Convex Stochastic Bounds and Stochastic Optimisation on Graphs

JM Fourneau for J. Cohen

LRI France

J. Cohen (LRI-CNRS, France), A. Fauquette (LRI-CentraleSupelec, France), J.M. Fourneau (DAVID, France), N. Pekergin (LACL, France)

Motivation

- Optimisation on capacited networks
- where capacity (lengths, costs, weights) are random variables
- For instance Max-Flow when the capacity of the arcs are random integers.
- We only consider problems which are polynomial where the capacities are deterministic (a large class of problems with many applications: Max Flow, Shortest Path, "s-t" reliability, Completion Time of a Task graph).
- Most of these problems are NP hard when the capacities are random.
- We propose a new approach based on stochastic comparison and monotonicity to provide stochastic bounds on the distribution of the result.
- Most of the classical approches in the littrature are based on the structure of the graph (for instance Serie-Parallel subgraph)

Key Ideas

- We consider discrete distributions on a totally ordered space, for instance $\{1, 2, \ldots, n\}$ and $\leq$.
- Conditioning on each random variable

$$Pr(Rand - Res(D1, \ldots, D_{Nv}) = T) = \sum_{(d1 \ldots dNv) \in \Omega} Pr(d1, \ldots, dNv) 1_{Det-Res(d1, \ldots, dNv)=T}$$

- The complexity of the distribution which appears in the computation is the number of atoms in the distribution
- A deterministic R.V. is a distribution with a single atom.
- Question: Find a new distribution with less atoms which will be a bound according to a stochastic order.
- Here we consider convex and increasing convex ordering and we prove an algorithm to find optimal upper bounds according to the convex order.

The convex order: Why ?

- because it is supposed to be more precise than the strong stochastic order
- because "Max" is a convex operator and "Min" is a concave operator
- and operators "Min" and "Max" are the key operators in all the problems we consider
  - MAX-Flow: Min et + (because Max-Flow = Min Cut).
  - Shortest Path: Min et +
  - "s-t" Reliability: existence of a path from s to t: a variant of Max-Flow
  - Completion Time: Max and +

The convex order: Definition?

Definition (increasing convex ordering)

Let X and Y be two random variables, $X \preceq_{icx} Y$ if for all increasing convex function Φ, $\mathbf{E}[\phi(X)] \leq \mathbf{E}[\phi(Y)]$ if the expectations exist.

Definition (stochastic convex ordering)

Let X and Y be two random variables, $X \preceq_{cx} Y$ if $\mathbf{E}[X] = \mathbf{E}[Y]$ and $X \preceq_{icx} Y$.

Property (Stop Loss)

Let X and Y be two random variables, $X \preceq_{cx} Y$ if and only if $\mathbf{E}[X] = \mathbf{E}[Y]$ and, for all d we have, $\mathbf{E}[(X-d)^+] \leq \mathbf{E}[(Y-d)^+]$.

Stochastic Monotony

Definition ($\Psi-$Monotony)

A function f is $\Psi-$monotone if for all X and Y random variables such that $X \preceq_\psi Y$, then $f(X) \preceq_\psi f(Y)$.

- $+$, "max" are monotone for the increasing convex ordering.

Property

The CompletionTime of a task graph is monotone for the increasing convex stochastic ordering.

Proof: because it is defined with the "max" and "$+$" operators (see at the end of this talk) .

Measurements: typical approach

- Transform measurements into an average (the expectation)
- Solve the deterministic problem: ADD a systematic BIAIS.

Property

Consider the CompletionTime problem. Assuming that the durations of the task $\mathbf{D}_i$ are random. Replace the random durations by their expectations. The deterministic result is a lower bound for the increasing convex ordering for the distribution and a lower bound for the expectation.

$$\mathbf{E}[\mathbf{D}_i] \preceq_{cx} \mathbf{D}_i \quad \forall i$$

CompletionTime is monotone for the increasing convex ordering:

$$CT(\mathbf{E}[\mathbf{D}_1], \mathbf{E}[\mathbf{D}_2], .., \mathbf{E}[\mathbf{D}_{Nv}]) \preceq_{icx} CompletionTime(\mathbf{D}_1, \mathbf{D}_2, .., \mathbf{D}_{Nv}).$$

