Rutgers University
Center for Operations Research

L. Eötvös University
Department of Operations Research

SINGLE COMMODITY STOCHASTIC NETWORK DESIGN PROBLEM UNDER PROBABILISTIC CONSTRAINT

ANDRÁS PRÉKOPA
MERVE UNUVAR

RUTCOR, RUTGERS CENTER FOR OPERATIONS RESEARCH

L. EÖTVÖS UNIVERSITY,
DEPARTMENT OF OPERATIONS RESEARCH

Summary

1. Definition of a (single commodity) network according to Gale, 1957.
2. Examples:
   - Interconnected Power Systems.
   - Flood Control Reservoir System Design.
   - Parking Lots, Transportation, Location Problems.
3. The Gale–Hoffman Feasibility Theorem.
4. The Prékopa–Boros and the Wallace–Wets Theorems.
5. Elimination of the Redundant Gale–Hoffman Inequalities.
6. A Theorem on *p*-efficient Points.
7. The Capacity Design Problem.
8. Solution by the Dentcheva–Prékopa–Ruszczyński Algorithm.
9. Numerical Example.

A *network* $G = (N, A)$ is a finite collection of nodes N and a subset A of $N \times N$, which is the collection of arcs. We assume that if $(i,k) \in A$, then also $(k,i) \in A$. The *arc capacity* function is a real-valued function $y(i,k)$, $(i,k) \in A$ on the set of arcs. A *flow* is a real-valued function $f(i,k)$, $(i,k) \in A$ which satisfies the conditions

$$\begin{aligned} &f(i,k) + f(k,i) = 0 \\ &f(i,k) \le y(i,k) \text{ for } (i,k) \in A. \end{aligned} \tag{1}$$

The definition of y and f can be extended to the entire set $N \times N$, so we write $f(i,k) = y(i,k) = 0$ for $(i,k) \in N \times N$, and $(i,k) \notin A$. We will use the notation

$$\begin{aligned} y(B,C) &= \sum_{i \in B,\, k \in C} y(i,k) \\ f(B,C) &= \sum_{i \in B,\, k \in C} f(i,k), \end{aligned}$$

where B and C are subsets of N.

A *demand function* $d(i)$, $i \in N$ is a real-valued function on the set of nodes. If $B \subseteq N$, then we assign a demand value $d(B)$, to B which is defined by

$$d(B) = \sum_{i \in B} d(i).$$

A demand function (briefly: *demand*) is said to be *feasible* if there exists a flow f such that

$$f(N, i) \geq d(i) \text{ for every } i \in N. \tag{2}$$

Relations (1) and (2) contain the variables $f(i,k)$, $y(i,k)$ and $d(i)$. It is an important problem to find the projection of the convex polyhedron defined by (1) and (2) onto the space of the variables $y(i,k)$ and $d(i)$, i.e., to give the necessary and sufficient condition in terms of these variables for the existence of a flow satisfying (1) and (2). This problem was solved by Gale (1957) and Hoffman (1960) and the result is contained in the following theorem.

Theorem (Gale and Hoffman). *The demand function* $d(i)$, $i \in N$ *is feasible if and only if, for every set* $S \subseteq N$, *we have the inequality*

$$d(S) \le y(\bar{S}, S).$$

For a short proof the reader is referred to Gale (1957).

In power system engineering one node of the network represents one area. To each node i a deterministic generating capacity x_i is assigned, which is diminished by a random deficiency ζ_i, so that the available generating capacity is $x_i - \zeta_i$. Moreover, there exists a random local demand η_i, corresponding to node i, which is to be satisfied first by the use of the generating capacity $x_i - \zeta_i$.

Let $\xi_i = \eta_i + \zeta_i$, $i \in N$. The function

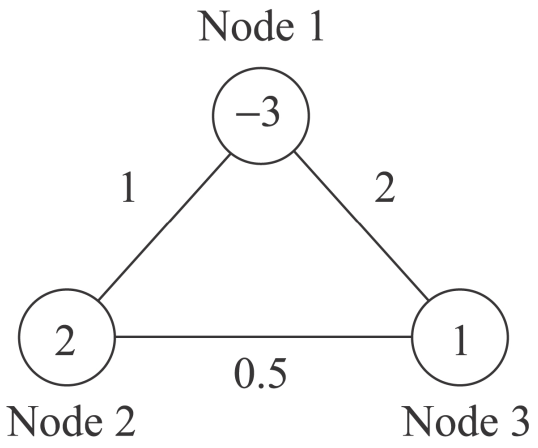
$$d(i) = \xi_i - x_i, \quad i \in N$$

is a demand function corresponding to the network (network demand). If $\xi_i - x_i > 0$, then at node i we need an amount of power $\xi_i - x_i$; and if $\xi_i - x_i < 0$, then at node i there is a surplus generating capacity of $x_i - \xi_i$, which we call the supply. If

$$\sum_{i \in N} x_i \geq \sum_{i \in N} \xi_i$$

then the total available power generating capacity is enough to meet the total demand. However, the transmission system may not be able to allow the individual areas to assist each other to the extent that is necessary. The above stated theorem by Gale and Hoffman provides us with a necessary and sufficient condition for this, i.e., for the existence of a feasible flow.

Simple example: $|N| = 3$



$$d(1) = -3, \quad d(2) = 2, \quad d(3) = 1.$$

There is enough supply but if $y(1,2) = 1$, $y(1,3) = 2$, $y(2,3) = 0.5$ then there is no feasible flow

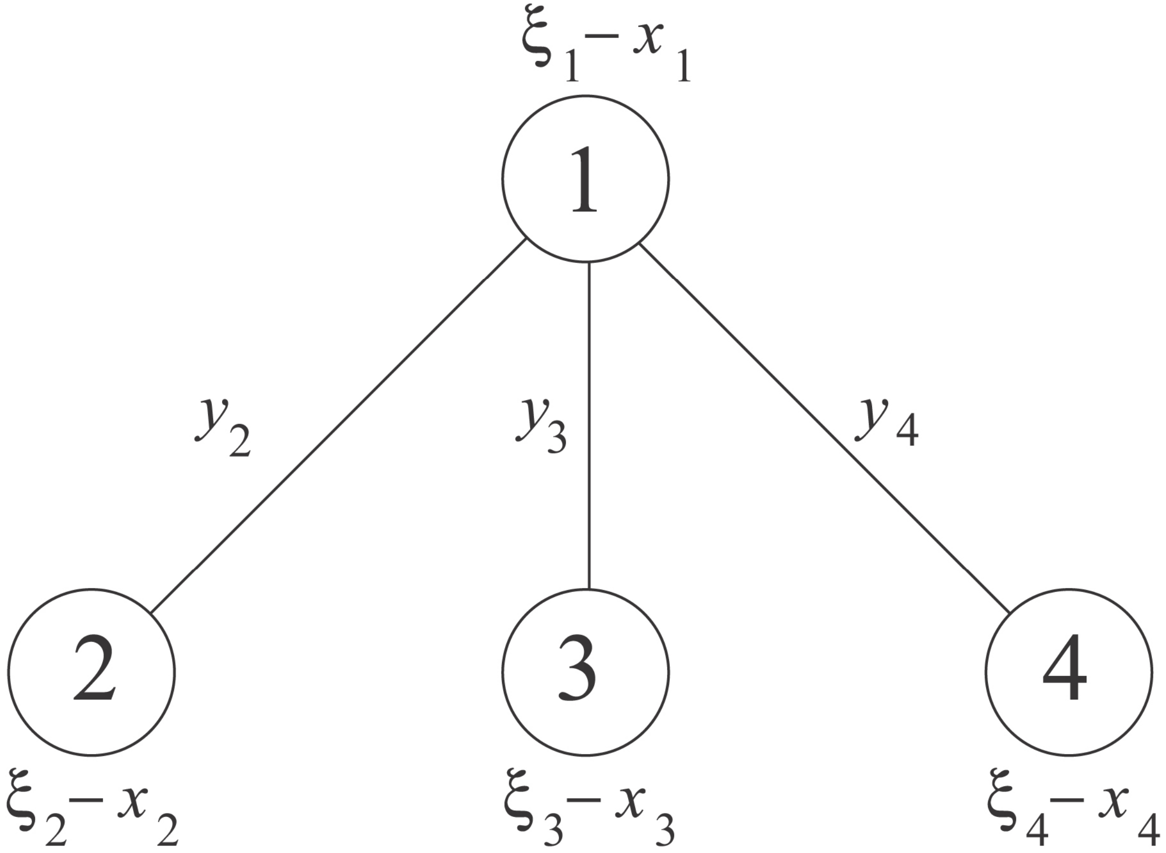
$$d(2) \quad \not\leq \left(\{1,3\}, \ \{1,2\} \right) .$$

Elimination of the Redundant Inequalities

Each Gale–Hoffman inequality corresponds to a subset $S \subset N$. Let (S) designate that inequality.

Theorem (**Prékopa–Boros, 1991,Wallace, Wets, 1993**). The inequality (S) is redundant, among the Gale–Hoffman inequalities, iff at least one of the subgraphs $G(S), G(\bar{S})$ is not connected. In that case the inequality $d(S) \leq y(\bar{S}, S)$ is the sum of other Gale–Hoffman inequalities.

Other elimination procedures also apply, for details see Prékopa, Boros (1989).


