ISMP 2012. Berlin

Maximizing Expected Utility for Stochastic Combinatorial Optimization Problems

Jian Li

Institute of Interdisiplinary Information Sciences

Tsinghua University

Aug. 2012

Joint work with Amol Deshpande (UMD)

lijian83@mail.tsinghua.edu.cn

Inadequacy of Expected Value

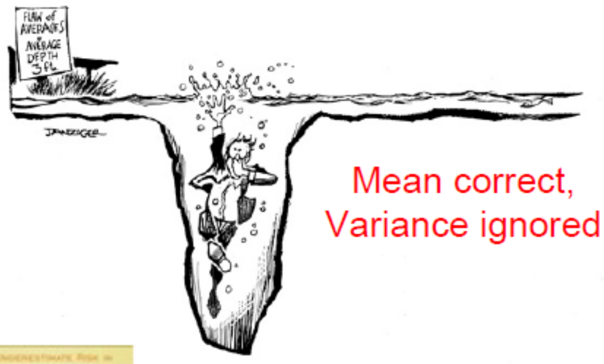
- Stochastic Optimization
  - Some part of the input are probabilistic
  - Most common objective: Optimizing the expected value

- Inadequacy of expected value:
  - Unable to capture risk-averse or risk-prone behaviors
    - Action 1: \$100 VS Action 2: \$200 w.p. 0.5; \$0 w.p. 0.5
    - Risk-averse players prefer Action 1
    - Risk-prone players prefer Action 2 (e.g., a gambler spends \$100 to play Double-or-Nothing)

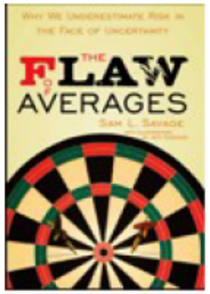
Inadequacy of Expected Value

- Be aware of risk!

Flaw of averages (weak form):



Mean correct,
Variance ignored



Flaw of averages (strong form):

Wrong value of mean:
$f(E[X]) \neq E[f(X)]$

21

- St. Petersburg Paradox

Expected Utility Maximization Principle

Remedy: Use a utility function

$\mu : R \rightarrow R$: The utility function: value (profit/cost)-> utility

Expected Utility Maximization Principle: the decision maker should choose the action that maximizes the **expected utility**

$$\text{maximize. } \mathbb{E}[\mu(\text{profit})]$$

- Proved quite useful to explain some popular choices that seem to contradict the expected value criterion
- An *axiomatization* of the principle (known as von Neumann-Morgenstern expected utility theorem).

Problem Definition

- Deterministic version:
  - A set of element $\{e_i\}$, each associated with a weight w_i
  - A solution *S* is a subset of elements (that satisfies some property)
  - **Goal:** Find a solution *S* such that the total weight of the solution $w(S) = \Sigma_{i \in S} w_i$ is minimized
  - E.g. shortest path, minimal spanning tree, top-k query, matroid base

Problem Definition

- Deterministic version:
  - A set of element $\{e_i\}$, each associated with a weight w_i
  - A solution S is a subset of elements (that satisfies some property)
  - **Goal:** Find a solution S such that the total weight of the solution $w(S) = \Sigma_{i \in S} w_i$ is minimized
  - E.g. shortest path, minimal spanning tree, top-k query, matroid base
- Stochastic version:
  - w_is are independent positive random variable
  - $\mu()$: $R^+ \rightarrow R^+$ is the utility function (assume $\lim_{x \rightarrow \infty} \mu(x) = 0$)
  - **Goal:** Find a solution S such that the expected utility $E[\mu(w(S))]$ is maximized

Our Results

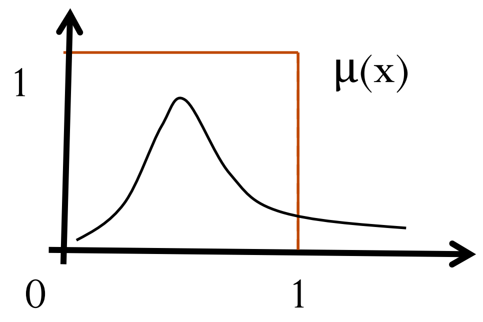
- **THM:** If the following two conditions hold
  - (1) there is a pseudo-polynomial time algorithm for the **exact version** of deterministic problem, and
  - (2) μ is bounded by a constant and satisfies *Hőlder condition* $|\mu(x) - \mu(y)| \le C|x-y|^{\alpha}$ for constant C and $\alpha \ge 0.5$,

  then we can obtain in polynomial time a solution S such that $E[\mu(w(S))] \ge OPT - \varepsilon$, for any fixed $\varepsilon > 0$

- Exact version: find a solution of weight exactly K
- Pseudo-polynomial time: polynomial in K
- Problems satisfy condition (1): shortest path, minimum spanning tree, matching, knapsack.