Taking the Expectation, and noting that as $CT(.)$ is deterministic, it is equal to its expectation:

$$CT(\mathbf{E}[\mathbf{D}_1], \mathbf{E}[\mathbf{D}_2], .., \mathbf{E}[\mathbf{D}_{Nv}]) \leq \mathbf{E}[CompletionTime(\mathbf{D}_1, \mathbf{D}_2, .., \mathbf{D}_{Nv})].$$

Methodology:

- Balance between the complexity (i.e. the number of atoms) and the accuracy of the bound.
- Expectation: lower bound with 1 atom. Worst bound....
- Step 1: Bound the input distributions to obtain upper and lower bound with a small number of atoms
- Step 2: Condition on the new random variables
  - Step 2.1: Solve the deterministic cases with a polynomial algorithm (they are well-known)
  - Step 2.2: Use the law of total probability to obtain the distribution of the bounds
- How to find convex bounds for a discrete distribution ?
- Convex rather than increasing convex: see the conclusion

Here

- For an arbitrary distribution D with N positive atoms and a convex function r, we prove an algorithm to find $D2$ such that
  - $\mathbf{D} \preceq_{\mathbf{cx}} \mathbf{D2}$
  - **D2** has size $K < N$.
  - **D2** is an optimal bound according to the expectation of function $r(x) = x^2$ (i.e. the second moment).
- Find some basic actions to remove atoms
- Prove the structure of the optimal solution

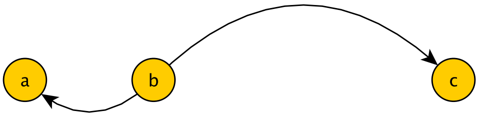
Basic Actions for an upper bound with less atoms

Lemma

We consider an arbitrary discrete distribution (say **D3***) on three atoms a, b, c ($a < b < c$) defined by the positive probabilities p_a, p_b and p_c. Let us define by* **D4** *the distribution of atoms a and c the probabilities of witch (denoted as q_a and q_c) are defined by*

$$q_a + q_c = 1 \quad and, \ aq_a + cq_c = ap_a + bp_b + cp_c.$$

Then, $\mathbf{D3} \preceq_{\mathbf{cx}} \mathbf{D4}$.



Basic Actions for a lower bound with less atoms

Lemma

We consider an arbitrary discrete distribution (say **D1***) on two atoms a and b (without loss of generality we assume that $a < b$) defined by the following positive probabilities p_a and p_b. Let us define by* **D2** *the distribution with a single atom M equal to* $\frac{(a\ p_a + b\ p_b)}{p_a + p_b}$*. Then,* **D2** *is a lower bound for the convex stochastic ordering of* **D1***:* $\mathbf{D2} \preceq_{\mathbf{cx}} \mathbf{D1}$.

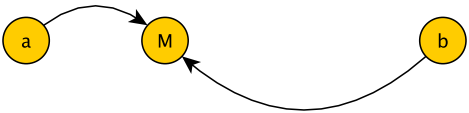


Figure: Fusion of two atoms for a lower bound

Proof: it is a simple application of $\mathbf{E}[X] \preceq_{cx} X$.

Structure of the solution for the optimal upper bound

Lemma

Let **D1** *a distribution on N atoms and* **D2** *a distribution on $K < N$ atoms, such that* **D1** $\preceq_{\mathbf{cx}}$ **D2**. *Let $x_1, x_2, \cdots, x_N$ be the atoms of* **D1** *sorted in increasing order, each value x_i is associated a probability p_i. Similarly, let $u_1, u_2, \cdots, u_K$ be the atoms of* **D2** *sorted in increasing order, associated to probabilities q_i. Then $y_1 = x_1$ and $y_K = x_N$.*

Thus, we know that we must keep x_1 and x_N.

Lemma

We consider distribution **D1** *with 4 atoms a, b, c, f such that $(a < b < c < f)$. Assume that p_a, p_b, p_c, p_f are positive. We apply Lemma 1 to atoms a, b et c to obtain an upper bound (i.e. say* **D2***) on the support $\{a, c, f\}$. We also apply Lemma 1 to atoms a, b et f to obtain* **D3***. This last distribution has also $\{a, c, f\}$ as a support and* **D1** $\preceq_{\mathbf{cx}}$ **D2** $\preceq_{\mathbf{cx}}$ **D3**.

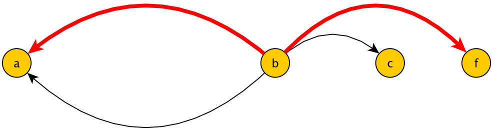


Figure: Comparing upper bounds.