ξ₁ − x₁
1
y₂
y₃
y₄
2
3
4
ξ₂ − x₂
ξ₃ − x₃
ξ₄ − x₄

Gale–Hoffman inequalities:

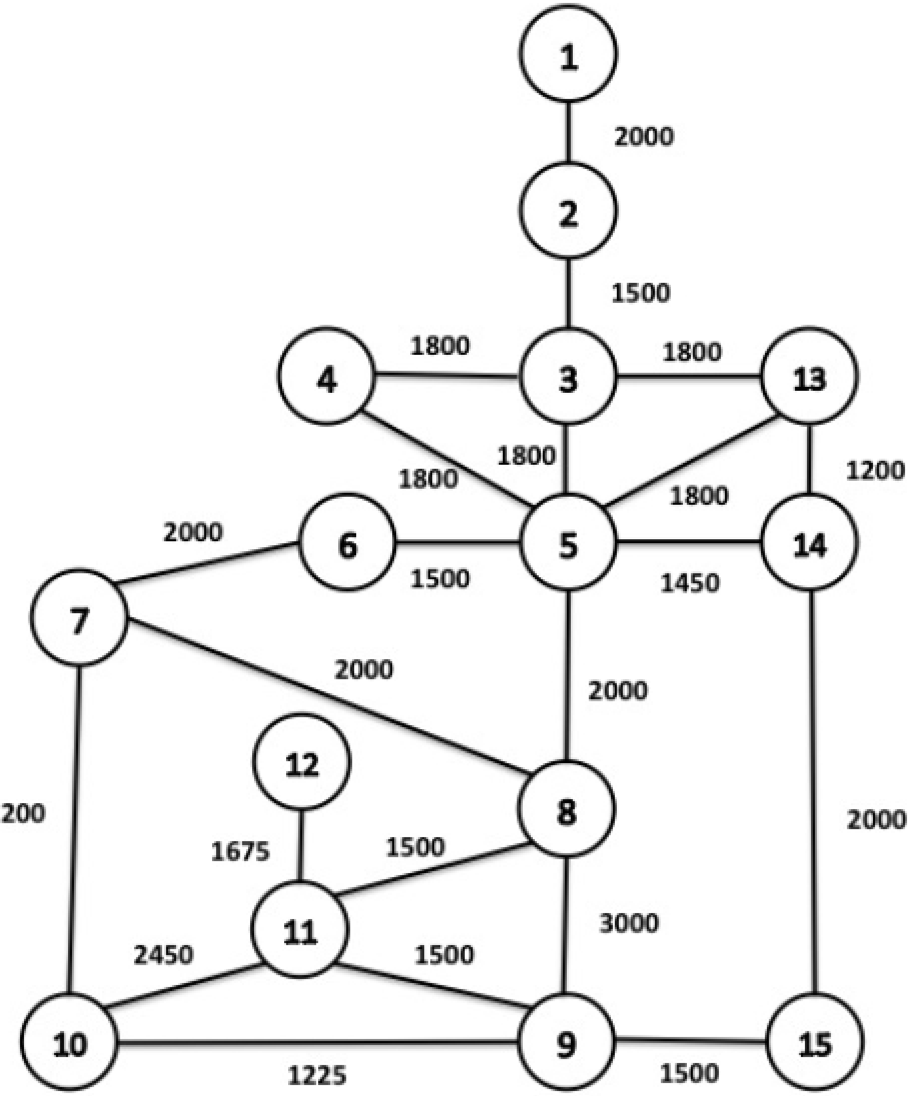
$$
\begin{aligned}
\xi_1 - x_1 + \xi_2 - x_2 + \xi_3 - x_3 + \xi_4 - x_4 &\le 0 \\
\xi_1 - x_1 &\le y_2 + y_3 + y_4 \\
\xi_2 - x_2 &\le y_2 \\
\xi_3 - x_3 &\le y_3 \\
\xi_4 - x_4 &\le y_4 \\
\xi_1 - x_1 + \xi_2 - x_2 &\le y_3 + y_4 \\
\xi_1 - x_1 + \xi_3 - x_3 &\le y_2 + y_4 \\
\xi_1 - x_1 + \xi_4 - x_4 &\le y_2 + y_3 \\
\xi_2 - x_2 + \xi_3 - x_3 &\le y_2 + y_3 \\
\xi_2 - x_2 + \xi_4 - x_4 &\le y_2 + y_4 \\
\xi_3 - x_3 + \xi_4 - x_4 &\le y_3 + y_4 \\
\xi_2 - x_2 + \xi_3 - x_3 + \xi_4 - x_4 &\le y_2 + y_3 + y_4 \\
\xi_1 - x_1 + \xi_2 - x_2 + \xi_3 - x_3 &\le y_4 \\
\xi_1 - x_1 + \xi_2 - x_2 + \xi_4 - x_4 &\le y_3 \\
\xi_1 - x_1 + \xi_3 - x_3 + \xi_4 - x_4 &\le y_2 .
\end{aligned}
$$

Inequalities 9, 10, 11, 12 are sums of others, hence they are redundant.

The Remaining Gale–Hoffman Inequalities in Case of the Four Node Network, After Elimination by Graph Structure

$$
\begin{array}{rrrrllll}
\xi_1 - x_1 & + \xi_2 - x_2 & + \xi_3 - x_3 & + \xi_4 - x_4 & \leq 0 & & & \\
\xi_1 - x_1 & & & & \leq & y_2 & + y_3 & + y_4 \\
 & \xi_2 - x_2 & & & \leq & y_2 & & \\
 & & \xi_3 - x_3 & & \leq & & y_3 & \\
 & & & \xi_4 - x_4 & \leq & & & y_4 \\
\xi_1 - x_1 & + \xi_2 - x_2 & & & \leq & & y_3 & + y_4 \\
\xi_1 - x_1 & & + \xi_3 - x_3 & & \leq & y_2 & & + y_4 \\
\xi_1 - x_1 & & & + \xi_4 - x_4 & \leq & y_2 & + y_3 & \\
\xi_1 - x_1 & + \xi_2 - x_2 & + \xi_3 - x_3 & & \leq & & & y_4 \\
\xi_1 - x_1 & + \xi_2 - x_2 & & + \xi_4 - x_4 & \leq & & y_3 & \\
\xi_1 - x_1 & & + \xi_3 - x_3 & + \xi_4 - x_4 & \leq & y_2 . & &
\end{array}
$$

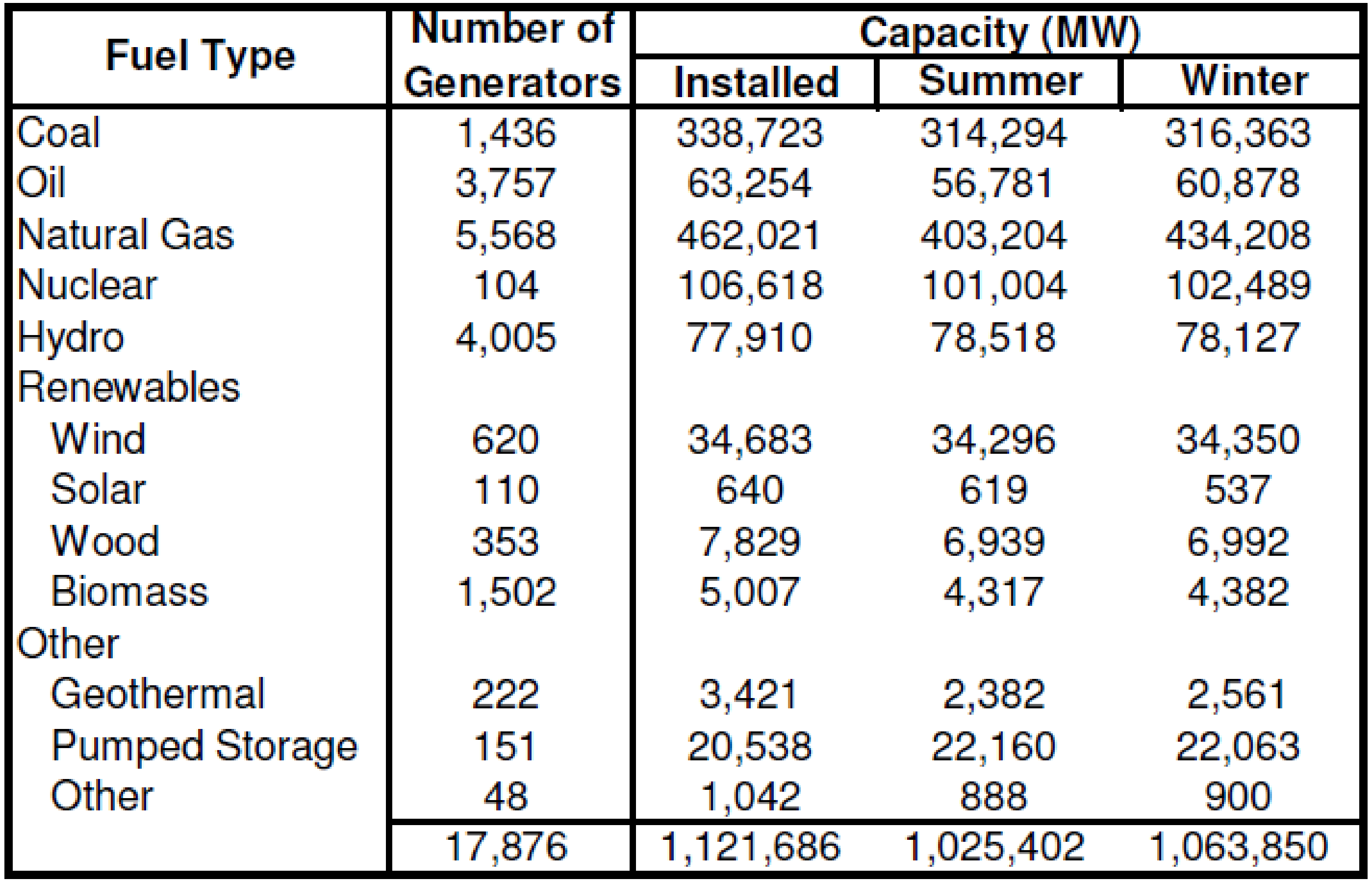
Here no inequality is the sum of others but on the left hand side there are four lines (2, 3, 4, 5), where $\xi_i - x_i$ stands alone. All left hand sides are sums of the left hand sides of these four inequalities.