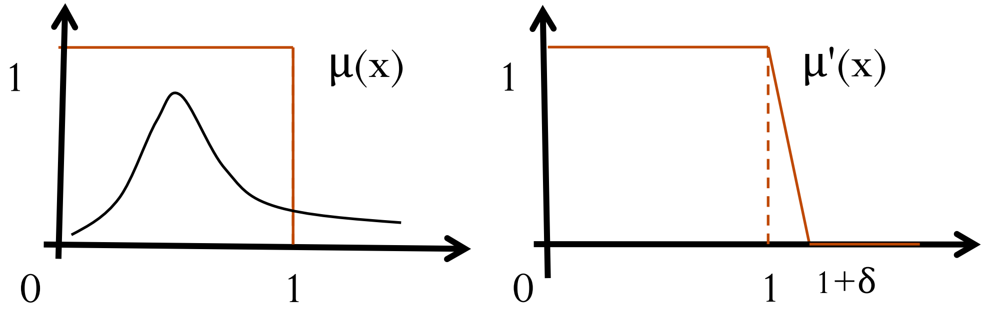
Our Results

- If *μ is a threshold function, maximizing* $E[\mu(w(S))]$ *is equivalent to maximizing* $Pr[w(S)<1]$
  - *minimizing overflow prob.* [Kleinberg, Rabani, Tardos. STOC'97] [Goel, Indyk. FOCS'99]
  - *chance-constrained stochastic optimization problem* [Swamy. SODA'11]



Our Results

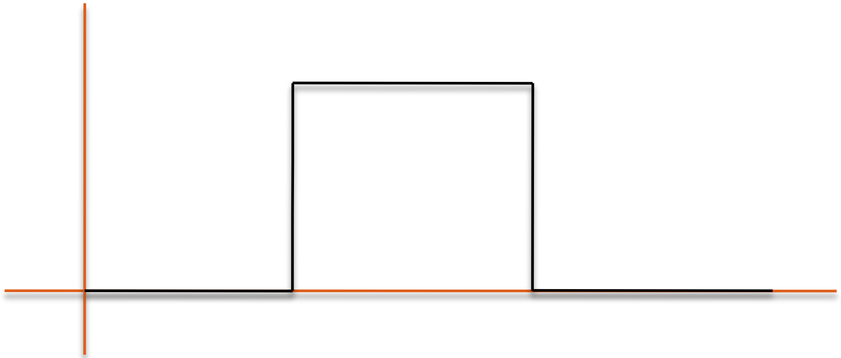
- If *μ is a threshold function, maximizing E[μ(w(S))] is equivalent to maximizing Pr[w(S)<1]*
  - *minimizing overflow prob.* [Kleinberg, Rabani, Tardos. STOC'97] [Goel, Indyk. FOCS'99]
  - *chance-constrained stochastic optimization problem* [Swamy. SODA'11]
- However, our technique can not handle discontinuous function directly
- So, we consider a continuous version *μ'*

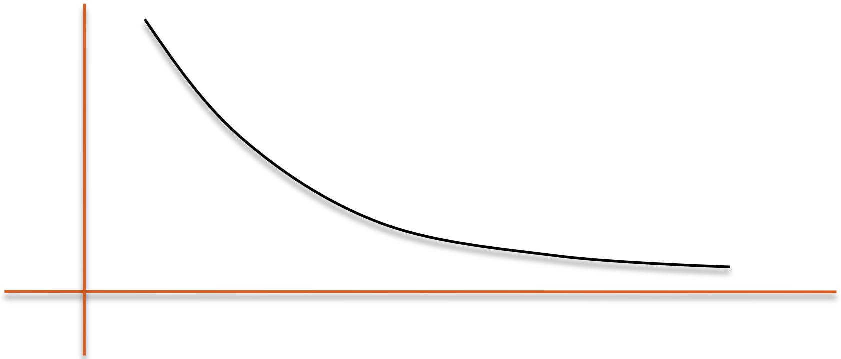


$$Pr[w(S) \le 1] \le E[\mu'(w(S))] \le Pr[w(S) \le 1+\delta]$$

Our Results

- Other Utility Functions

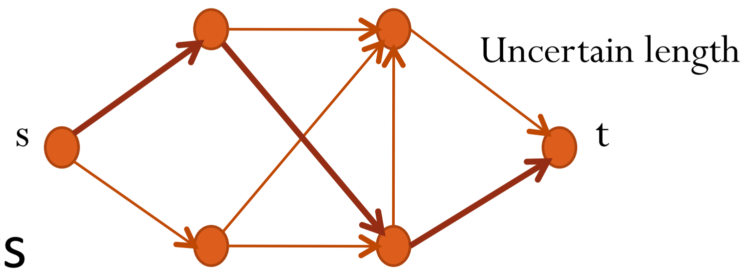




Exponential

Our Results

- **Stochastic shortest path** : find an s-t path P such that $Pr[w(P)<1]$ is maximized