Thus, we expect that it is better to operated on closest atoms. It is formally proved in the paper.

Algorithm for the optimal upper bound

- We represent a distribution as a path in a graph from x_1 to x_N (we know that they are in the optimal solution).
- The path has length $K-1$ (i.e. it has K nodes including x_1 and x_N).
- The cost of an arc from y_i to y_j in the solution is related to the atoms between y_i and y_j in the input distribution (easily computed because of the structural properties when we try to optimize the variance)
- Find a path of length $K-1$ with minimal cost (Bellman-Ford algorithm).

Algorithm for a lower bound

Compute $\Delta(a, b)$ the variation of r when we merge atoms a and b.
At each step find the couple of atoms which minimizes Δ

$$\Delta(a, b) = p_a r(a) + p_b r(b) - (p_a + p_b) r(\frac{p_a a + p_b b}{p_a + p_b}).$$

Require: input distribution **D1**, input size N, output size K
Ensure: Output distribution **D2**

```
D2 = D1.
for all atoms a and all atoms b do
    Compute Δ(a, b) and store it in a data structure.
end for
for i = N down to K + 1 do
    Search for the couple (a, b) which minimizes Δ(a, b).
    Fuse a and b into c in D2.
    Update the matrix Δ(x, y) (remove a and b, add c).
end for
```

Example: A task graph

- Let b_i (resp. e_i) be the time when node i begins (resp. completes) its work.
- $e_i - b_i = w_i$ where w_i is the execution time of task i.
- Classical induction: the time to complete the last node of the graph e_{Nv} can be recursively computed by the following sequence:

$$e_i = w_i + max_{j \in \Gamma^-(i)}(e_j)$$

where $\Gamma^-(i)$ is the set of predecessors of i and the sequence is initialized with

$$e_1 = b_1 + w_1 \quad and \quad b_1 = 0.$$

- A "Max" and "+" recurence formulation. Thus,

The exemple

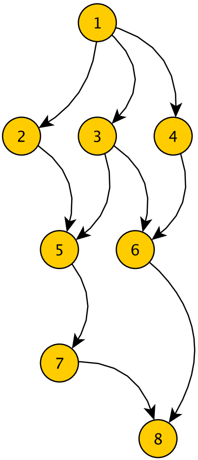
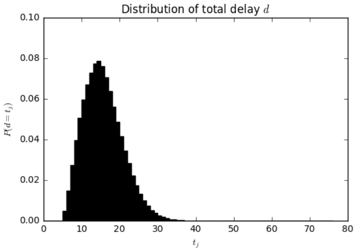


Figure: Example of a task graph with a small size such that we can also compute the exact solution.

Some numerical results for the output



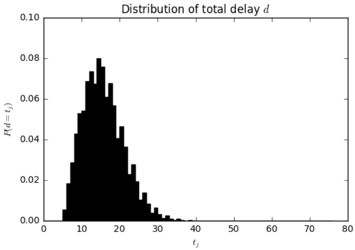


Figure: Distribution of the total delay for the execution of the task graph (left: exact result 10 atoms, right: upper bound with 5 atoms)

Reducing the complexity from 10^8 to 5^8.

Some numerical results 2

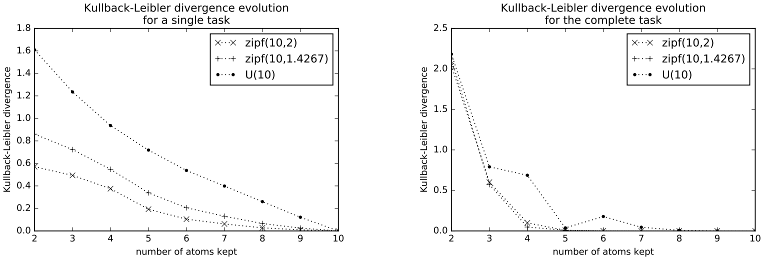


Figure: Kullback-Leibler divergence between the exact and optimal upper bound distributions. For input distributions (left), for the whole task graph distribution of time (right).

MIN based problems

- "Min" is not convex: it is concave.
- The concave ordering also exists.

Definition (stochastic concave ordering)

Let X and Y be two random variables, $X \preceq_{cv} Y$ if $\mathbf{E}[X] = \mathbf{E}[Y]$, and for all concave function Φ, $\mathbf{E}[\phi(X)] \leq \mathbf{E}[\phi(Y)]$ if the expectations exist.

