Improved Initialization of the Homogeneous Self-Dual Embedding Model for Solving Conic Optimization

Alexandre Belloni, Robert M. Freund, Kim-Chuan Toh, and Allison Chang

MIT and NUS

May, 2008

(paper not yet available)

Interior-Point Methods (IPMs) for LP

- 1984 - Karmarkar's paper
- 1985 - first IPM codes - 20-100 iterations on NETLIB suite, typically 35 iterations
- $\sim$1990 - Mehrotra predictor-corrector, 10-60 iterations on NETLIB suite, typically 25 iterations
- 1992-2007 - no significant computational improvements

IPMs for Convex Nonlinear Optimization

- 1991-94 - Nesterov and Nemirovsky - IPM theory for convex nonlinear optimization, and Alizadeh - identification of semidefinite programming (SDP)

- 1996 - software for SOCP, SDP - 10-60 iterations on NETLIB suite, typically $\sim$30 iterations

- Each IPM iteration is expensive to solve: $\begin{pmatrix} H^k & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$

- $O(n^3)$ work per iteration, managing numerical linear algebra very important

- no recent significant improvements in IPM computation

Goal of this Work

- Enhance understanding of the homogeneous self-dual (HSD) embedding model for conic optimization

- develop methods to compute a "better" initializing interior point in the HSD embedding model

- Reduce the number of IPM iterations needed to solve the HSD model

- Enabling technologies of sorts:
  - random walks on convex sets
  - projective transformations

Iterations and Running Times

Original IPM Iteration Range	Percentage of Instances	Average IPM Iterations: Original	Average IPM Iterations: After Random Walk	Average Total Running Time: Original	Average Total Running Time: After Random Walk	Net Savings (%)
16-18	29%	16.1	19.8	916.6	1167.8	-27%
19-21	16%	20.3	18.3	9.9	9.3	5%
22-24	19%	23.2	19.2	497.2	442.5	11%
25-27	16%	25.7	19.0	102.5	76.6	25%
28-30	6%	29.5	20.3	65.6	47.5	28%
31-35	14%	32.9	19.9	81.1	51.5	36%

Averages in table are arithmetic averages

Primal and Dual Conic Problem

We consider convex optimization in conic form:

$$\begin{array}{lrl} P: & \mathrm{VAL}_* := \min_x & c^T x \\ & \text{s.t.} & Ax = b \\ & & x \in C \end{array}$$

$$\begin{array}{lrl} D: & \mathrm{VAL}^* := \max_{y,z} & b^T y \\ & \text{s.t.} & A^T y + z = c \\ & & z \in C^* \end{array}$$

$A \in \mathbb{R}^{m \times n}$

$C \subset X$ is a regular cone: closed, convex, pointed, with nonempty interior

$C^* := \{z : z^T x \geq 0 \text{ for all } x \in C\}$

Homogeneous Self-Dual (HSD) Model Embedding

Given initial values (x^0, y^0, z^0) satisfying $x^0 \in \text{int}C, z^0 \in \text{int}C^*$, $\tau^0 > 0, \kappa^0 > 0, \theta^0 > 0$, consider the homogeneous self-dual (HSD) embedding:

$$
\begin{array}{llrrrrrrcr}
H: & \text{VAL}_H := \min_{x,y,z,\tau,\kappa,\theta} & & & & \bar{\alpha}\theta & & & & \\
& \text{s.t.} & & Ax & -b\tau & +\bar{b}\theta & & & = & 0 \\
& & -A^T y & & +c\tau & +\bar{c}\theta & -z & & = & 0 \\
& & b^T y & -c^T x & & +\bar{g}\theta & & -\kappa & = & 0 \\
& & -\bar{b}^T y & -\bar{c}^T x & -\bar{g}\tau & & & & = & -\bar{\alpha} \\
& & & x \in C & \tau \geq 0 & & z \in C^* & \kappa \geq 0 & &
\end{array}
$$

where:

$$
\begin{array}{rclcrcl}
\bar{b} & = & \frac{b\tau^0 - Ax^0}{\theta^0} & \qquad & \bar{c} & = & \frac{A^T y^0 + z^0 - c\tau^0}{\theta^0} \\
\bar{g} & = & \frac{c^T x^0 - b^T y^0 + \kappa^0}{\theta^0} & \qquad & \bar{\alpha} & = & \frac{(z^0)^T x^0 + \tau^0 \kappa^0}{\theta^0}
\end{array}
$$