- Previous results
  - Many heuristics
  - Poly-time approximation scheme (PTAS) if (1) all edge weights are normally distributed r.v.s (2) $OPT>0.5$ [Nikolova, Kelner, Brand, Mitzenmacher. ESA'06] [Nikolova. APPROX'10]
  - Bicriterion PTAS ($Pr[w(P)<1+\delta]>(1\text{-eps})OPT$) for exponential distributions [Nikolova, Kelner, Brand, Mitzenmacher. ESA'06]
- Our result
  - Bicriterion PTAS if $OPT= Const$

Our Results

- **Stochastic knapsack**: find a collection S of items such that $Pr[w(S)<1]>\gamma$ and the total profit is maximized

Each item has a deterministic profit and a (uncertain) size

Knapsack, capacity=1

- Previous results
  - $log(1/(1-\gamma))$-approximation [Kleinberg, Rabani, Tardos. STOC'97]
  - Bicriterion PTAS for exponential distributions [Goel, Indyk. FOCS'99]
  - PTAS for Bernouli distributions if $\gamma= Const$ [Goel, Indyk. FOCS'99] [Chekuri, Khanna. SODA'00]
  - Bicriterion PTAS if $\gamma= Const$ [Bhalgat, Goel, Khanna. SODA'11]
- Our result
  - Bicriterion PTAS if $\gamma= Const$ (with a better running time than Bhalgat et al.)
  - Stochastic partial-ordered knapsack problem with tree constraints

Our Algorithm

- Observation: We first note that the exponential utility functions are tractable

$$\mathbb{E}[\phi^{w(S)}] = \mathbb{E}[\phi^{\sum_{e \in S} w_e}] = \mathbb{E}[\prod_{e \in S} \phi_{w_e}] = \prod_{e \in S} \mathbb{E}[\phi^{w_e}]$$

Our Algorithm

- Observation: We first note that the exponential utility functions are tractable

$$\mathbb{E}[\phi^{w(S)}] = \mathbb{E}[\phi^{\sum_{e\in S} w_e}] = \mathbb{E}[\prod_{e\in S} \phi_{w_e}] = \prod_{e\in S} \mathbb{E}[\phi^{w_e}]$$

- Approximate the expected utility by a short linear sum of exponential utility $\mu(x) \simeq \sum_{k=1}^{L} c_k \phi_k^x, \quad L = O(1)$
  - Show that the error $|\mu(x) - \sum_{k=1}^{L} c_k \phi_k^x|$ can be bounded by ϵ

Our Algorithm

- Observation: We first note that the exponential utility functions are tractable

$$\mathbb{E}[\phi^{w(S)}] = \mathbb{E}[\phi^{\sum_{e \in S} w_e}] = \mathbb{E}[\prod_{e \in S} \phi_{w_e}] = \prod_{e \in S} \mathbb{E}[\phi^{w_e}]$$

- Approximate the expected utility by a short linear sum of exponential utility $\mu(x) \simeq \sum_{k=1}^{L} c_k \phi_k^x, \quad L = O(1)$
  - Show that the error $|\mu(x) - \sum_{k=1}^{L} c_k \phi_k^x|$ can be bounded by ϵ
- By linearity of expectation

$$\mathbb{E}[\mu(w(S))] \simeq \mathbb{E}[\sum_{k=1}^{L} c_k \phi_k^{w(S)}] = \sum_{k=1}^{L} c_k \mathbb{E}[\phi_k^{w(S)}]$$

Our Algorithm

- Problem: Find a solution *S* minimizing

$$\sum_{k=1}^{L} c_k \mathbb{E}[\phi_k^{w(S)}]$$

Our Algorithm

- Problem: Find a solution *S* minimizing

$$\sum_{k=1}^{L} c_k \mathbb{E}[\phi_k^{w(S)}]$$

- Suffices to find a set *S* such that

Opt

$$\langle \mathbb{E}[\phi_1^{w(S)}], \ldots, \mathbb{E}[\phi_L^{w(S)}] \rangle \simeq \langle \mathbb{E}[\phi_1^{w(S^*)}], \ldots, \mathbb{E}[\phi_L^{w(S^*)}] \rangle$$

Our Algorithm

- Problem: Find a solution *S* minimizing

$$\sum_{k=1}^{L} c_k \mathbb{E}[\phi_k^{w(S)}]$$

- Suffices to find a set *S* such that

Opt

$$\langle \mathbb{E}[\phi_1^{w(S)}], \ldots, \mathbb{E}[\phi_L^{w(S)}] \rangle \simeq \langle \mathbb{E}[\phi_1^{w(S^*)}], \ldots, \mathbb{E}[\phi_L^{w(S^*)}] \rangle$$

- (Approximately) Solve the multi-objective optimization problem with objectives $\mathbb{E}[\phi_i^{w(S)}]$ for *k=1,…,L*

Approximating Utility Functions

How to approximate $\mu()$ by a short sum of exponentials?