- And convex ordering and concave ordering ($\preceq_{cv}$) are closely related:

Property

If $X \preceq_{cx} Y$ *then* $Y \preceq_{cv} X$

- Because the expectations are equal (thus it is not true anymore for $\preceq_{icx}$ and $\preceq_{icv}$ ordering).

Concave ordering and $\preceq_{icv}$- Monotonicity

Definition (increasing concave ordering)

Let X and Y be two random variables, $X \preceq_{icv} Y$ if for all increasing concave function Φ, $\mathbf{E}[\phi(X)] \leq \mathbf{E}[\phi(Y)]$ if the expectations exist.

Property

Let X and Y be two random variables, $X \preceq_{cv} Y$ impies that $X \preceq_{icv} Y$.

Definition ($\preceq_{icv}$ −Monotony)

A function f is $\preceq_{icv}$ −monotone if for all X and Y random variables such that $X \preceq_{icv} Y$, then $f(X) \preceq_{icv} f(Y)$.

- $+$, "min" are monotone for the increasing concave ordering.

Checking concave ordering

Property

Let X and Y be two random variables, $X \preceq_{icv} Y$ if and only if for all d,

$$\mathbf{E}[min(d, X)] \leq \mathbf{E}[min(d, Y)]$$

- Compare to

Property

Let X and Y be two random variables, $X \preceq_{icx} Y$ if and only if for all d,

$$\mathbf{E}[max(d, X)] \leq \mathbf{E}[max(d, Y)]$$

Max-Flow and Shortest Path problems

Property

The MAX-Flow and Shortest Path problems are monotone for the increasing concave ordering.

- Proof: because it is defined with the "MIN" and "+" operators. Remember that MAX-Flow = MIN-Cut.
- Thus the algorithm we provided can also be used to obtain lower bound for the concave ordering and lower bound for the MAX-FLOW and Shortest-Path problems as well, using the concavity of MIN operator.
- Next step: Combine structural approches (based on the graph) and stochastic ordering.

Typical Bias

- For input distributions, $X \preceq_{cv} \mathbf{E}[X]$
- Thus replacing input distributions by their expectations, we get

$$MaxFlow(RandomCapacity) \preceq_{icv} MaxFlow(DeterministicCapacity)$$

where $DeterministicCapacity = \mathbf{E}[RandomCapacity]$

- Taking the expectation: (a Jensen-like inequality)

$$\mathbf{E}[MaxFlow(RandomCapacity)] \leq MaxFlow(\mathbf{E}[RandomCapacity])$$

Small Exemple (Thanks to Yann Strozecki)

- Consider a path of length $N \geq 2$
- Each arc has a random capacity equal to 0 with probability 1/2 and 2 with probability 1/2.
- thus the expectation of the capacity is equal to 1.
- And the Max-Flow for the deterministic network (with capacity equal to the expectation of the capacity) is also 1.
- If we consider the stochastic network, the arcs are independent. Thus the flow is equal to 2 with probability $(\frac{1}{2})^{N-1}$.
- And the expectation is $(\frac{1}{2})^{N-2}$
- Clearly $(\frac{1}{2})^{N-2} \leq 1$.

Exemple

- But it also work when we bound a limited (i.e. not all) input distributions
- Algorithm (not completely original)
  1. Replace all the distributions by their expectation
  2. Find a minimal cut-set
  3. Consider the initial problem and replace all the distributions except the arcs in the cut-set by their expectations
  4. Compute the $\preceq_{icv}$ bound
  5. Iterate at step 2 with another cut-set to obtain several $\preceq_{icv}$ bounds (at least to improve the first and second moments).
- Can we combine (i.e. improve) $\preceq_{icv}$ bounds like we did for strong stochastic bounds ???

Conclusion

- We have optimal upper bounds for $\preceq_{cx}$ ordering and optimal lower bounds for $\preceq_{cv}$ ordering (same algorithms because $X \preceq_{cv} Y$ impies that $X \preceq_{icv} Y$).
- We have $\preceq_{icx}$ ordering for MAX,+ based problems (Critical Path, Pert)
- We have $\preceq_{icv}$ ordering for MIN,+ based problems (Shortest Path, Max-Flow)
- The methods also works when the random variables are not independent (however we must know the conditional distributions). Typically for transportation problems.
- How to combine structural approaches and stochastic bounds approach ?