Properties of the HSD Model

- H is self-dual
- $(x, y, z, \tau, \kappa, \theta) = (x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ is a strictly feasible primal (and hence dual) solution of H
- $\mathrm{VAL}_H = 0$ and H attains its optimum
- Let $(x^*, y^*, z^*, \tau^*, \kappa^*, \theta^*)$ be any optimal solution of H. Then $(x^*)^T z^* = 0$ and $\tau^* \cdot \kappa^* = 0$, and
  - If $\tau^* > 0$, then x^*/τ^*, $(y^*/\tau^*, z^*/\tau^*)$ are optimal for P, D
  - If $\kappa^* > 0$, then either $c^T x^* < 0$ (and D is infeasible) or $-b^T y^* < 0$ (and P is infeasible)

Stopping Rule Theory for HSD Model

$(x, y, z, \tau, \kappa, \theta)$ is any (feasible) iterate of HSD model

Trial primal and dual values: $(\bar{x}, \bar{y}, \bar{z}) := (x/\tau, y/\tau, z/\tau)$

Compute residuals:
$$\begin{array}{rcl} r_p & := & b - A\bar{x} \\ r_d & := & A^T\bar{y} + \bar{z} - c \\ r_g & := & c^T\bar{x} - b^T\bar{y} + \bar{\kappa} \end{array}$$

Stopping rule: stop if

$$\text{RESID} := \frac{\|r_p\|}{\max\{1, \|b\|\}} + \frac{\|r_d\|}{\max\{1, \|c\|\}} + \frac{(r_g)^+}{\max\{1, \text{OPTVAL}\}} \le r_{\max}$$

Typically $r_{\max} = 10^{-8}\ ,\ 10^{-6}$

SDPT3 and SeDuMi stopping rules are similar to the above

Stopping Rule Theory for HSD Model, continued

Residuals:
$$\begin{aligned} r_p &:= b - A\bar{x} \\ r_d &:= A^T\bar{y} + \bar{z} - c \\ r_g &:= c^T\bar{x} - b^T\bar{y} + \bar{\kappa} \end{aligned}$$

Stopping rule: stop if

$$\text{RESID} := \frac{\|r_p\|}{\max\{1, \|b\|\}} + \frac{\|r_d\|}{\max\{1, \|c\|\}} + \frac{(r_g)^+}{\max\{1, \text{OPTVAL}\}} \le r_{\max}$$

Particular choice of norm and combination rule is not so important

What is important is that RESID be positively homogenous (of degree 1) in r_p, r_d, and $(r_g)^+$

Initial Residual and an Equivalence Lemma

$$\text{RESID}^0 = \left(\frac{\|b - Ax^0/\tau^0\|}{\max\{1, \|b\|\}} + \frac{\|A^T y^0/\tau^0 + z^0/\tau^0 - c\|_*}{\max\{1, \|c\|_*\}} + \frac{\left(c^T x^0/\tau^0 - b^T y^0/\tau^0 + \kappa^0/\tau^0\right)^+}{\max\{1, |\text{OPTVAL}|\}} \right)$$

Equivalence Lemma: Suppose that $(x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ is the starting point, and $(x, y, z, \pi, \tau, \kappa, \theta)$ is a feasible iterate of an interior-point method for solving H. Let $(\bar{x}, \bar{y}, \bar{z}) := (x/\tau, y/\tau, z/\tau)$ be the trial solution of P and D. Then the stopping rule is equivalent to:

$$\frac{\theta}{\theta^0 + \theta} \leq \frac{r_{\max}}{\text{RESID}^0} \left(\frac{(x^0)^T z^0 + \tau^0 \kappa^0}{(z^0)^T \bar{x} + (x^0)^T \bar{z} + \kappa^0 + \frac{\tau^0 \kappa}{\tau}} \right) \left(\frac{1}{\tau^0} \right)$$

Proof is just arithmetic.

An Iteration Count Identity

Stopping Rule: $\frac{\theta}{\theta^0+\theta} \le \frac{r_{\max}}{\text{RESID}^0}\left(\frac{(x^0)^T z^0+\tau^0\kappa^0}{(z^0)^T\bar{x}+(x^0)^T\bar{z}+\kappa^0+\frac{\tau^0\kappa}{\tau}}\right)\left(\frac{1}{\tau^0}\right)$

Let T denote the total number of iterations to solve H

Let β denote the average duality gap decrease over all iterations:

$$\beta := \sqrt[T]{\frac{2\bar{\alpha}\theta}{2\bar{\alpha}\theta^0}} = \sqrt[T]{\frac{\theta}{\theta^0}}$$

Corollary:

$$T = \left\lceil \frac{\ln\left(\frac{\theta^0}{\theta^0+\theta}\right)+\ln(\tau^0)+\ln\left(\frac{(z^0)^T\bar{x}+(x^0)^T\bar{z}+\kappa^0+\frac{\tau^0\kappa}{\tau}}{(z^0)^T x^0+\tau^0\kappa^0}\right)+\ln\left(\text{RESID}^0\right)+|\ln(r_{\max})|}{|\ln(\beta)|} \right\rceil$$