(with lim μ(x)->0) $\mu(x) \approx \sum_{k=1}^{L} c_k \phi_k^x$

Approximating Utility Functions

How to approximate $\mu()$ by a short sum of exponentials?

(with lim μ(x)->0) $\mu(x) \approx \sum_{k=1}^{L} c_k \phi_k^x$

- A scheme based on Fourier Series decomposition:

  For a periodic function f(x) with period 2π

$$f(x) \sim \sum_{k=-\infty}^{+\infty} c_k e^{ikx}$$

  where the Fourier coefficient $c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} \mathrm{d}x.$

Approximating Utility Functions

How to approximate $\mu()$ by a short sum of exponentials?

(with lim μ(x)->0) $\mu(x) \approx \sum_{k=1}^{L} c_k \phi_k^x$

- A scheme based on Fourier Series decomposition:

  For a periodic function f(x) with period 2π

$$f(x) \sim \sum_{k=-\infty}^{+\infty} c_k e^{ikx}$$

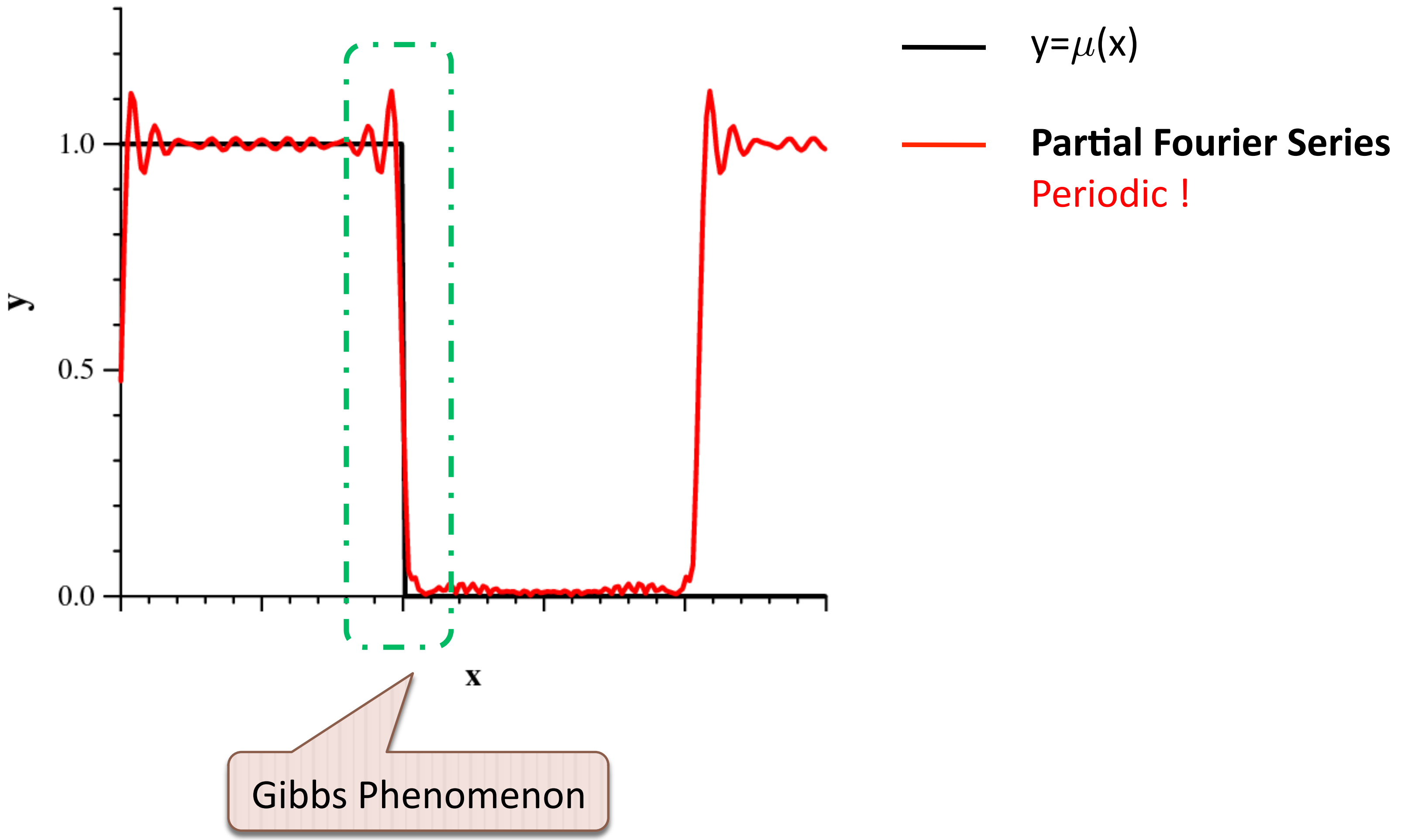
  where the Fourier coefficient $c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} \mathrm{d}x.$

