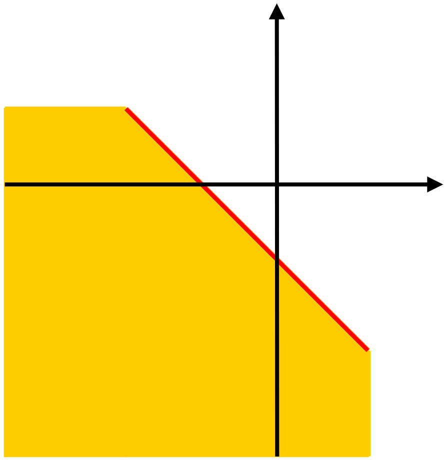
$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y(v_1) \\ y(v_2) \\ \vdots \\ y(v_n) \end{bmatrix} = \begin{bmatrix} f(L(v_1)) \\ f(L(v_2)) \\ \vdots \\ f(L(v_n)) \end{bmatrix}$$

Min-Max Theorem

Theorem Edmonds (1970)

$$\min_{Y \subseteq V} f(Y) = \max\{x^{-}(V) \mid x \in B(f)\}$$

$$x^{-}(v) := \min\{0, x(v)\}$$



$$x^{-}(V) \leq x(Y) \leq f(Y)$$

Combinatorial Approach

Extreme Base $y_L \in B(f)$

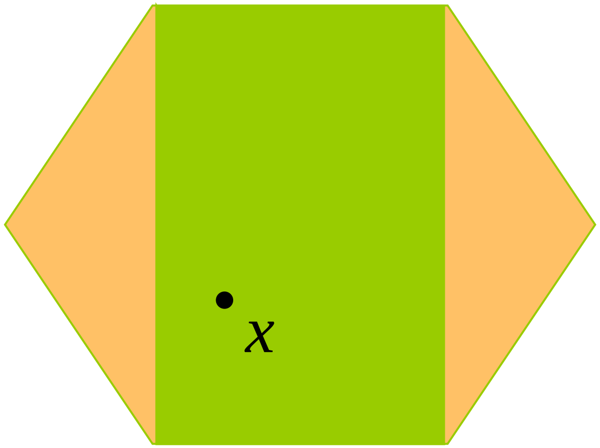
Convex Combination

$$x = \sum_{L \in \Lambda} \lambda_L y_L$$

Cunningham (1985)