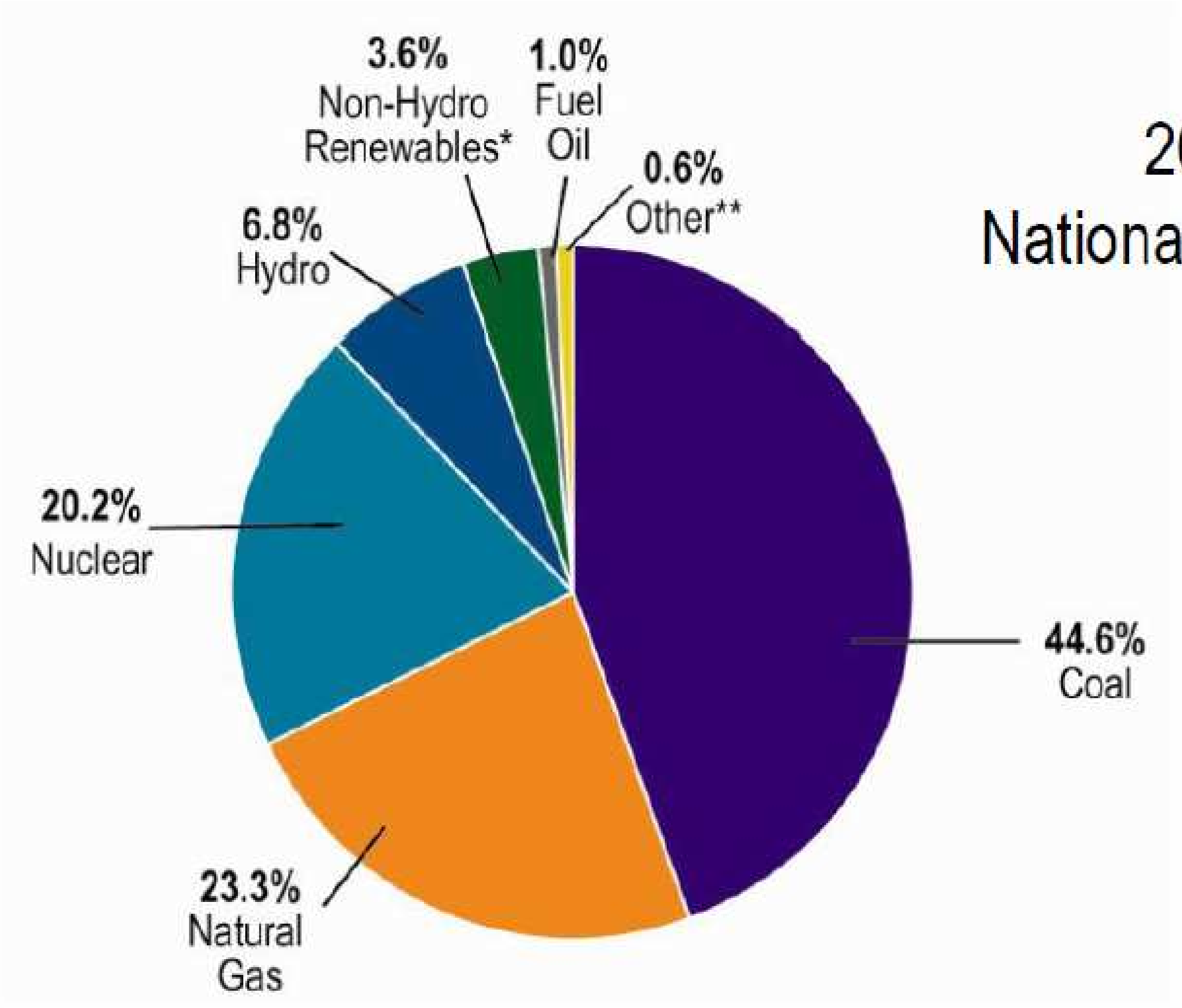
Pacific Gas and Electric Interconnected Power System

Source: Prékopa – Boros, 1989

The table below provides the installed Generation capacity (in MW), by fuel type, in the U.S. in 2009.[3]



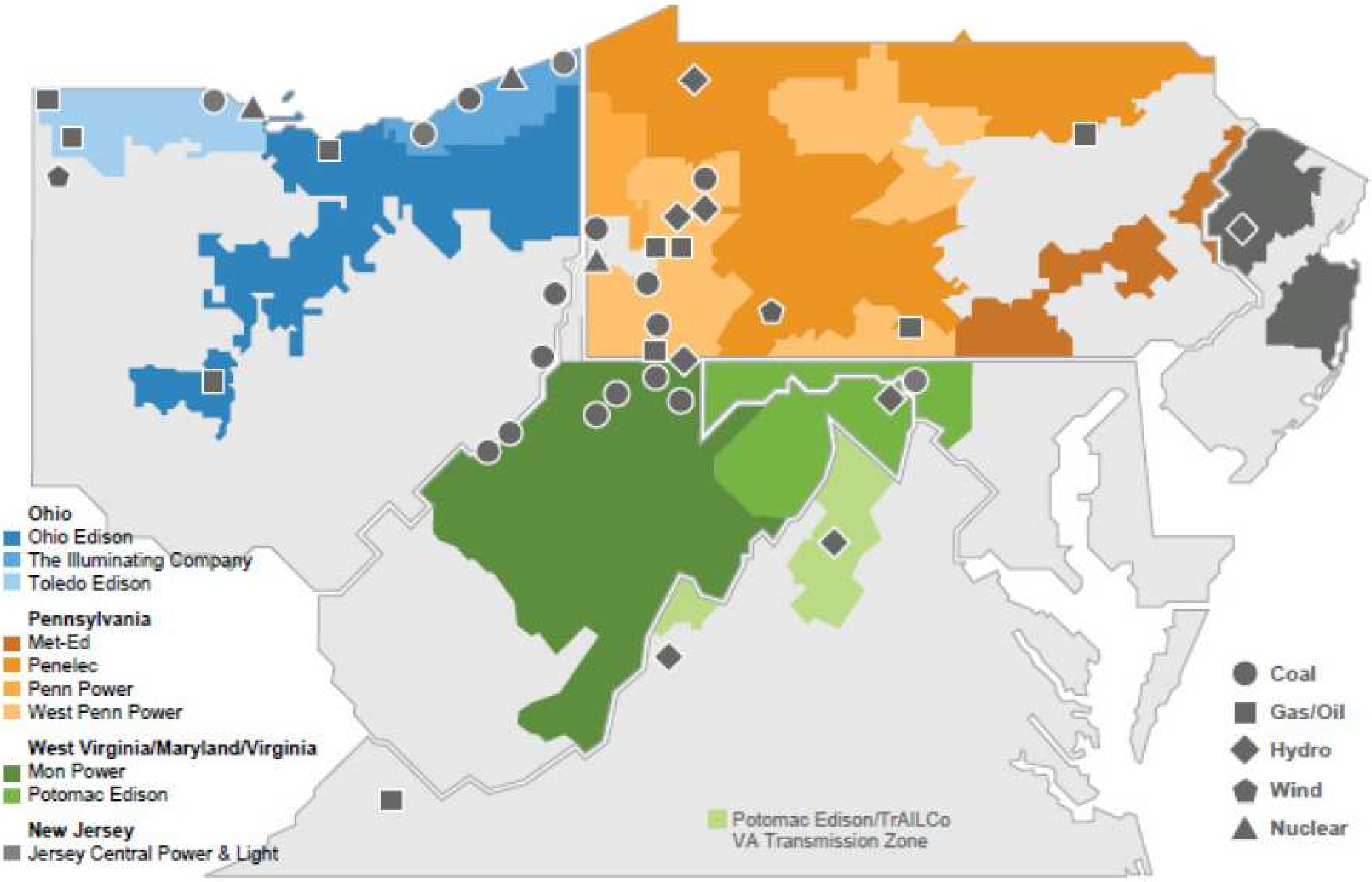
Fuel Type	Number of Generators	Capacity (MW)		
		Installed	Summer	Winter
Coal	1,436	338,723	314,294	316,363
Oil	3,757	63,254	56,781	60,878
Natural Gas	5,568	462,021	403,204	434,208
Nuclear	104	106,618	101,004	102,489
Hydro	4,005	77,910	78,518	78,127
Renewables				
Wind	620	34,683	34,296	34,350
Solar	110	640	619	537
Wood	353	7,829	6,939	6,992
Biomass	1,502	5,007	4,317	4,382
Other				
Geothermal	222	3,421	2,382	2,561
Pumped Storage	151	20,538	22,160	22,063
Other	48	1,042	888	900
	17,876	1,121,686	1,025,402	1,063,850



2009
National Fuel Mix

Estimated Transmission Capacity Additions (Miles)						
2011	2012	2013	2014	2015	2016	Average
2,230	3,831	5,503	3,616	3,970	3,451	3,767