- Consider the partial sum (*L=2N+1*)

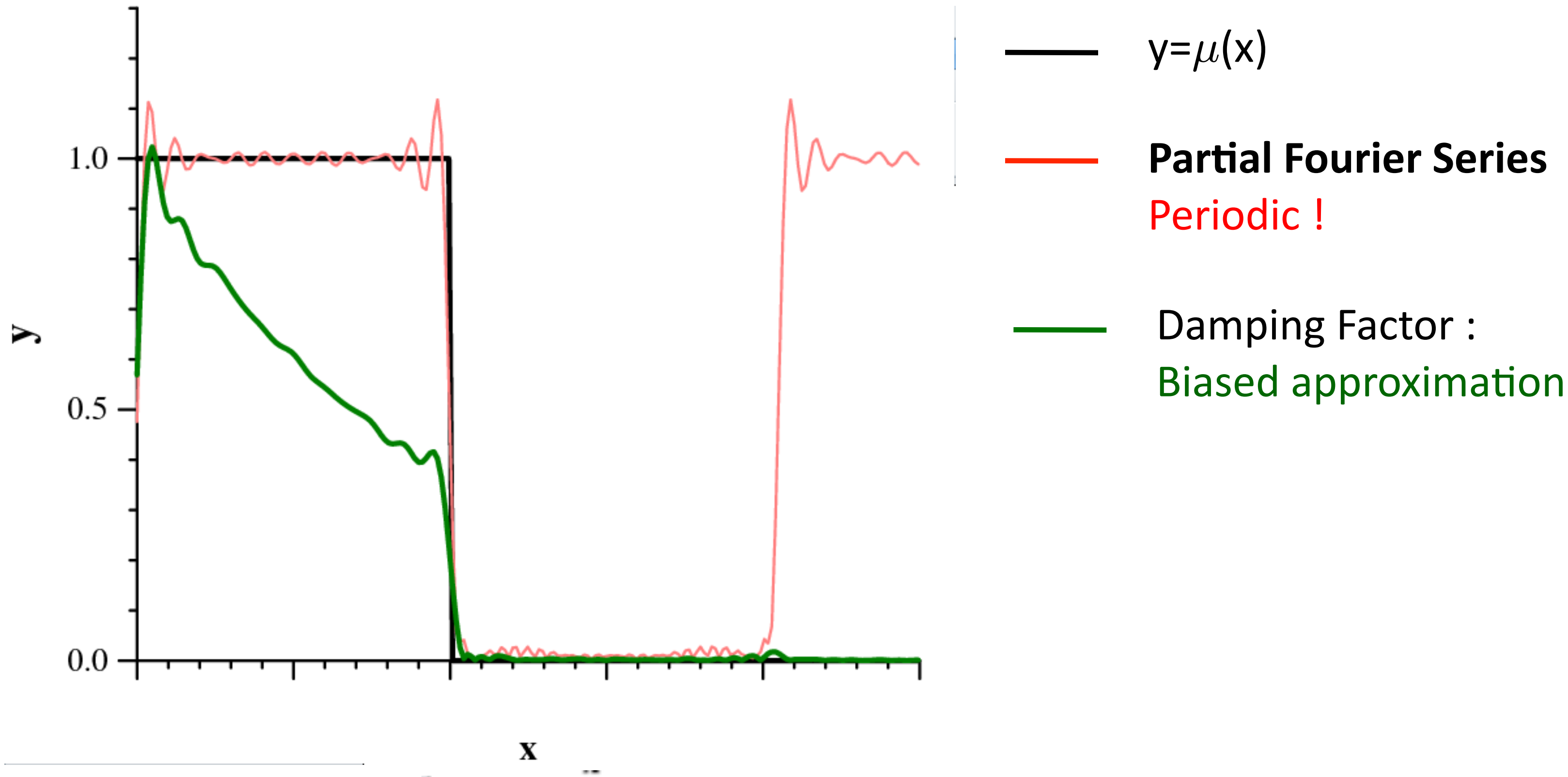
$$(S_N f)(x) = \sum_{k=-N}^{N} c_k e^{ikx}$$