$$O(n^6 M\gamma \log nM)$$

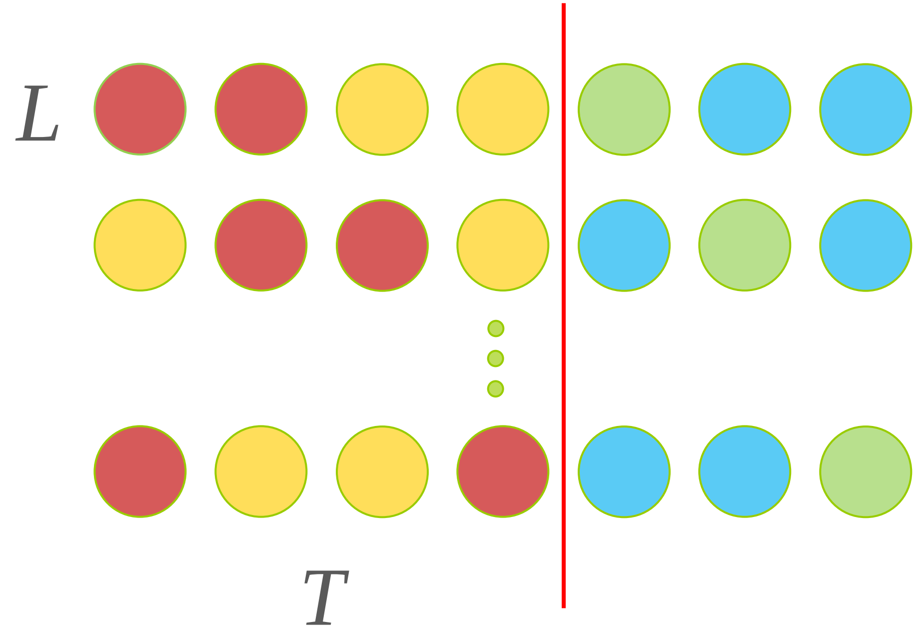
$$M = \max_{X \subseteq V} | f(X) |$$



Combinatorial Approach

$$x = \sum_{L \in \Lambda} \lambda_L y_L$$

y_L : Extreme Base



$$x(v) \le 0, \quad \forall v \in T$$

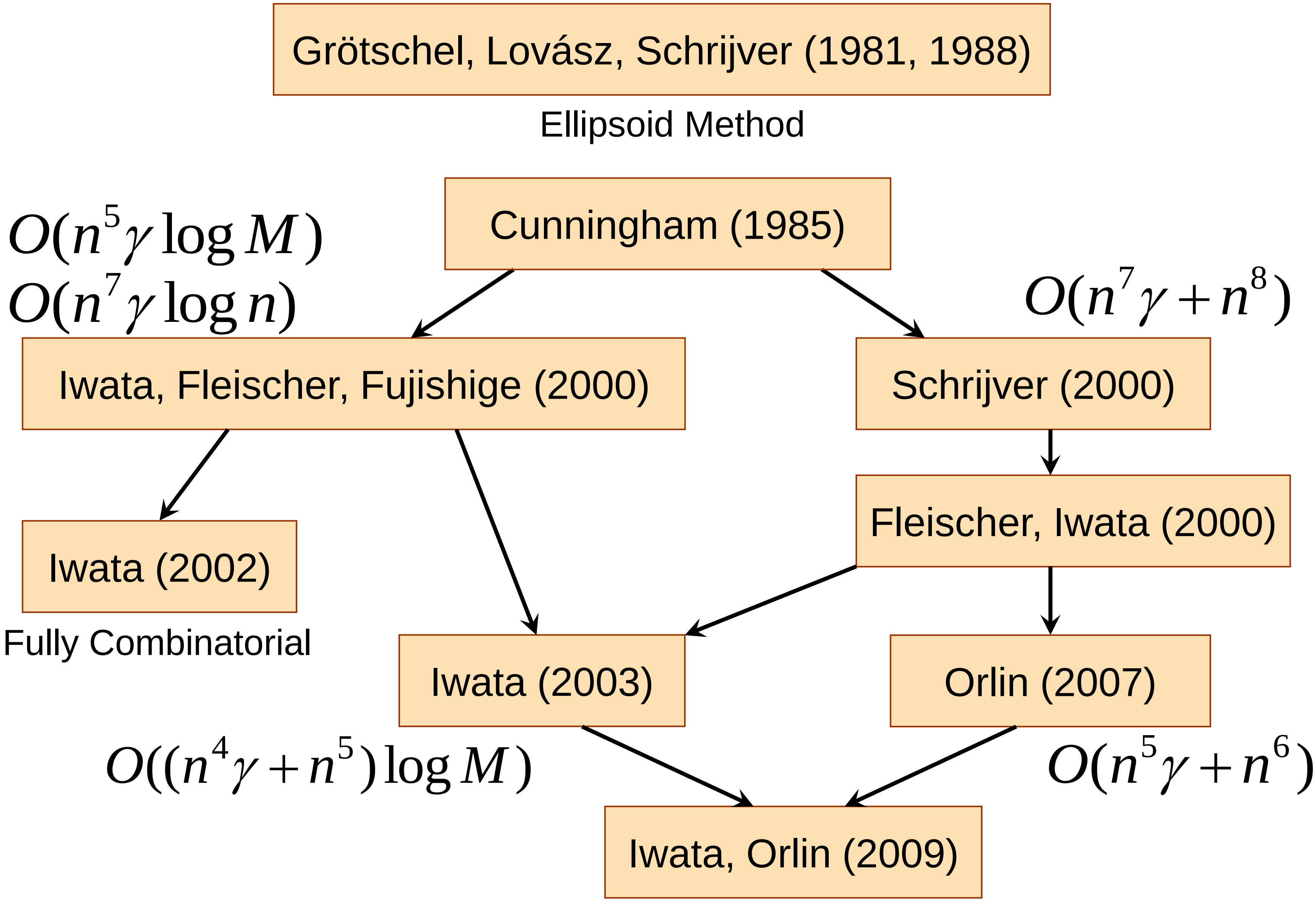
$$x(v) \ge 0, \quad \forall v \notin T$$

$$y_L(T) = f(T), \quad \forall L \in \Lambda. \quad \therefore x(T) = f(T).$$

$$x^-(V) = x(T) = f(T)$$

T : Minimizer

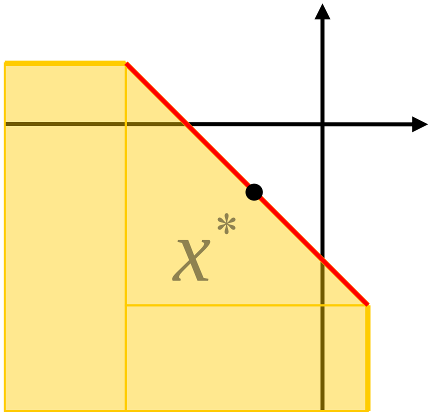
Submodular Function Minimization



The Fujishige-Wolfe Algorithm

$$\text{Minimize } \|x\|^2$$
$$\text{subject to } x \in B(f)$$

Fujishige (1984)



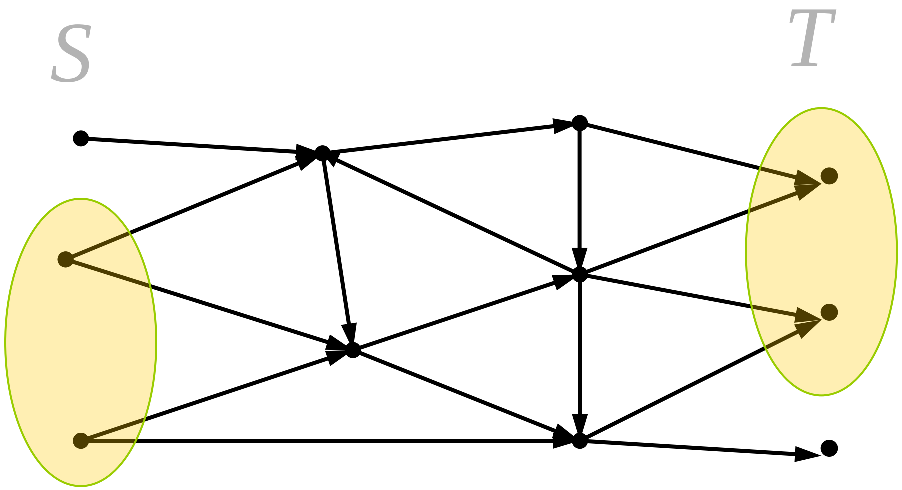
x^* : opt.sol.

$S := \{v \mid x^*(v) < 0\}$ ⟶ Minimal Minimizer

$T := \{v \mid x^*(v) \leq 0\}$ ⟶ Maximal Minimizer

$$f(S) = f(T) = \min_{X \subseteq V} f(X)$$

Evacuation Problem（Dynamic Flow）



Hoppe, Tardos (2000)