U.S. Peak Demand (MW)				
2005	2006	2007	2008	2009
758,876	789,475	782,227	752,470	725,958


Ohio
Ohio Edison
The Illuminating Company
Toledo Edison
Pennsylvania
Met-Ed
Penelec
Penn Power
West Penn Power
West Virginia/Maryland/Virginia
Mon Power
Potomac Edison
New Jersey
Jersey Central Power & Light
Potomac Edison/TrAILCo VA Transmission Zone
Coal
Gas/Oil
Hydro
Wind
Nuclear

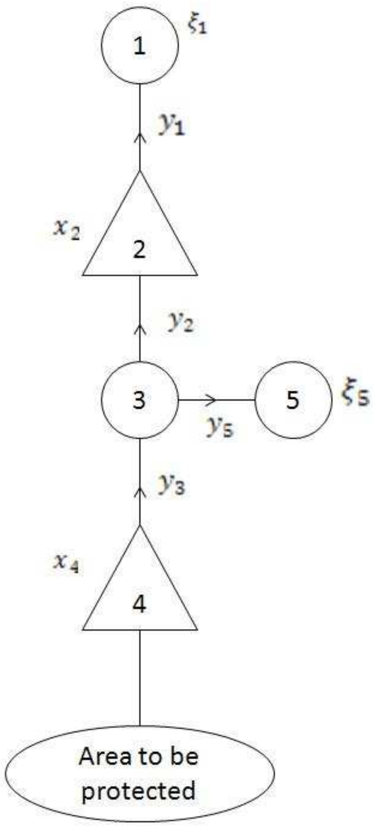
Other examples for networks, from the point of view of application:

- Flood control hydraulic networks.
- Evacuation network
- Transportation networks with parking facilities.

In these cases

ξ_i = demand for freeboard, parking place, shelter room

x_i = capacity for the same at node .



Flood control hydralic network.

Source: Prékopa–Szántai,1978.

ξ_1, ξ_5 flood amounts to be retained or demands for freeboard, $\xi_2 = \xi_3 = \xi_4 = 0$

x_4, x_5 reservoir capacities, $x_1 = x_3 = x_5 = 0$

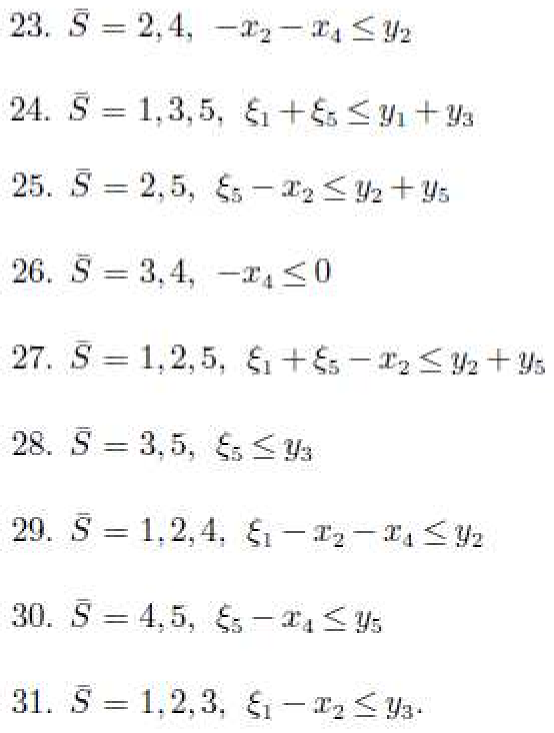
$y_1, \mathrm{y}_2, y_3, \mathrm{y}_5$ arc capacities

$d(i) = \xi_i - x_i$, $\mathrm{i} = 1, \ldots, 5$ system demand function

Gale-Hoffman Inequalities 2^5-1=31

1. $S = \emptyset$, trivial
2. $\bar{S} = N$, $\xi_1 + \xi_5 \leq x_2 + x_4$
3. $\bar{S} = 1$, $\xi_1 \leq y_1$
4. $\bar{S} = 2, 3, 4, 5$, $\xi_5 - x_2 - x_4 \leq 0$
5. $\bar{S} = 2$, $-x_2 \leq y_2$
6. $\bar{S} = 1, 3, 4, 5$, $\xi_1 + \xi_5 - x_4 \leq y_1$
7. $\bar{S} = 3$, $0 \leq y_3$
8. $\bar{S} = 1, 2, 4, 5$, $\xi_1 + \xi_5 - x_2 - x_4 \leq y_2 + y_5$
9. $\bar{S} = 4$, $-x_4 \leq 0$
10. $\bar{S} = 1, 2, 3, 5$, $\xi_1 + \xi_5 - x_2 \leq y_3$
11. $\bar{S} = 5$, $\xi_5 \leq y_5$
12. $\bar{S} = 1, 2, 3, 4$, $\xi_1 - x_2 - x_4 \leq 0$
13. $\bar{S} = 1, 2$, $\xi_1 - x_2 \leq y_2$
14. $\bar{S} = 3, 4, 5$, $\xi_5 - x_4 \leq 0$
15. $\bar{S} = 1, 3$, $\xi_1 \leq y_1 + y_3$
16. $\bar{S} = 2, 4, 5$, $\xi_5 - x_2 - x_4 \leq y_2 + y_5$
17. $\bar{S} = 1, 4$, $\xi_1 - x_4 \leq y_1$
18. $\bar{S} = 2, 3, 5$, $\xi_5 - x_2 \leq y_2 + y_5$
19. $\bar{S} = 1, 5$, $\xi_1 + \xi_5 \leq y_1 + y_5$
20. $\bar{S} = 2, 3, 4$, $-x_2 - x_4 \leq 0$
21. $\bar{S} = 2, 3$, $-x_2 \leq y_3$
22. $\bar{S} = 1, 4, 5$, $\xi_1 + \xi_5 - x_4 \leq y_1 + y_5$

Gale-Hoffman Inequalities 2^5-1=31 (cont')



23. $\bar{S} = 2, 4,\ \ -x_2 - x_4 \le y_2$

24. $\bar{S} = 1, 3, 5,\ \ \xi_1 + \xi_5 \le y_1 + y_3$

25. $\bar{S} = 2, 5,\ \ \xi_5 - x_2 \le y_2 + y_5$

26. $\bar{S} = 3, 4,\ \ -x_4 \le 0$

27. $\bar{S} = 1, 2, 5,\ \ \xi_1 + \xi_5 - x_2 \le y_2 + y_5$

28. $\bar{S} = 3, 5,\ \ \xi_5 \le y_3$

29. $\bar{S} = 1, 2, 4,\ \ \xi_1 - x_2 - x_4 \le y_2$

30. $\bar{S} = 4, 5,\ \ \xi_5 - x_4 \le y_5$

31. $\bar{S} = 1, 2, 3,\ \ \xi_1 - x_2 \le y_3.$

Remaining Inequalities After Elimination by Network Topology

2. $\xi_1 + \xi_5 \le x_2 + x_4$

3. $\xi_1 \le y_1$

10. $\xi_1 + \xi_5 \le y_3 + x_2$

11. $\xi_5 \le y_5$

13. $\xi_1 \le x_2 + y_2$

14. $\xi_5 \le x_4$

28. $\xi_5 \le y_3.$

More Concise Form

$$\xi_1 \le \min(y_1, x_2 + y_2)$$

$$\xi_5 \le \min(y_3, y_5, x_4)$$

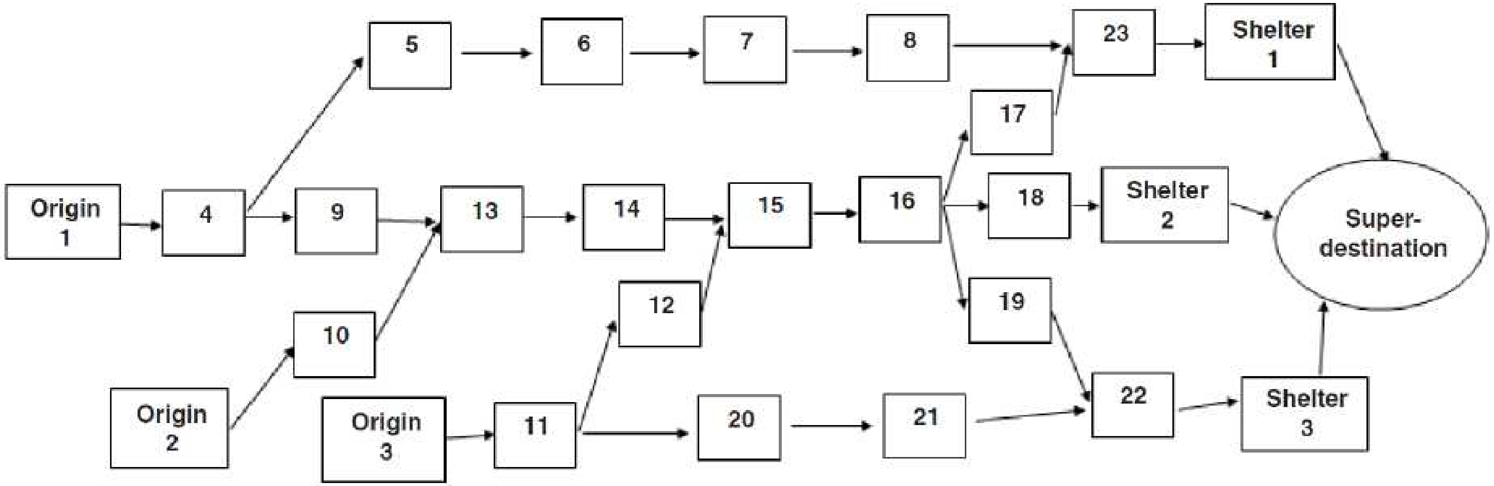
$$\xi_1 + \xi_5 \le \min(x_2 + x_4, y_3 + x_2).$$

If $y_1 = y_2 = y_3 = \infty$, then

$$\xi_1 + \xi_5 \le x_2 + x_4$$

$$\xi_5 \le x_4.$$

Simplified Cell Representation of Cape May Evacuation Network



p-efficient Points of a Multivariate Discrete Probability Distribution

$$\xi = (\xi_1, \ldots, \xi_n)^T, \qquad F(z) = P(\xi \le z), \;\; z \in R^n.$$

Assumption. Each ξ_i has finite support Z_i. Let $Z = Z_1 \times \cdots \times Z_m$.