Approximating The Threshold Function

Need *L=poly(1/ϵ)* terms

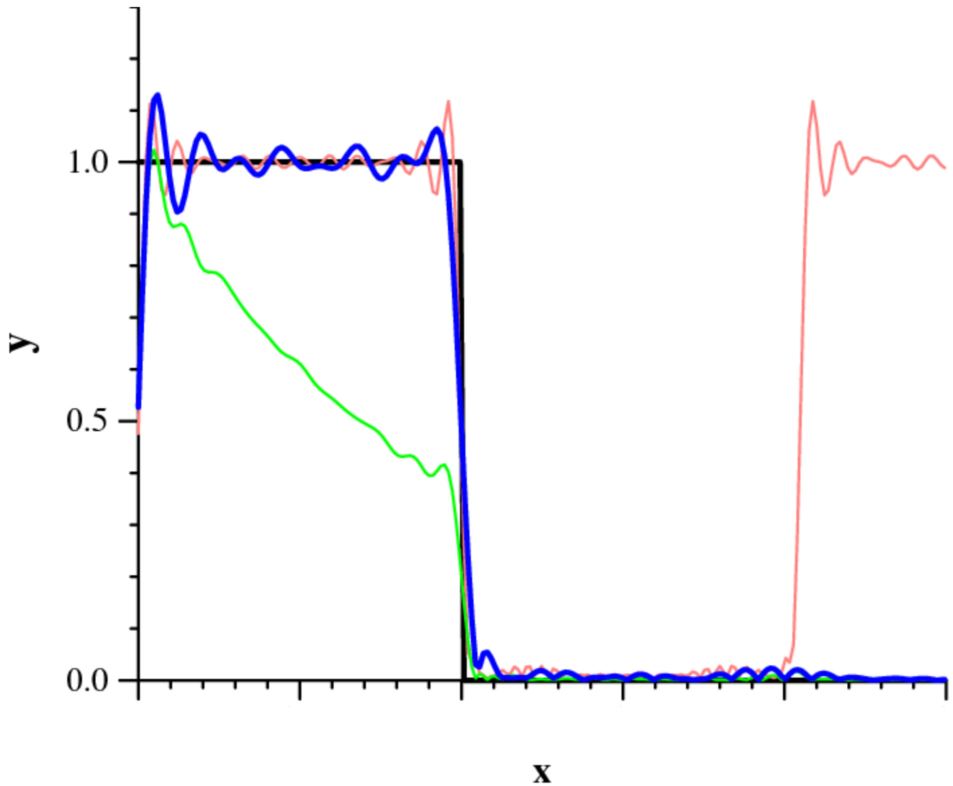


Approximating The Threshold Function



$$\widehat{\mu}(x) = \frac{1}{\eta^x} \sum_{k=-N}^{N} c_k e^{ikx}, \quad \eta > 1$$

Approximating The Threshold Function



y=μ(x)

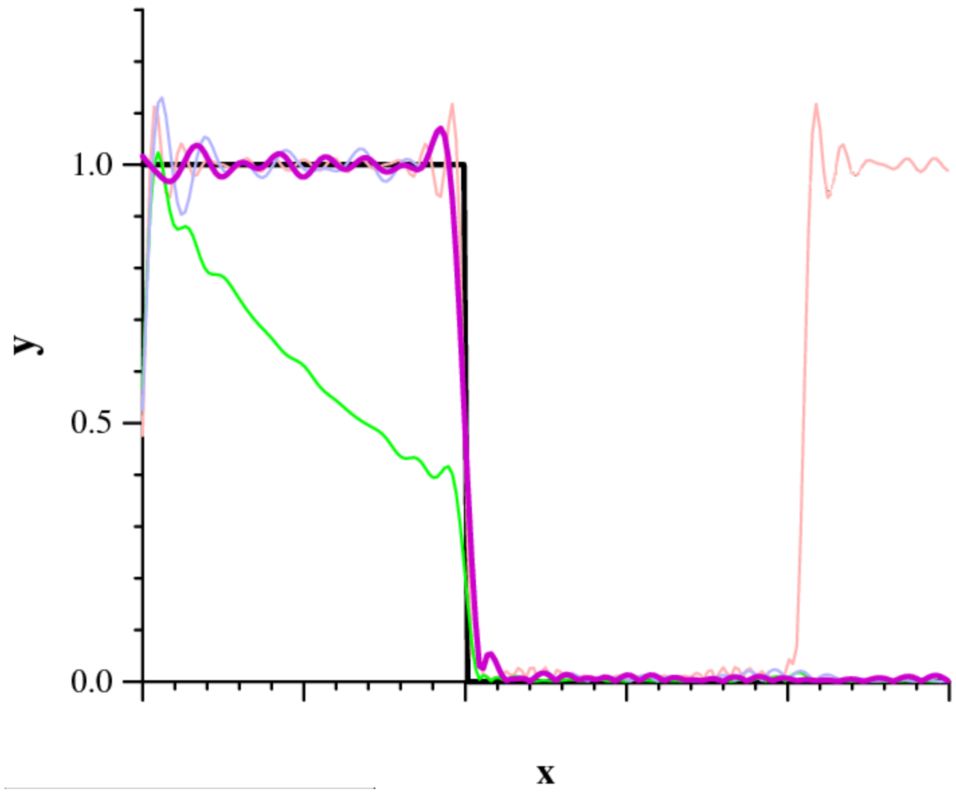
Partial Fourier Series
Periodic !

Damping Factor :
Biased approximation

Initial Scaling :
Unbiased

Apply Fourier series expansion on $\eta^x \mu(x)$

Approximating The Threshold Function



y=μ(x)

Partial Fourier Series
Periodic !