$c(a)$: Capacity

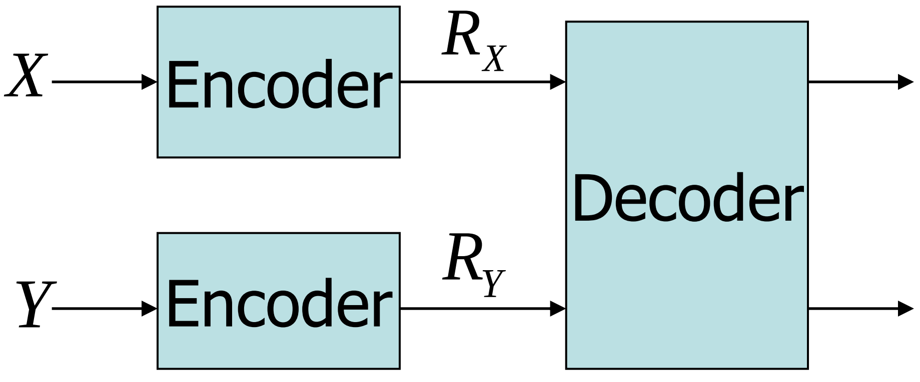
$\tau(a)$: Transit Time

$b(v)$: Supply/Demand

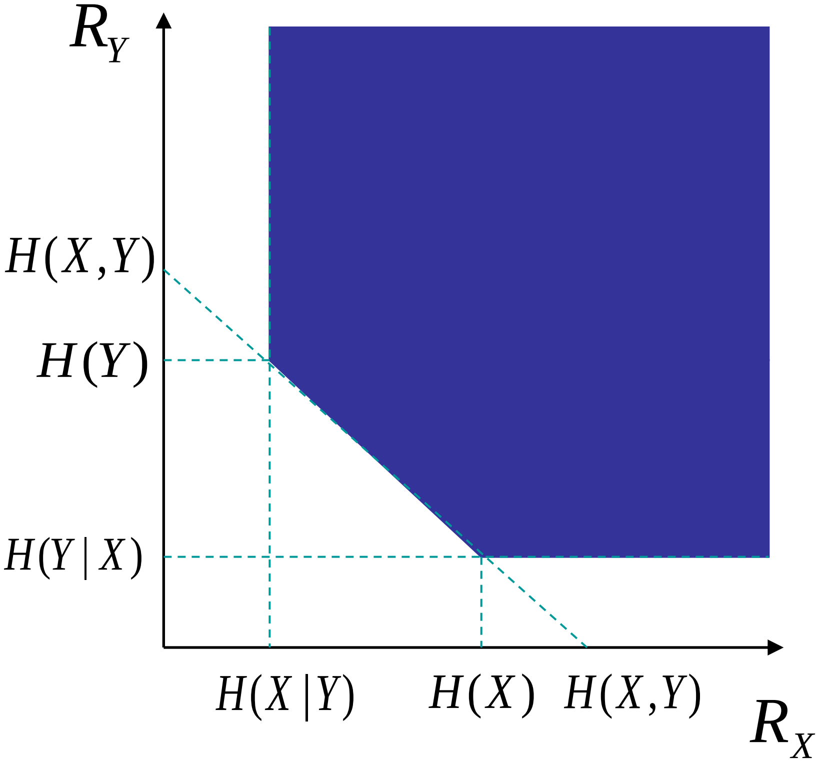
$o(X)$: Maximum Amount of Flow from $X \cap S$ to $T \setminus X$.

Feasible $\Longleftrightarrow$ $b(X) \leq o(X), \forall X \subseteq S \cup T$

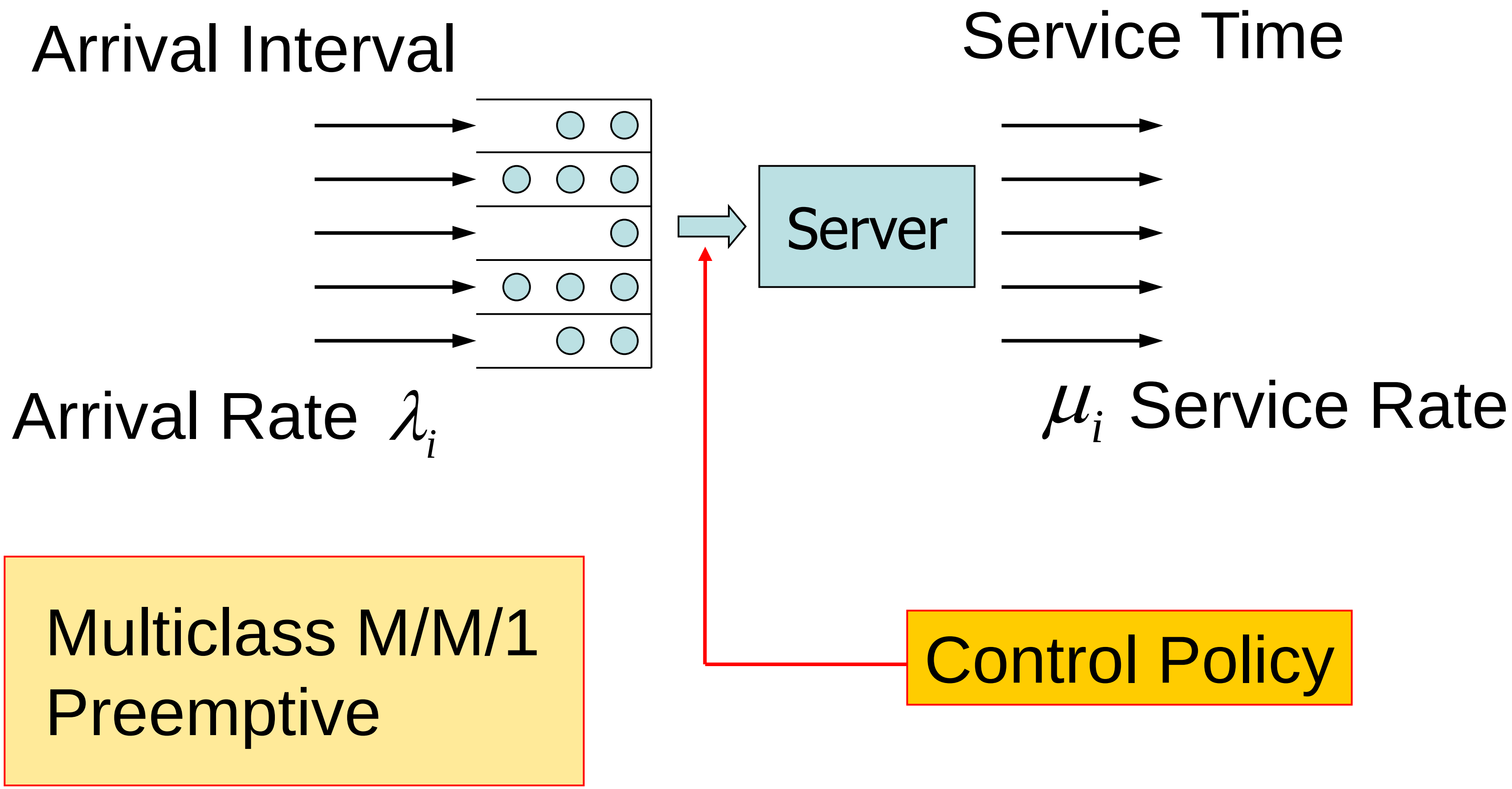
Multiterminal Source Coding



Slepian, Wolf (1973)