Simplified Iteration Count Identity

$$T = \left\lceil \frac{\ln\left(\frac{\theta^0}{\theta^0+\theta}\right) + \ln(\tau^0) + \ln\left(\frac{(z^0)^T\bar{x}+(x^0)^T\bar{z}+\kappa^0+\frac{\tau^0\kappa}{\tau}}{(z^0)^Tx^0+\tau^0\kappa^0}\right) + \ln\left(\mathsf{RESID}^0\right) + |\ln(r_{\max})|}{|\ln(\beta)|} \right\rceil$$

We can assume $\tau^0 = 1$ without loss of generality

Presume that at stopping: $\theta \approx 0$, $\kappa \approx 0$, $\tau > \varepsilon$ for some $\varepsilon > 0$, $\bar{x} \approx x^{\mathrm{opt}}$, $(\bar{y}, \bar{z}) \approx (y^{\mathrm{opt}}, z^{\mathrm{opt}})$ (optimal solutions of P/D)

Iteration count simplifies to:

$$T \approx \frac{\ln\left(\mathsf{RESID}^0\right) + |\ln(r_{\max})| + \ln\left(1 - \frac{(z^0-z^*)^T(x^0-x^*)}{(z^0)^Tx^0+\kappa^0}\right)}{|\ln(\beta)|}$$

Simplified Iteration Count, continued

$$T \approx \frac{\ln\left(\text{RESID}^0\right) + |\ln(r_{\max})| + \ln\left(1 - \frac{(z^0 - z^*)^T(x^0 - x^*)}{(z^0)^T x^0 + \kappa^0}\right)}{|\ln(\beta)|}$$

Iteration count depends only on inner products involving "cone variables" (and not on their norms)

$$Q := 1 - \frac{(z^0 - z^*)^T(x^0 - x^*)}{(z^0)^T x^0 + \kappa^0}$$

$$Q \leq 1 + \frac{\max\{1, \|b\|, \|c\|_*\}\left(\|x^* - x^0\| + \|y^* - y^0\|_*\right)}{(z^0)^T x^0 + \kappa^0} \text{RESID}^0 \ .$$

Substituting in the iteration count we can write:

Simplified Iteration Count, continued

$$T \lessapprox \frac{\ln\left(\text{RESID}^0\right) + |\ln(r_{\max})| + \ln\left(1 + \frac{\max\{1,\|b\|,\|c\|_*\}\left(\|x^*-x^0\|+\|y^*-y^0\|_*\right)}{(z^0)^T x^0+\kappa^0}\text{RESID}^0\right)}{|\ln(\beta)|}$$

IPM iterations of HSD model is driven by four quantities:

(1) $r_{\max}$ desired feasibility/optimality tolerance

(2) RESID^0 initial feasibility/optimality tolerance

(3) convergence rate β

(4) how close the starting values are to the optimal values

Note that if we use a theoretically guaranteed gap decrease $\beta = 1 - \frac{1}{8\vartheta_C}$ we obtain:

$$T \lessapprox 8\sqrt{\vartheta_C}\left(\ln\left(\text{RESID}^0\right) + |\ln(r_{\max})| + \ln\left(1 + \frac{\max\{1,\|b\|,\|c\|_*\}\left(\|x^*-x^0\|+\|y^*-y^0\|_*\right)}{(z^0)^T x^0+\kappa^0}\text{RESID}^0\right)\right)$$

A Strategy for Reducing IPM Iterations

Simplified Iteration Count:

$$T \lessapprox \frac{\ln\left(\text{RESID}^0\right) + |\ln(r_{\max})| + \ln\left(1 + \frac{\max\{1, \|b\|, \|c\|_*\}\left(\|x^* - x^0\| + \|y^* - y^0\|_*\right)}{(z^0)^T x^0 + \kappa^0}\text{RESID}^0\right)}{|\ln(\beta)|}$$

Try to replace $(x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ with $(x^1, y^1, z^1, \tau^1, \kappa^1, \theta^1)$ so that

$$\text{RESID}^1 << \text{RESID}^0 \ .$$

$$\text{RESID}^1 = \left(\frac{\|b - Ax^1/\tau^1\|}{\max\{1, \|b\|\}} + \frac{\|A^T y^1/\tau^1 + z^1/\tau^1 - c\|_*}{\max\{1, \|c\|_*\}} + \frac{\left(c^T x^1/\tau^1 - b^T y^1/\tau^1 + \kappa^1/\tau^1\right)^+}{\max\{1, |\text{OPTVAL}|\}}\right)$$

Our method is motivated by properties of skew-symmetric conic feasibility systems, for which the HSD model is a special case . . . .