Damping Factor :
Biased approximation

Initial Scaling :
Unbiased

Extending the func.
Unbiased and better quality around origin

The Multi-objective Optimization

- Want to find a set *S* such that

$$\langle \mathbb{E}[\phi_1^{w(S)}], \ldots, \mathbb{E}[\phi_L^{w(S)}] \rangle \simeq \langle \mathbb{E}[\phi_1^{w(S^*)}], \ldots, \mathbb{E}[\phi_L^{w(S^*)}] \rangle$$

- (Approximately) Solve the multi-objective optimization problem with objectives $\mathbb{E}[\phi_i^{w(S)}]$ for *i=1,…,L*

(Similar to [Papadimitriou, Yannakakis FOCS'00])

- Each element *e* has a multi-dimensional weight

$$\langle \mathbb{E}[\phi_1^{w_e}], \ldots, \mathbb{E}[\phi_L^{w_e}] \rangle$$

The Multi-objective Optimization

- Discretize the vector

$$\alpha_{\mathbf{e}} = \left\langle \lfloor \frac{-\ln|\mathbb{E}[\phi_1^{w_e}]|}{\gamma} \rfloor, \lfloor \frac{\arg(\mathbb{E}[\phi_1^{w_e}])}{\gamma} \rfloor, \ldots, \lfloor \frac{-\ln|\mathbb{E}[\phi_L^{w_e}]|}{\gamma} \rfloor, \lfloor \frac{\arg(\mathbb{E}[\phi_L^{w_e}])}{\gamma} \rfloor \right\rangle$$

- *Enumerate all possible 2L-dimensional vector A (poly many)*
  - *Think A as the approximate version of* $\langle \mathbb{E}[\phi_1^{w(S^*)}], \ldots, \mathbb{E}[\phi_L^{w(S^*)}] \rangle$
- *Check if there is any feasible solution S such that*

$$\sum_{e \in S} \alpha_{\mathbf{e}} = A$$

  - *We use the pseudo-poly algorithm for the exact problem*

Summary

- We need L=poly(1/ϵ) terms
- We solve an O(L) dim optimization problem
- The overall running time is $n^{poly(1/\epsilon)}$
- This improves the $n^{2^{poly(1/\epsilon)}}$ running time in [Bhalgat, Goel, Khanna. SODA'11]

Extension: Multi-dimensional Weights

Stochastic Multidimensional Knapsack:

- Each element *e* is associated with a random vector $\omega_e = \langle \omega_{e1}, \ldots, \omega_{ed} \rangle$ (entries can be correlated) *d* =O(1)
- Each element *e* is associated with a profit p_e
- Objective: Find a set *S* of items such that

$$\Pr(\bigwedge_{j=1}^{d}(\sum_{i\in S} w_{ij} \leq 1)) \geq \gamma$$

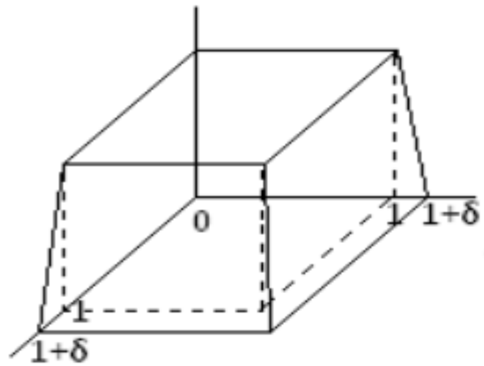
and the total profit $\sum_{e\in S} p_e$ is maximized

Our result: We can find a set *S* of items in poly time such that the total profit is at least (1-ϵ) of the optimum and

$$\Pr(\bigwedge_{j=1}^{d}(\sum_{i\in S} w_{ij} \leq 1+\epsilon)) \geq (1-\epsilon)\gamma$$

Extension: Multi-dimensional Weights

- Consider the 2-dimensional utility function



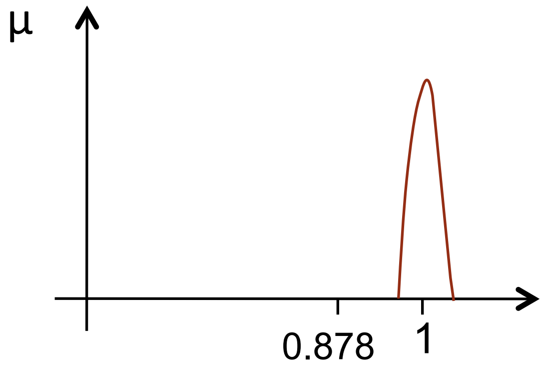
A continuous version of the 2d threshold function

- Consider the rectangular partial sum of the 2d Fourier series expansion

$$S_N f(x,y) = \sum_{|k_1|\leq N} \sum_{|k_2|\leq N} c_k e^{i(k_1 x + k_2 y)}.$$

Recent Progress

- Other problems in P? min-cut, matroid base
  - There is no pseudo-poly time algorithm for EXACT-min-cut
    - Why? EXACT-min-cut is strongly NP-hard (harder than MAX-CUT)
  - It is a long standing open problem whether there is a pseudo-poly-time algo for EXACT-matroid base
- Can we get the same result for these problems?
  - Not for min-cut
  - Consider a deterministic instance where MAX-CUT=1



Recent Progress

- (1) Monotone and Lipschitz nonincreasing utility function
  (2) PTAS for the deterministic multi-dimensional version
  then we can obtain S such that

$$E[\mu(w(S))] \geq OPT-\varepsilon,$$

for any fixed $\varepsilon > 0$.