Multiclass Queueing Systems



Performance Region

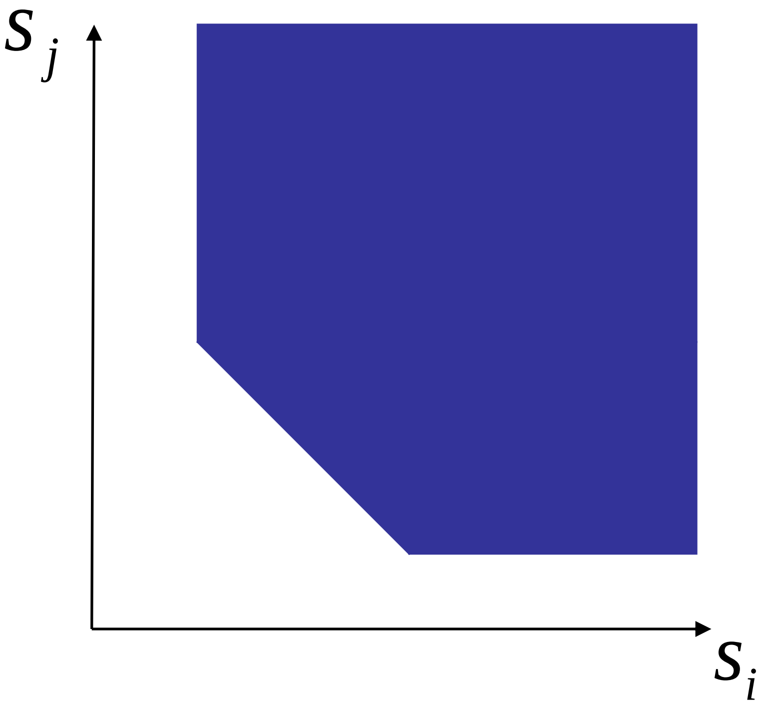
S_j : Expected Staying Time of a Job in *j*

S : Achievable

$$\sum_{i \in X} \rho_i S_i \ge \frac{\sum_{i \in X} \rho_i / \mu_i}{1 - \sum_{i \in X} \rho_i}, \forall X \subseteq V$$

Coffman, Mitrani (1980)

$$\rho_i := \lambda_i / \mu_i, \quad \sum_{i \in V} \rho_i < 1$$



A Class of Submodular Functions

Itoko & Iwata (2007)

$x, y, z \in \mathrm{R}_+^V$

$h:$ Nonnegative, Nondecreasing, Convex

$$f(X) = z(X) - y(X)h(x(X)) \quad (X \subseteq V)$$

→ Submodular

$$\sum_{i \in X} \rho_i S_i \geq \frac{\sum_{i \in X} \rho_i / \mu_i}{1 - \sum_{i \in X} \rho_i}, \forall X \subseteq V$$

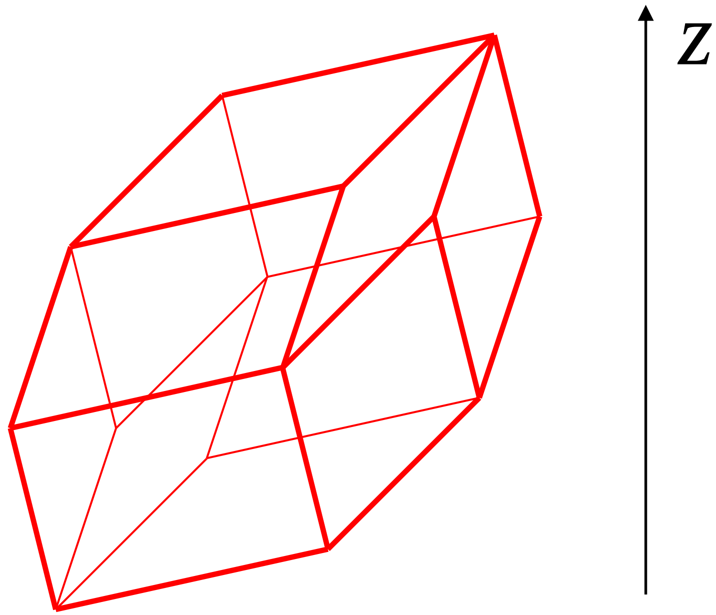
$$z_i := \rho_i S_i \qquad y_i := \frac{\rho_i}{\mu_i}$$