Definition. The point $z \in Z$ is a p-efficient point of the distribution of ξ, if

$$F(z) \ge p$$

and there is no $y < z$ such that $F(y) \ge p$.

Algorithms to enumerate all *p*-efficient points:

Prékopa, Vizvári, Badics (1998)

Boros, Elbassioni, Gurvich, Khachiyan, Makino (2003).

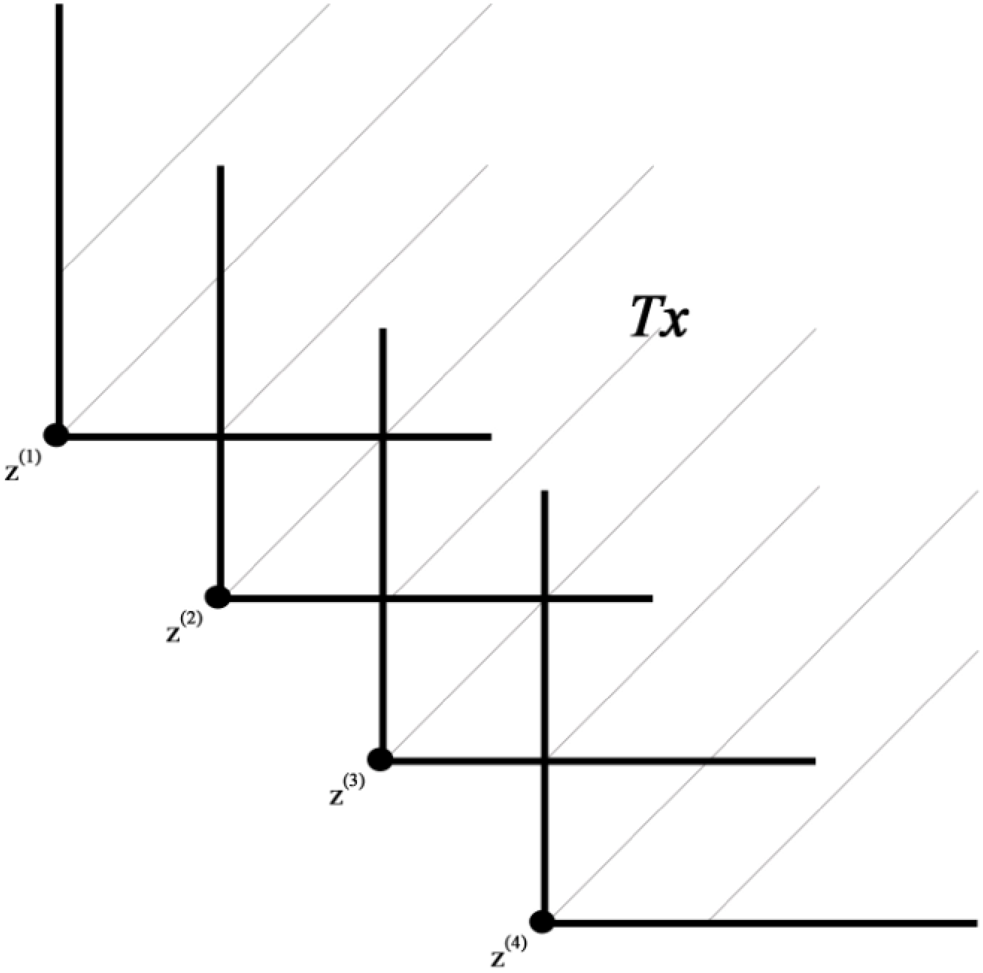
The concept of a p-efficient point has successfully been applied in programming under probabilistic constraint with discrete right hand side random vector:

$$\begin{aligned}&\text{Min } c^T x\\&\text{subject to}\\&P(Tx \geq \xi) \geq p\\&Ax = b,\ x \geq 0.\end{aligned}$$

If $z^{(1)}, \dots, z^{(M)}$ are the p-efficient points of ξ, then the probabilistic constraint is equivalent to the following:

$$Tx \geq z^{(i)},\ \text{for at least one } i = 1, \dots, M.$$

The Constraint $P(Tx \geq \xi) \geq p$ is equivalent to the requirement that Tx is an element of the shaded set, where $z^{(1)}, z^{(2)}, z^{(3)}, z^{(4)}$ are the p-efficient points



The original problem is a disjunctive problem:

$$\begin{aligned}&\text{Min } c^T x\\&\text{subject to}\\&Tx \geq z^{(i)},\ \text{for at least one } i = 1,\dots,M\\&Ax = b,\ x \geq 0.\end{aligned}$$

Through “convexification’’ we obtain a relaxation of it:

$$\begin{aligned}&\text{Min } c^T x\\&\text{subject to}\\&Tx \geq \sum_{i=1}^{M} \lambda_i z^{(i)}\\&\sum_{i=1}^{M} \lambda_i = 1\\&Ax = b,\ x \geq 0,\ \lambda \geq 0,\end{aligned}$$

where the decision variables are the components of x and λ.

A Basic Theorem on *p*-Efficient Points

Let $\xi = (\xi_1, \dots, \xi_n)^T$ be a random vector, where the support of ξ_i is a finite set Z_i, $i = 1, \dots, n$. Let

$$Z = Z_1 \times \cdots \times Z_n$$

and

$$z^{(i)} = (z_1^{(i)}, \dots, z_n^{(i)})^T, \quad i = 1, \dots, M$$

the *p*-efficient points of the random vector ξ.

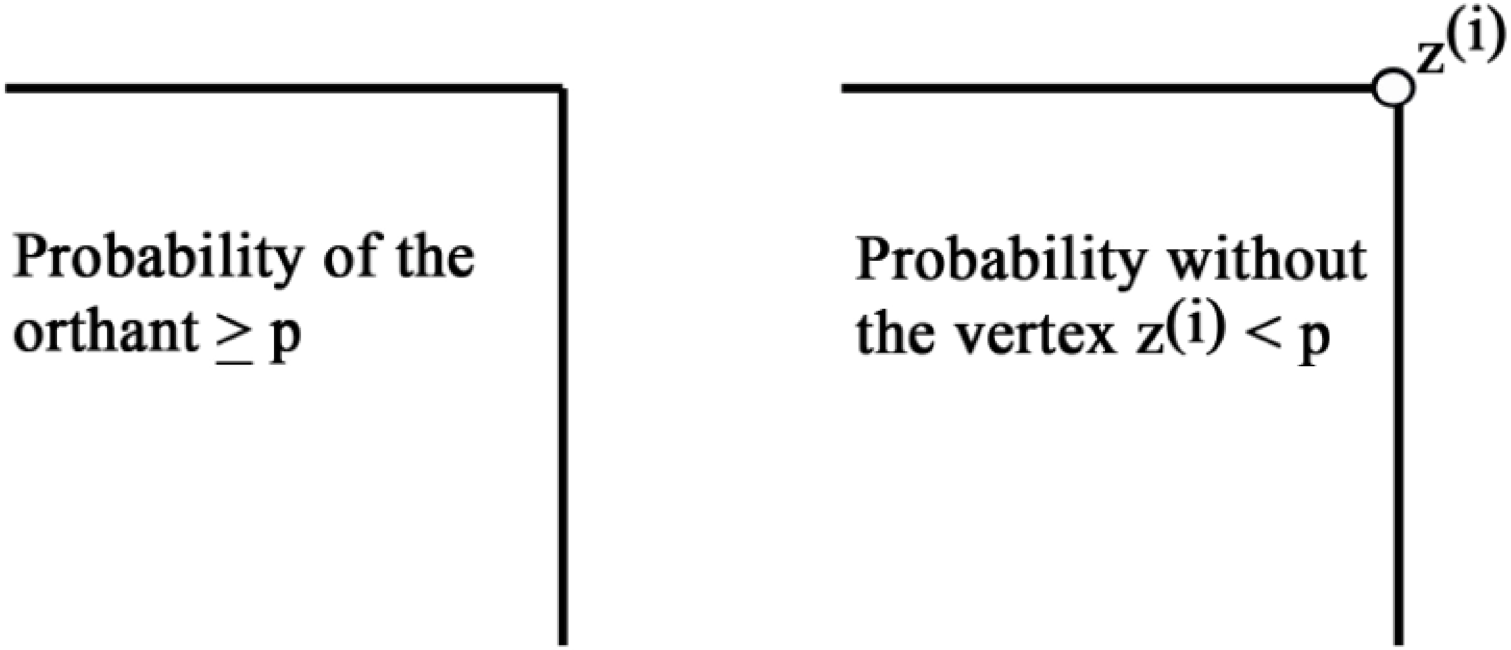
Let $B \geq 0$ be an $n \times M$ matrix that has positive entry in each row.