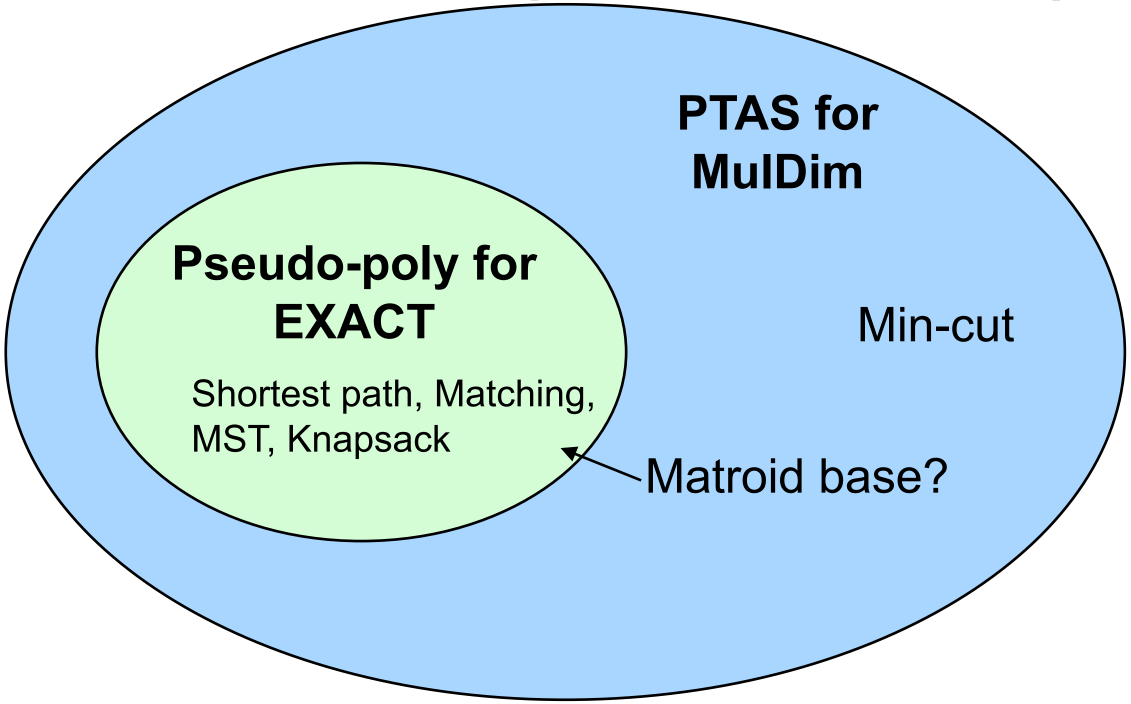
- ✓ Deterministic multi-dimensional version:
  Each edge e has a weight (const dim) vector v_e.
  Given a target vector V, find a feasible set S s.t. $v(S) <= V$
- ✓ A PTAS should
  either return a feasible set S such that $v(S) <= (1+\varepsilon)V$
  or claim that there is no feasible set S such that $v(S) <= V$

Recent Progress

- *KNOWN:*

  Pseudo-poly for EXACT-A => PTAS for MulDim-A

  [Papadimitriou, Yannakakis FOCS'00]
- PTAS for MulDim-MinCut [Armon, Zwick Algorithmica06]
- PTAS for MulDim-Matroid Base [Ravi, Goemans SWAT96]



Open Problems

- OPEN: Can we extend the technique to the adaptive problem (considered in [Dean,Geomans,Vondrak, FOCS'04] [Bhalgat, Goel, Khanna. SODA'11])?
- OPEN: Can we get a PTAS? (with a multiplicative error)
  - Even for optimizing the overflow probability
- Consider maximization problems and increasing utility functions
- Optimizing $\mathbb{E}[\mu(\max_{e\in S} w_e)]$ or $\mathbb{E}[\mu(\min_{e\in S} w_e)]$
  - For $\mathbb{E}[\max_{e\in S} w_e]$ and $\mathbb{E}[\min_{e\in S} w_e]$, see [Kleinberg, Rabani, Tardos. STOC'97], [Goel, Guha, Munagala PODS'06]

Thanks

lijian83@mail.tsinghua.edu.cn