$$x_i := \rho_i \qquad h(x) := \frac{1}{1-x}$$

Zonotope in 3D

$$w(X) = (x(X), y(X), z(X))$$

$$Z = \mathrm{conv}\{w(X) \mid X \subseteq V\}$$

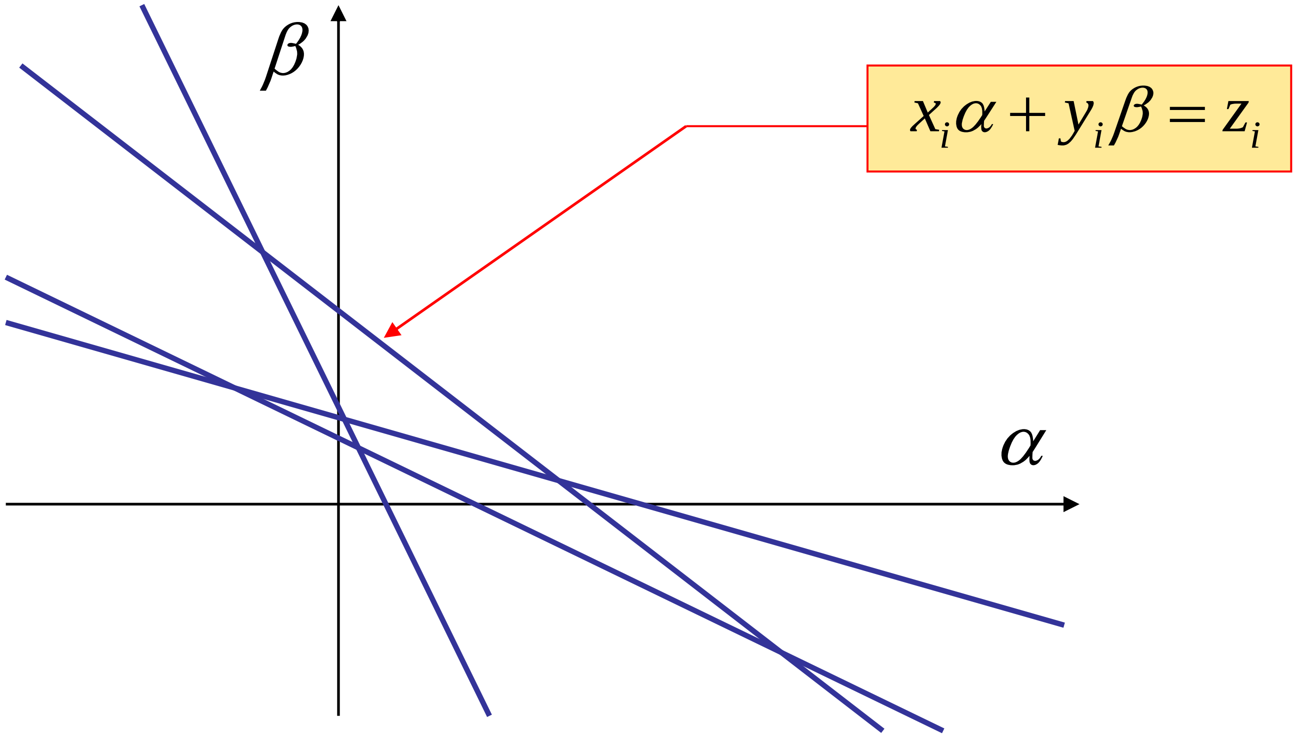


$$\tilde{f}(x, y, z) = z - yh(x)$$

$$\min\{f(X) \mid X \subseteq V\}$$
$$= \min\{\tilde{f}(x, y, z) \mid (x, y, z) : \text{Lower Extreme Point of } Z\}$$

Remark: $\tilde{f}(x, y, z)$ is NOT concave!

Line Arrangement



Enumerating All the Cells

Topological Sweeping Method
Edelsbrunner, Guibas (1989)

$O(n^2)$

Symmetric Submodular Functions

$$f : 2^V \to \mathbf{R}$$

Symmetric $f(X) = f(V \setminus X), \quad \forall X \subseteq V.$

Symmetric Submodular Function Minimization

$$\min\{ f(X) \mid \phi \neq X \neq V \}?$$

$O(n^3 \gamma)$ Queyranne (1998)

Minimum Cut Algorithm by MA-ordering

Nagamochi & Ibaraki (1992)

Minimum Degree Ordering Nagamochi (2007)

Submodular Partition

Minimize $\sum_{i=1}^{k} f(V_i)$

subject to $V = V_1 \cup \cdots \cup V_k$

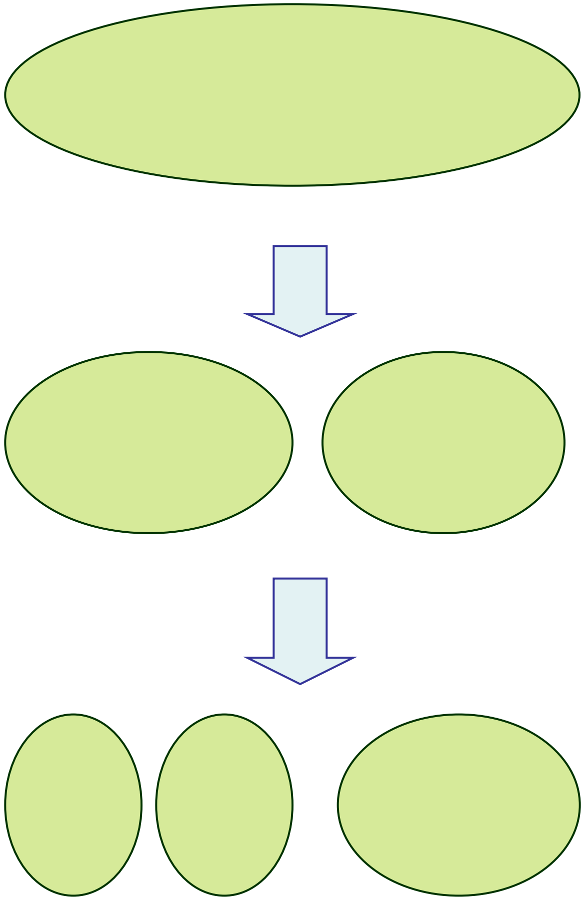
$V_i \cap V_j = \phi \ (i \neq j)$

f : Monotone or Symmetric

$2(1-1/k)$-Approximation

Zhao, Nagamochi, Ibaraki (2005)

Greedy Split



Questions

What Kind of Approximation Algorithms
Can Be Extended to Optimization Problems
with Submodular Cost or Constraints ?

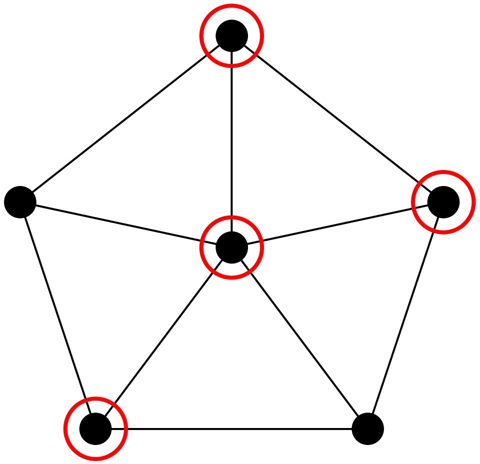
Cf. Submodular Flow (Edmonds & Giles, 1977)

How Can We Exploit Discrete Convexity
in Design of Approximation Algorithms?