Assertion: If $P(\{z \in Z \mid z \leq z^{(i)}\} \setminus \{z^{(i)}\}) < p$, $i = 1, \dots, M$, then the *p*-efficient points of the random vector

$$\begin{pmatrix} \xi \\ B\xi \end{pmatrix}$$

are

$$\begin{pmatrix} z^{(i)} \\ Bz^{(i)} \end{pmatrix}, \quad i = 1, \dots, M.$$



- Minor restriction from practical point of view. Slight perturbation of the probability distribution can make the condition satisfied.

Example.

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \qquad B = (1,1)$$

ξ has *p*-efficient points, taken with positive probabilities:

$$\begin{pmatrix} z_1^{(1)} \\ z_2^{(1)} \end{pmatrix}, \quad \begin{pmatrix} z_1^{(2)} \\ z_2^{(2)} \end{pmatrix}, \quad \begin{pmatrix} z_1^{(3)} \\ z_2^{(3)} \end{pmatrix}.$$

Then the random vector

$$\begin{pmatrix} \xi \\ B\xi \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_1 + \xi_2 \end{pmatrix}$$

has *p*-efficient points:

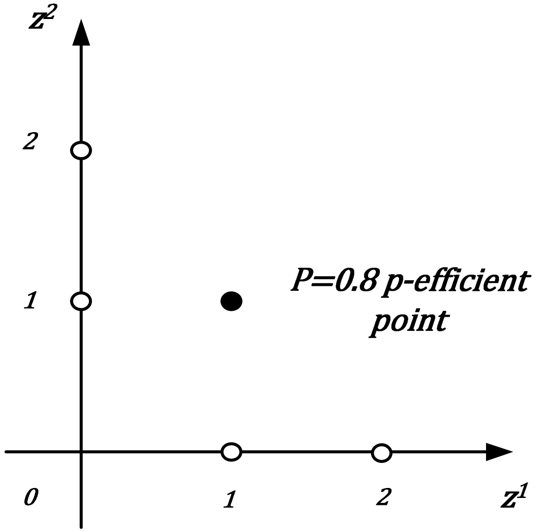
$$\begin{pmatrix} z_1^{(1)} \\ z_2^{(1)} \\ z_1^{(1)} + z_2^{(1)} \end{pmatrix}, \quad \begin{pmatrix} z_1^{(2)} \\ z_2^{(2)} \\ z_1^{(2)} + z_2^{(2)} \end{pmatrix}, \quad \begin{pmatrix} z_1^{(3)} \\ z_2^{(3)} \\ z_1^{(3)} + z_2^{(3)} \end{pmatrix}.$$

The theorem is very important from the point of view of network reliability calculation. If the number of nodes is *n*, then the number of Gale-Hoffman inequalities is 2^n - 1. Still, it is enough to determine the *p*-efficient points of the *n* random demands at the *n* nodes because then we can easily generate the *p*-efficient points for the entire collection of the random variables.

The condition $P(\xi = z^{(i)}) > 0$, $i = 1, \ldots, N$ is essential.

Example: Let $r = 2$, $Z_1 = Z_2 = \{0, 1, 2\}$

$$p_{ij} = P(\xi_1 = i, \xi_2 = j),$$
$$p_{00} = 0.6,\ \ p_{01} = p_{02} = p_{10} = p_{12} = 0.1,$$
$$p_{11} = p_{12} = p_{21} = p_{22} = 0,\ \ p = 0.8$$



ξ has one 0.8-efficient point: (1,1) but (1,1,2) is not 0.8-efficient for $(\xi_1, \xi_2, \xi_1 + \xi_2)$. In fact, $P(\xi_1 \leq 1, \xi_2 \leq 1, \xi_1 + \xi_2 \leq 1) = 0.8$.

The Stochastic Network Design Problem

Decision variables: $x_1, \ldots, x_n$, node capacities, $n = |N|$.

$$\min \left\{ \sum_{i \in N} c_i(x_i) + \sum_{(i,j) \in N \times N} c_{ij}(y_{ij}) \right\}$$

subject to

$$P(d(S) \leq y((S, \bar{S}), S \subset N)) \geq p$$

$$A_1 x + A_2 y \geq b$$

$$x \geq 0, y \geq 0.$$

The static stochastic programming problems can be of probabilistic constrained, recourse(penalty) or hybrid type. Instead of above, we may easily construct a hybrid type model, where the expectation of the measure of violation of the stochastic constraints is incorporated into the objective function.

After the elimination of the redundant Gale-Hoffman inequalities, problem becomes:

$$\min\left\{\sum_{i\in N} c_i(x_i)+\sum_{(i,j)\in NxN} c_{ij}(y_{ij})\right\}$$

subject to

$$P\begin{pmatrix} \xi_k \le x_k+\sum_{(j,k)\in N\times N} y_{jk}, \quad k=1,\ldots,n \\ \sum_{k\in I_j}\xi_k \le \sum_{k\in I_j} x_k+\sum_{k\in I_j}\sum_{(j,k)\in N\times N} y_{jk}, \quad j=1,\ldots,t \end{pmatrix}\ge p$$

$$A_1x+A_2y\ge b$$

$$x\ge 0, y\ge 0,$$

Where I_j, $j=1,\ldots t$ are the subscript sets corresponding to the non-eliminated inequalities, respectively.

Let $\{z^{(1)},\ldots,z^{(M)}\}$ be the set of p-efficient points of ξ. Then the p-efficient points of the random vector

$$\begin{pmatrix} \xi \\ \sum_{k\in I_j} \xi_k \\ j=1,\ldots,t \end{pmatrix}$$

are the points

$$v^{(i)} = \begin{pmatrix} z^{(i)} \\ \sum_{k\in I_j} z_k^{(i)} \\ j=1,\ldots,t \end{pmatrix}, \quad i=1,\ldots,M$$

Assuming linear objective function, the form of the problem is:

$$\min \{c_1^T x + c_2^T y\}$$

$$\text{subject to}$$

$$T_1 x + T_2 y \geq \sum_{j\in J} \lambda_j v^{(j)}$$

$$A_1 x + A_2 y \geq b$$

$$\sum_{i=1}^{M} \lambda_i = 1$$

$$x \geq 0, \ y \geq 0, \ \lambda \geq 0.$$

Two solutions have been developed for the problem

I. Application of a Dentcheva-Prékopa-Ruszczynski algorithm (2000) for the probabilistic constrained stochastic programming problem with random right hand side and discrete distribution. The *p*-efficient points are generated simultaneously with the solution algorithm

II. Application of the Prékopa-Vizvári-Badics algorithm for the same problem. In this case the *p*-efficient points are supposed to be known (enumerated).

We present Solution I., using ideas from Vizvári (2002).

Note that Solution II. has recently been improved by Yamangil-Prékopa (2012): no need to enumerate the *p*-efficient points, they can be generated simultaneously with the algorithm as in solution I.

Introducing the slack variables, the problem can be written in the form:

$$\min\{c_1^T x + c_2^T y + 0^T u + 0^T w + 0^T \lambda\}$$

$$\text{subject to}$$

$$\text{(P)} \qquad \begin{pmatrix} T_1 & T_2 & -E & 0 & 0 \\ A_1 & A_2 & 0 & -E & -V \\ 0^T & 0^T & 0^T & 0^T & e^T \end{pmatrix} \begin{pmatrix} x \\ y \\ u \\ w \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ b \\ 1 \end{pmatrix}$$

$$x \geq 0, \; y \geq 0, \; u \geq 0, \; w \geq 0, \; \lambda \geq 0,$$

Where $V = (v^j, j \in J)$ and $e^T = (1,\ldots,1)$. We subsequently generate the columns of V. Let J_h designate the correct subscript set and $V_h = \{v^{(j)}, j \in J_k\}$. Let P_k designate the corresponding problem (P).

Solve (P_h) and let α be the optimal dual vector. Partition α into α_1 , α_2, α_3, consistent with the constraints. It can be shown that a new column(and variable) can enter the basis in (P_h) iff

$$\min_{i \in J_h} \alpha_2^T v^{(i)} > \min_{i \in J} \alpha_2^T v^{(i)}.$$

and any $v^{(j)}$ for which $\alpha_2^T v^{(i)} < \min_{i \in J} \alpha_2^T v^{(i)}$ is suitable to enter the basis. Note that $\alpha_2 \geq 0$ but under general condition it is possible to guarantee that $\alpha_2 >> 0$.

Note that $z^{(i)} >> 0, i = 1, \ldots, m$ can be achieved if we use $\xi + k$ instead of ξ, where $k >> 0$ is a suitable constant vector. The next theorem provides us with a condition for $w_1 >> 0$.

Theorem. Suppose that there are *n* vectors in the final (primal-dual feasible) basis with linearly independent *p*-efficient upper parts $v^{(i)} >> 0, i = 1, \ldots, n$ such that they are not contained in any of the hyperplanes $e_{\mathrm{i}}^{\mathrm{T}} v = \mathrm{const.}$, $i = 1, \ldots, n.$ Then $\alpha_2 >> 0.$

Proof. We already have the inequality $w_1 > 0.$ The stronger inequality can be obtained by the applications of the theorems of Gordan (1873) and Stiemke (1915).