Cf. Ellipsoid Method
(Grötschel, Lovász & Schrijver, 1981)

Submodular Vertex Cover

Graph $G = (V, E)$



Submodular Function

$$f : 2^V \to \mathbf{R}_+$$

Find a Vertex Cover $S \subseteq V$ Minimizing $f(S)$

2-Approximation Algorithm

Goel, Karande, Tripathi, Wang (FOCS 2009)

Iwata & Nagano (FOCS 2009)

Relaxation Problem

Convex Programming Relaxation (CPR)

Minimize $\hat{f}(x)$

subject to $x(u) + x(v) \geq 1 \quad (\forall e = (u, v) \in E)$

$$x(v) \geq 0 \quad (\forall v \in V)$$

CPR has a half-integral optimal solution.

→ 2-Approximation Algorithm

Submodular Cost Set Cover

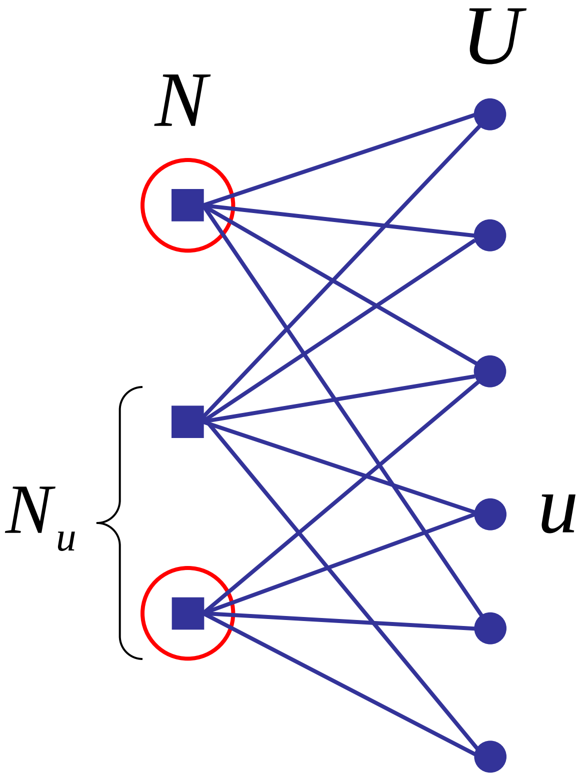
Find $X \subseteq N$ Covering U with Minimum $f(X)$.

$$\eta := \max_{u \in U} |N_u|$$

η-Approximation

Rounding Algorithm

Primal-Dual Algorithm



Submodular Multiway Partition

Chekuri & Ene (FOCS 2011)

Minimize $\sum_{i=1}^{k} f(V_i)$

subject to $V = V_1 \cup \cdots \cup V_k$

$V_i \cap V_j = \phi \ (i \neq j)$

$s_i \in V_i \quad (i = 1, \ldots, k)$

f : Nonnegative Submodular

Relaxation Problem

Convex Programming Relaxation

Minimize $\sum_{i=1}^{k} \hat{f}(x_i)$

subject to $\sum_{i=1}^{k} x_i(v) = 1 \quad (\forall v \in V)$

$x_i(s_i) = 1 \quad (i = 1, \dots, k)$

$x_i(v) \geq 0 \quad (i = 1, \dots, k, \forall v \in V)$

Ellipsoid Method → x_i^* : Optimal Solution

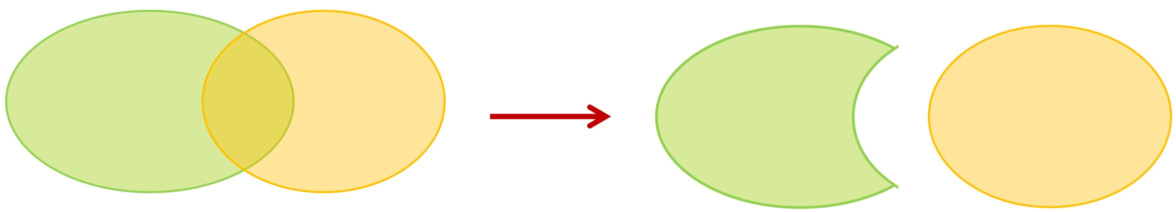
Rounding Scheme

f : Symmetric Submodular

$\theta \in [0,1]$: Chosen Uniformly at Random

$$A_i := \{v \mid x_i(v) \geq \theta\} \quad (i = 1, \ldots, k), \qquad U := V - \bigcup_{i=1}^{k} A_i$$

Uncross $(A_1, \ldots, A_k)$



Return $(A_1, \ldots, A_{k-1}, A_k \cup U)$

→ 3/2-Approximate Solution

Rounding Scheme

f : Nonnegative Submodular

$\delta \in (\frac{1}{2}, 1]$: Chosen Uniformly at Random

$A_i := \{v \mid x_i(v) \geq \theta\} \quad (i = 1, \ldots, k), \qquad U := V - \bigcup_{i=1}^{k} A_i$

Return $(A_1, \ldots, A_{k-1}, A_k \cup U)$

$\longrightarrow$ 2-Approximate Solution

Improvement over the $(k-1)$-Approximation by Zhao, Nagamochi, & Ibaraki (2005)

Submodular Function Maximization

Monotone Submodular Functions

Nemhauser, Wolsey, Fisher (1978)

Cardinality Constraint

(1−1/e)-Approximation

Calinescu, Chekuri, Pál, Vondrák (IPCO 2007)

Vondrák (STOC 2008)

Filmes and Ward (FOCS 2012)

Matroid Constraint

(1−1/e)-Approximation

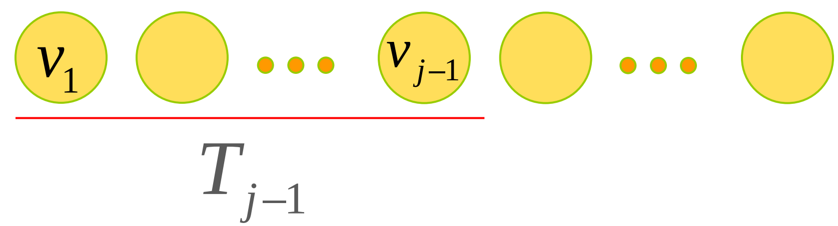
Greedy Approximation Algorithm

Nemhauser, Wolsey, Fisher (1978)

f : Monotone Submodular Function

Maximize $f(S)$ subject to $|S| \le k$

Greedy Algorithm



$$T_0 := \phi$$

$$\left.\begin{aligned} v_j &:= \arg\max f(T_{j-1} \cup \{v\}) \\ T_j &:= T_{j-1} \cup \{v_j\} \end{aligned}\right\} \quad j = 1, \ldots, k$$

Greedy Approximation Algorithm

S^* : Optimal Solution $\eta := f(S^*)$

$$\eta \le k\rho_j + \sum_{i=1}^{j-1} \rho_i \quad (j = 1, \ldots, k)$$

$$\rho_j := f(T_j) - f(T_{j-1})$$

$$f(T_k) = \sum_{i=1}^{k} \rho_i$$

$$\because)\ f(S^*) \le f(S^* \cup T_{j-1})$$

$$\le f(T_{j-1}) + \sum_{u \in S^* \setminus T_{j-1}} \left[f(T_{j-1} \cup \{u\}) - f(T_{j-1}) \right]$$

$$\le \underline{f(T_{j-1})} + k\rho_j$$

$$\underline{f(T_{j-1})} \rightarrow \sum_{i=1}^{j-1} \rho_i$$

Greedy Approximation Algorithm

Minimize $\sum_{i=1}^{k} \rho_i$

subject to
$$\begin{bmatrix} k & 0 & \cdots & 0 \\ 1 & k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1 & \cdots & 1 & k \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} \geq \begin{bmatrix} \eta \\ \eta \\ \vdots \\ \eta \end{bmatrix}$$

$$\rho_j \geq 0 \qquad (j = 1, \ldots, k)$$

$$\sum_{i=1}^{k} \rho_i = f(T_k)$$

$$\eta := f(S^*)$$

$$\hat{\rho}_j := \frac{\eta}{k}\left(1 - \frac{1}{k}\right)^j \qquad (j = 1, \ldots, k)$$

$$\sum_{i=1}^{k} \hat{\rho}_i = \eta\left[1 - \left(1 - \frac{1}{k}\right)^k\right] \geq \eta\left(1 - \frac{1}{\mathrm{e}}\right)$$

Submodular Welfare Problem

Utility Functions $f_1, \ldots, f_k$ (Monotone, Submodular)

Maximize $\sum_{i=1}^{k} f_i(V_i)$

subject to $V = V_1 \cup \cdots \cup V_k$

$V_i \cap V_j = \phi \ (i \neq j)$

$(1 - 1/e)$ -Approximation Vondrák (2008)

Submodular Function Maximization

Nonnegative Submodular Functions

Feige, Mirrokni, Vondrák (FOCS 2007)

2/5-Approximation

Lee, Mirrokni, Nagarajan, Sviridenko (STOC 2009)

1/4-Approximation (Matroid Constraint)

1/5-Approximation (Knapsack Constraints)

Buchbinder, Feldman, Naor, Schwartz (FOCS 2012)

1/2-Approximation

Double-Greedy Algorithm

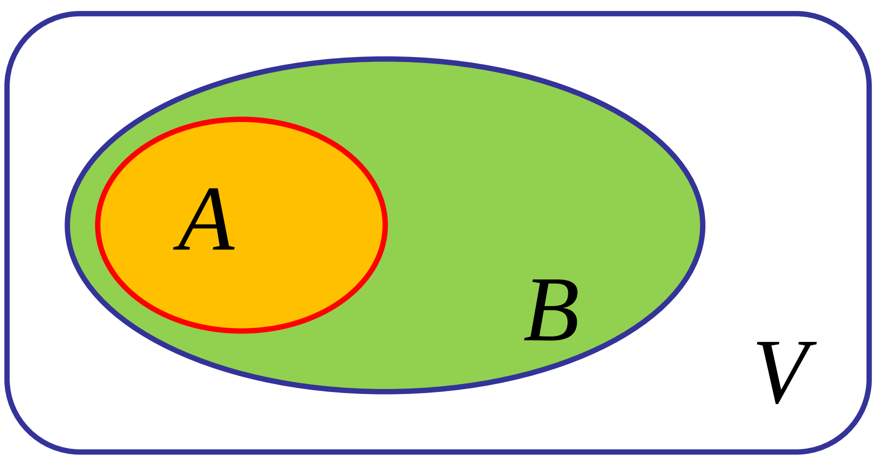
Buchbinder, Feldman, Naor, Schwartz (FOCS 2012)

Initialize: $A := \phi, B := V, D := \phi.$

Iteration: grow A or shrink B.

Invariants: $A \subseteq B, D \subseteq B \setminus A.$

Output: either A or B that has the greater value.



Double-Greedy Algorithm

While $A \cup D \neq B$

Select $u \in B \setminus (D \cup A)$.

$\alpha := \max\{ f(A \cup \{u\}) - f(A), 0\}$.

$\beta := \max\{ f(B \setminus \{u\}) - f(B), 0\}$.

If $\alpha + \beta > 0$, then

$A := A \cup \{u\}$ with probability $\frac{\alpha}{\alpha+\beta}$,

$B := B \setminus \{u\}$ with probability $\frac{\beta}{\alpha+\beta}$.

Else $D := D \cup \{u\}$.