The p-efficient point $v^{(i)}$ has size $r=n+t$. Its upper n components form the p-efficient point $z^{(i)}$ of ξ. Let F be the c.d.f. of ξ. Since

$$\begin{aligned}\alpha_2^T v^{(i)} &= \sum_{j=1}^{n} \alpha_{2j} z_j^{(i)} + \sum_{h=n+1}^{r} \alpha_{2h} \sum_{j\in I_h} z_j^{(i)} \\ &= \sum_{j=1}^{n} \left(\alpha_{2j} + \sum_{I_h \ni j} \alpha_{2h} \right) z_j^{(i)} = \sum_{j=1}^{n} \gamma_j z_j^{(i)}\end{aligned}$$

and $\gamma = (\gamma_1, \ldots, \gamma_n)^T >> 0$ it follows that $\min_{i\in J} \alpha_2^T v^{(i)}$ can be solved by the solution of the problem:

$$\min \gamma^T z$$

$$\text{subject to}$$

$$F(z) \geq p$$

$$z \in Z^M.$$

For simplicity, assume that ξ takes its values from the integer lattice points of R^n.

When we apply this algorithm for the solution of the network design problem, then the random variable η has size r, equal to the number of non-eliminated Gale–Hoffman inequalities.

Some of the components of η are ξ_1 , …. , ξ_n , the local demands at the nodes of the network, the others are partial sums of them.

By the basic theorem of p-efficient points, it is enough to formulate and solve the knapsack problem based on $\xi = (\xi_1 , \ldots , \xi_n)$, where $n = |N|$, rather than based on η, the number of components of which (after the elimination) may still be very large.

This fact contributes greatly to the efficiency of the solution of the network design problem.

If $\xi_1,\dots,\xi_n$ are independent, then:

$$\min \gamma^T z$$

subject to

$$\sum_{i=1}^{n} (-\log F_i(z_i)) \leq -\log p$$

$$z \in Z, \; l_i \leq z_i \leq u_i, \; i = 1, \dots, n,$$

We also assume that $\xi_1,\dots,\xi_n$ are integer valued (or take values from an arithmetic sequence).

Equivalent Formulation of the Problem

write: $a_{ik} = -\log F_i(k),\ d = -\log p,\ z_i = \sum_{k=l_i}^{u_i} k\delta_{ik}$

$$\min \sum_{i=1}^{n} \sum_{k=l_i}^{u_i} \gamma_i k \delta_{ik}$$

subject to

$$\sum_{i=1}^{n} \sum_{k=l_i}^{u_i} a_{ik}\delta_{ik} \le d$$

$$\sum_{k=l_i}^{u_i} \delta_{ik} = 1,\ \ i = 1, \ldots, M$$

$$\delta_{ik} \in \{0, 1\},\ \ \text{all } i, k,$$

Multiple Choice Knapsack Problem, MCKP

Solution of the MCKP

Relaxation, allowing $0 \le \delta_{ik} \le 1, \text{ all } i,k$

The problem is called Linear Multiple Choice Knapsack Problem, LMCKP

$$\min \sum_{i=1}^{n} \sum_{k=l_i}^{u_i} \gamma_i k \delta_{ik}$$

$$\text{subject to}$$

$$\sum_{i=1}^{n} \sum_{k=l_i}^{u_i} a_{ik} \delta_{ik} \le d$$

$$\sum_{k=l_i}^{u_i} \delta_{ik} = 1, \ \ i = 1, \ldots, n$$

$$0 \le \delta_{ik} \le 1, \ \text{ all } i,k.$$

For efficient solution see Pisinger (1995). However, we have our own, more efficient solution, using ideas from stochastic programming

Introduce slack variable u in the inequality constraint, then split the sum into m terms, each term corresponds to a component of ξ. Changing the summation range to 1 through m_i, $i=1,\ldots,n$ and designate the coefficients in the new sums by $h_{i1}, \ldots, h_{im_i}$, $i=1,\ldots,n$.

The new problems is:

$$\min\left\{0u+0u_1+\cdots+0u_n+h_{11}\delta_{11}+\cdots+h_{1m_1}\delta_{1m_1}+\cdots+h_{n1}\delta_{n1}+\cdots+h_{nm_n}\delta_{nm_n}\right\}$$

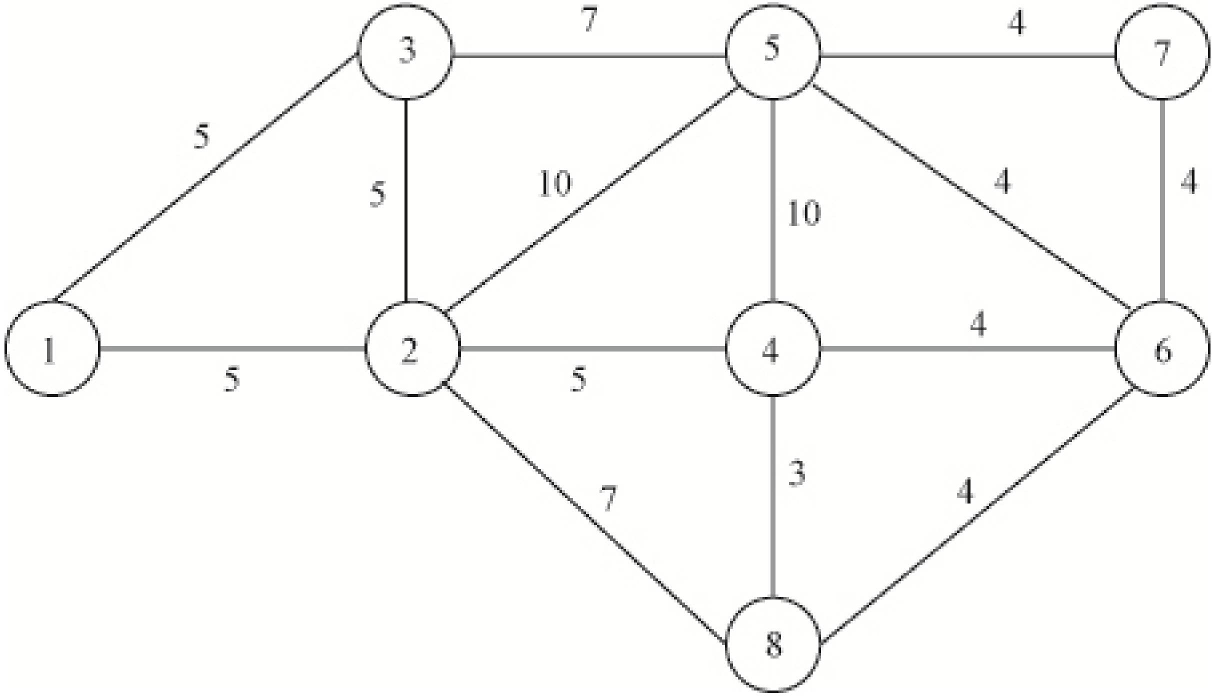
$$\text{subject to}$$

$$\begin{array}{llllr}
u+u_1+\cdots+u_n & & & & =d \\
\quad u_1 & -a_{11}\delta_{11}-\cdots-a_{1m_1}\delta_{1m_1} & & & =0 \\
\qquad \ddots & & & & \vdots \\
\qquad\quad u_n & & & -a_{n1}\delta_{n1}-\cdots-a_{nm_n}\delta_{nm_n} & =0 \\
 & \delta_{11}+\cdots+\delta_{n1} & & & =1 \\
 & & \ddots & & \vdots \\
 & & & \delta_{n1}+\cdots+\delta_{nm_n} & =1
\end{array}$$

$$u_i\ge 0,\ \ i=1,\ldots,n,\ \ \delta_{ik}\ge 0,\ \ \text{all } i,k.$$

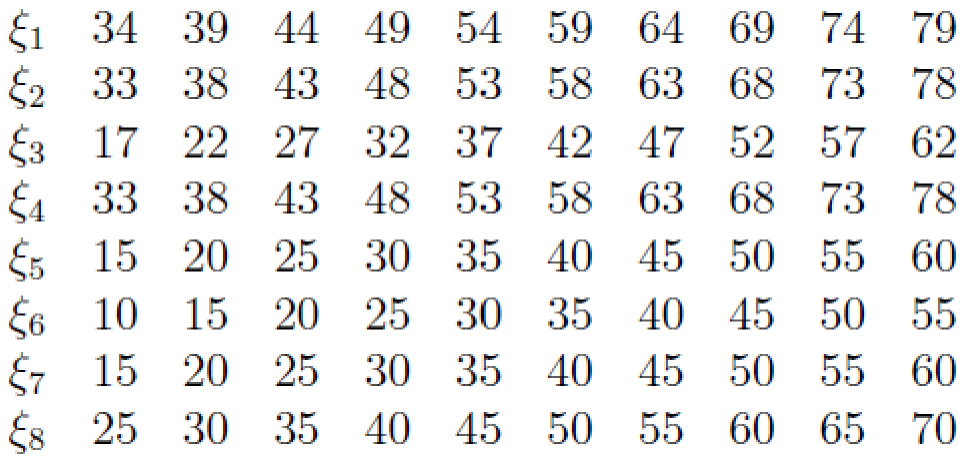
Simple recourse type problem. Fast algorithms and fast bounds in Prékopa (1990, 1995), Fábián, Prékopa, Ruff-Fiedler (1995). In the optimal solution one δ or a pair appears in each block. In the latter case a simple cost-efficient argument given the solution to the MCKP.