Double-Greedy Algorithm

The expectation $\mathrm{E}[f(A)+f(B)+2f(Z)]$

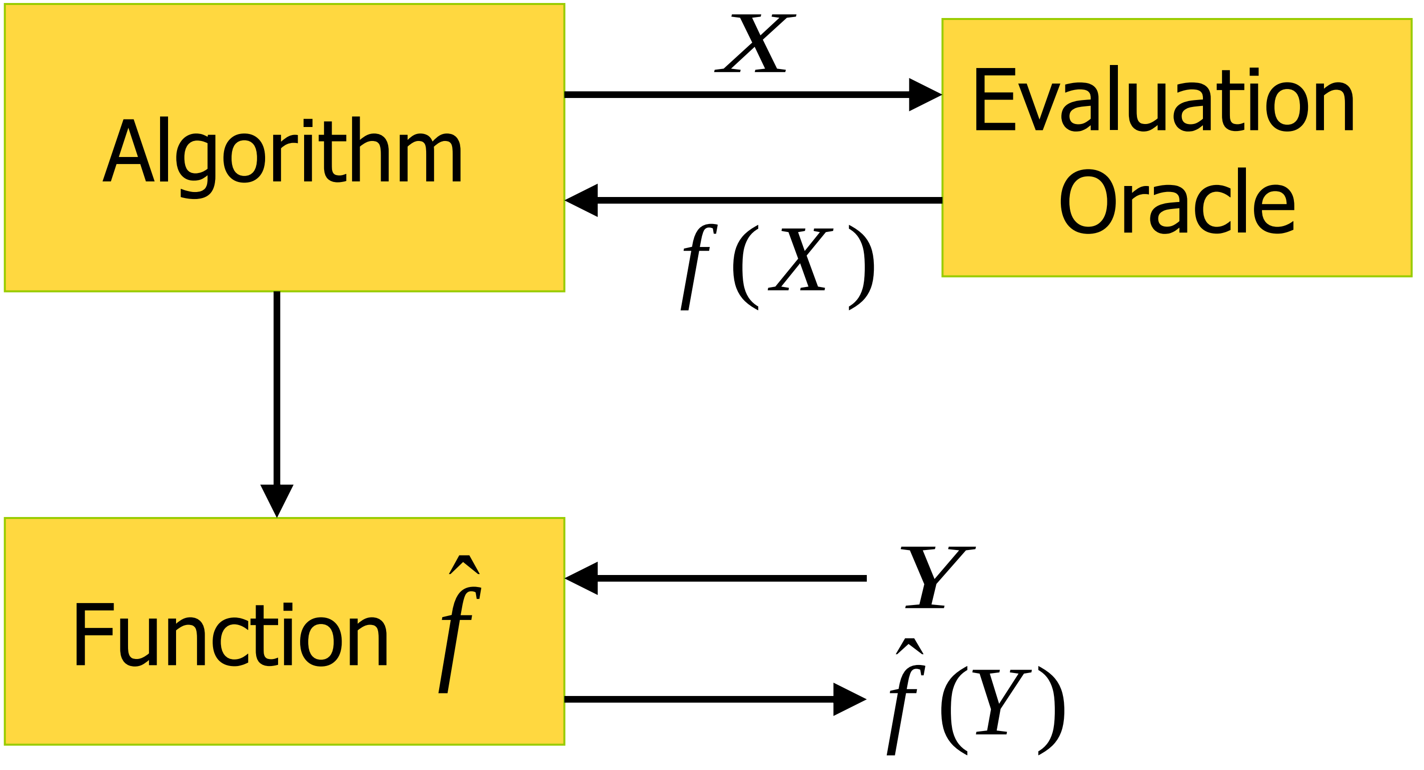
never decreases in the process.

∵) Expected changes of $f(A)+f(B)+2f(Z)$

at each iteration is at least

$$\frac{\alpha}{\alpha+\beta}\cdot\alpha+\frac{\beta}{\alpha+\beta}\cdot\beta-\frac{2\alpha\beta}{\alpha+\beta}=\frac{(\alpha-\beta)^2}{\alpha+\beta}\geq 0.$$

Approximating Submodular Functions



Approximating Submodular Functions

Goemans, Harvey, Iwata & Mirrokni (SODA 2009)

$$f(\phi) = 0, \quad f(X) \geq 0, \forall X \subseteq V.$$

Construct a set function $\hat{f}$ such that

$$\hat{f}(X) \leq f(X) \leq \alpha(n)\hat{f}(X), \quad \forall X \subseteq V.$$

For what function α is this possible?

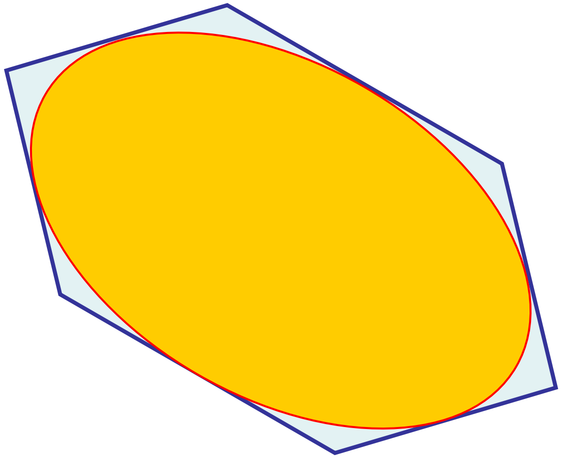
Algorithm with $\alpha(n) = O(\sqrt{n}\log n)$
for monotone submodular functions

Ellipsoidal Approximation

K : Centrally Symmetric Convex Body

$(x \in K \Rightarrow -x \in K)$

$E(A)$: Maximum Volume Inscribed Ellipsoid
(The John Ellipsoid)



$$E(A) \subseteq K \subseteq \sqrt{n}E(A)$$

Submodular Load Balancing

Svitkina & Fleischer (FOCS 2008)

$f_1, \ldots, f_m$: Monotone Submodular Functions

$$\min_{\{V_1, \ldots, V_m\}} \max_j f_j(V_j) ?$$

$$f_j(X) := \sum_{v \in X} p_j(v)$$ → Scheduling

2-Approximation Algorithm

Lenstra, Shmoys, Tardos (1990)

$O(\sqrt{n} \log n)$-Approximation Algorithm

Learning Submodular Functions

Balcan & Harvey (STOC 2010)

PMAC-learning

Construct a set function $\hat{f}$ such that

$$\hat{f}(X) \le f(X) \le \alpha(n)\hat{f}(X)$$

holds for most ($1-\varepsilon$ fraction) of X.
For what function α is this possible with probability $1-\delta$ in $\mathrm{poly}(n, 1/\varepsilon, 1/\delta)$ time?

Summary

- Submodular Functions Arise Everywhere.
- Discrete Analogue of Convexity.
- Combinatorial Algorithms for Minimization.
- Exploit Special Structures of Problems.
- Approx. Algorithms by Discrete Convexity
- Approx. Algorithms for Maximization
- Learning Submodular Functions