Numerical Example



Eight-node Network from Prékopa–Boros, 1989

Possible Values of the Random Demands



ξ_1	34	39	44	49	54	59	64	69	74	79
ξ_2	33	38	43	48	53	58	63	68	73	78
ξ_3	17	22	27	32	37	42	47	52	57	62
ξ_4	33	38	43	48	53	58	63	68	73	78
ξ_5	15	20	25	30	35	40	45	50	55	60
ξ_6	10	15	20	25	30	35	40	45	50	55
ξ_7	15	20	25	30	35	40	45	50	55	60
ξ_8	25	30	35	40	45	50	55	60	65	70

Corresponding Binomial Distribution Function Values

$p(1)$ $(n = 9,\ p = 0.4)$	0.01	0.0704	0.2316	0.4824	0.7332	0.9004	0.9747	0.9959	0.9994	1
$p(2)$ $(n = 9,\ p = 0.45)$	0.0046	0.0385	0.1494	0.3612	0.6212	0.8339	0.9499	0.9905	0.9988	1
$p(3)$ $(n = 9,\ p = 0.5)$	0.0019	0.0194	0.0897	0.2537	0.4997	0.7457	0.9097	0.98	0.9975	1
$p(4)$ $(n = 9,\ p = 0.6)$	0.0002	0.0037	0.0249	0.0992	0.2664	0.5172	0.768	0.9292	0.9896	1
$p(5)$ $(n = 9,\ p = 0.48)$	0.0027	0.0257	0.1109	0.2945	0.5488	0.7835	0.9279	0.985	0.9981	1
$p(6)$ $(n = 9,\ p = 0.35)$	0.0207	0.121	0.3371	0.6087	0.828	0.9461	0.9885	0.9982	0.9995	1
$p(7)$ $(n = 9,\ p = 0.42)$	0.0074	0.0558	0.196	0.4329	0.6902	0.8765	0.9664	0.9943	0.9993	1
$p(8)$ $(n = 9,\ p = 0.38)$	0.0135	0.0881	0.2711	0.5329	0.7735	0.921	0.9812	0.997	0.9994	1

All eight distributions are binomial hence all discrete functions F_i, $i = 1, 2, 3, 4, 5, 6, 7, 8$ are logconcave in the supports of ξ_i, $i = 1, 2, 3, 4, 5, 6, 7, 8$, respectively. The Stochastic Programming Problem to be solved is:

$$\min c^T x$$

$$\text{subject to}$$

$$P(y(S, \bar{S})) \geq d(S),\ (S)\ \text{non-eliminated}) \geq 0.95,\ S \subset N$$

$$0 \leq x_i \leq 100,\ i = 1, \ldots, 8,$$

$$\text{where}\quad d(s) = \sum_{i \in S} (\xi_i - x_i)$$

$$c^T = (2, 3, 4, 2, 1, 1, 7, 4)$$

161 inequalities are non-eliminated. The z vectors of the optimal basic vectors with p-efficient upper parts are:

$$\begin{array}{lcl}
z^{(1)} & = & (64,\quad 68,\quad 57,\quad 73,\quad 55,\quad 45,\quad 55,\quad 65) \\
z^{(2)} & = & (64,\quad 73,\quad 62,\quad 78,\quad 50,\quad 45,\quad 50,\quad 65) \\
z^{(3)} & = & (69,\quad 68,\quad 57,\quad 78,\quad 50,\quad 40,\quad 55,\quad 60) \\
z^{(4)} & = & (59,\quad 73,\quad 62,\quad 73,\quad 55,\quad 40,\quad 55,\quad 55) \\
z^{(5)} & = & (74,\quad 73,\quad 62,\quad 73,\quad 55,\quad 40,\quad 55,\quad 55) \\
z^{(6)} & = & (79,\quad 73,\quad 52,\quad 73,\quad 60,\quad 40,\quad 50,\quad 60) \\
z^{(7)} & = & (74,\quad 68,\quad 57,\quad 73,\quad 50,\quad 50,\quad 50,\quad 60)
\end{array}$$

Optimal Solution:

$$x = (74,\quad 60,\quad 40,\quad 67,\quad 86.1799,\quad 56.8201\quad 47\quad 51)$$

Optimal Value: 1298.

Two-Stage Stochastic Network Design Problem

The second stage problem is a problem that is solved in each period, every day, say. It is the following:

$$\min \quad \sum_{i\in N} c_{1i}\left[f(i,N)\right]_+ \qquad \text{cost of generation}$$

$$+\sum_{i\in N} c_{2i}(z_i) \qquad \text{cost of outage or feasibility}$$

$$+\sum_{(i,j)\in N\times N} c_{ij}\left(\left|f_{ij}\right|\right) \qquad \text{cost of transmission}$$

(S) s.t.

$$f(N,i)+z_i \ge \xi_i - x_i,\ z_i \ge 0,\ i\in N$$

$$f_{ij}+f_{ji}=0$$

$$\left|f_{ij}\right| \le y_{ij},\ (i,j)\in N\times N$$

$$\sum_{i\in N} z_i \le z_u$$

The decision variables are the $z_i,\ i\in N$ and $f_{ij},\ (i,j)\in N\times N$

The first stage problem is the following:

$$
\text{(F)}\quad
\begin{array}{lll}
\min & \sum_{i \in N} c_{3i} x_i & \text{cost of investment in generating capacities} \\
 & + \sum_{(i,j) \in N \times N} g_{ij}\left(y_{ij}\right) & \text{cost of investment in transmission capacities} \\
 & + \mathrm{E}\left(\mathrm{g}\left(\mathrm{x}, \mathrm{y}, \xi\right)\right) & \text{long term average operating cost} \\
\text{s.t.} & & \\
 & \mathrm{P}\left(\mathrm{d}\left(\mathrm{S}\right) \le y\left(S, \bar{S}\right), \text{ all } S \in \varphi\right) \ge p & \\
 & x_L \le x \le x_U & \\
 & y_L \le y \le y_U &
\end{array}
$$

References

1. G. Boole (1854). Laws of Thought. American Reprint of 1854 Edition, Dover, New York.
2. E. Boros, A. Prékopa (1989). Closed Form Two-Sided Bounds for Probabilities that Exactly *r* and at Least *r* out of *n* Events Occur. *Math. Oper. Res.* **14**, 317–342.
3. E. Boros, K. Elbassioni, V. Gurvich, L. Khachiyan, K. Makino (2003). An Intersection Inequality for Discrete Distributions and Related Generation Problems. Lecture Notes in Computer Science 2719, Springer, Berlin-Heidelberg, 543–555.
4. D. A. Dawson, A. Sankoff (1967). An Inequality for Probabilities. *Proc. Amer. Math. Soc.* **18**, 504–507.
5. D. Dentcheva, A. Prékopa, A. Ruszczynski (2000). Concavity and Efficient Points of Discrete Distributions in Probabilistic Programming. *Math, Prog. Series A*, **89**, 55–77.
6. D. Gale (1957). A Theorem on Flows in Networks. *Pacific J. of Math*. **7,** 1073–1082.
7. Th. Hailperin (1965). Best Possible Inequalities for the Probability of a Logical Function of Events. *Amer. Math. Month*. **72**, 343–359.

8. A. J. Hoffman (1960). Some Recent Applications of the Theory of Linear Inequalities to Extremal Combinatorial Analysis. In: *Proceedings of Symposia of Applied Math*. Vol. X. American Math. Soc., 113–117.
9. D. Hunter (1976). An Upper Bound for the Probability of a Union. *J. Appl. Probab*. **13**, 597–603.
10. S. M. Kwerel (1975a). Most Stringent Bounds on Aggregated Probabilities of Partially Specified dependent Probability Systems. *J. Amer. Statist. Assoc*. **70**, 472–479.
11. S. M. Kwerel (1975b). Bounds on Probability of a Union and Intersection of *m* Events. *Adv. Appl. Probab*. **7**, 431–448.
12. A. Prékopa (1980). Network Planning Using Two-Stage Programming Under Uncertainty. Lecture Notes in Economics and Math. Systems **179**, 216–237.
13. A. Prékopa (1988). Boole–Bonferroni Inequalities and Linear Programming. *Oper. Res*. **36**, 145–162.
14. A. Prékopa (1995). *Stochastic Programming*. Kluwer.
15. A. Prékopa, E. Boros (1989). On the Existence of a Feasible Flow in a Stochastic Transportation Network. *Oper. Res*. **39**, 119–129.

16. A. Prékopa, T. Szántai (1978). Flood Control Reservoir System Design Using Stochastic Programming. *Math. Prog. Study* **9**, 138–151.

17. A. Prékopa, T. Rapcsák, I. Zsuffa (1978). Serially Linked Reservoir System Design Using Stochastic Programming. *Water Resources Research* **14**, 672–678.

18. A. Prékopa, B. Vizvári, T. Badics (1998). Programming under Probabilistic Constraint with Discrete Random Variable. In: *New Trends in Mathematical Programming* (F. Giannessi et al. eds.), Kluwer, 235–255. See also: RUTCOR Research Report 10–96.

19. B. Vizvári (2002) The integer programming background of a stochastic integer programming algorithm of Dentcheva-Prékopa-Ruszczynski. *Optimization Methods and Software*, **17**, 543-559.

20. Term project works by Ünüvar, Yazici (RUTCOR), Z. Csizmadia (ELTE, Budapest).

21. S. W. Wallace, R. J.-B. Wets (1993). The Facets of the Polyhedral Set Determined by the Gale–Hoffman Inequalities. *Math. Prog*. **62**, 215–222.

22. M. A. Yazici, K. Özbay (2007). Determination of Hurricane Evacuation Shelter Capacities and Locations with Probabilistic Road Capacity Constraints. 86th Annual Meeting of the Transportation Research Board, Washington, D. C., 2007.