Skew-Symmetric Feasibility Problem

Given a cone K and a skew-symmetric matrix M $(M^T = -M)$

Solve for (v, s):

$$
\begin{array}{rrrcl}
SFP: & Mv & +w & = & 0 \\
& v \in K \ , & w \in K^* & & .
\end{array}
$$

Properties of SFP:

- self-alternative
- always has a non-trivial solution
- is ill-posed (infinitesimal changes in RHS will render SFP infeasible)

Skew-Symmetric Feasibility Problem, continued

Solve for (v, s):

$$\begin{array}{rcccc} SFP: & Mv & +w & = & 0 \\ & v \in K \ , & w \in K^* & & . \end{array}$$

We can write homogenized P/D optimality conditions as an instance of SFP with assignments:

$$M = \begin{bmatrix} 0 & -A & +b \\ +A^T & 0 & -c \\ -b^T & c^T & 0 \end{bmatrix} \ , \quad v = \begin{pmatrix} y \\ x \\ \tau \end{pmatrix} \ , \ w = \begin{pmatrix} \pi \\ z \\ \kappa \end{pmatrix}$$

and

$$K = \Re^m \times C \times \Re_+ \ , \ K^* = \{0\} \times C^* \times \Re_+ \ .$$

Normalized Version of SFP

Given $(v^0, w^0) \in \mathrm{relint} K \times \mathrm{relint} K^*$, define the normalized version of SFP:

$$
\begin{array}{rrrcl}
NSFP: & Mv & +w & = & 0 \\
 & (w^0)^T v & +(v^0)^T w & = & 1 \\
 & v \in K \ , & w \in K^* & & .
\end{array}
$$

Normalized Version of SFP and Image Set $\mathcal{H}$

$$\begin{array}{rcccl} NSFP: & Mv & +w & = & 0 \\ & (w^0)^T v & +(v^0)^T w & = & 1 \\ & v \in K \ , & w \in K^* & & . \end{array}$$

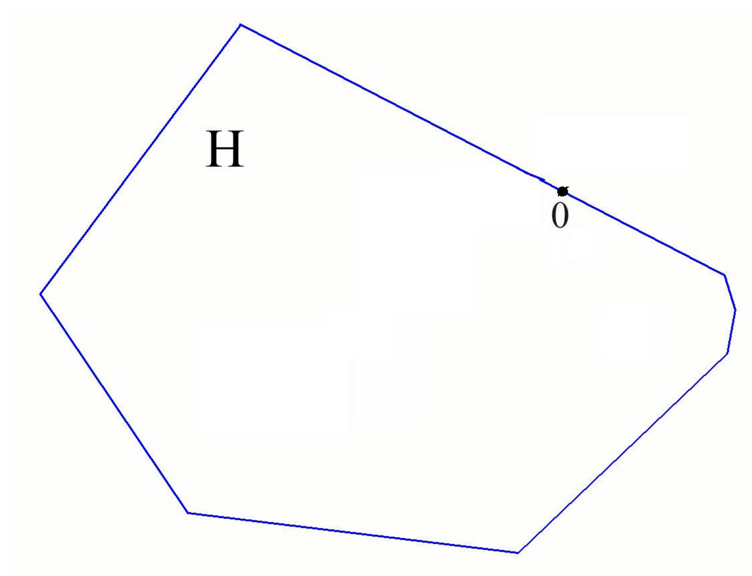
Image Set: $\mathcal{H} := \{-Mv - w : (w^0)^T v + (v^0)^T w = 1, \ v \in K, \ w \in K^*\}$

The *polar* of a convex set S is $S^\circ := \{y : y^T x \leq 1 \text{ for all } x \in S\}$

Properties of Image Set and its Polar:

- $\mathcal{H}^\circ = \{v : w^0 + M^T v \in K^*, \ v^0 + v \in K\}$
- $\text{rec}\mathcal{H}^\circ = \{v : \exists w \text{ satisfying } Mv + w = 0, \ v \in K, \ w \in K^*\}$

Image Set, Illustrated



Solving SFP via Optimization

Given initial values (v^0, w^0) satisfying $v^0 \in \mathrm{relint} K, w^0 \in \mathrm{relint} K^*$, and $\theta^0 > 0$, consider:

$$
\begin{array}{lllllll}
OSHF: & \mathrm{VAL}_S := \min_{v,w,\theta} & & \frac{(v^0)^T w^0}{\theta^0}\theta & & & \\
& \text{s.t.} & -Mv & + \left(\frac{Mv^0+w^0}{\theta^0}\right)\theta & -w & = & 0 \\
& & -\left(\frac{Mv^0+w^0}{\theta^0}\right)^T v & & & = & -\frac{(v^0)^T w^0}{\theta^0} \\
& & v \in K\ , & & w \in K^*\ , & &
\end{array}
$$

Properties of $OSHF$:

- translates exactly to HSD model when M arises from conic primal/dual problems

- $OSHF$ is self-dual, $(v, w, \theta) = (v^0, w^0, \theta^0)$ is a strictly feasible primal (and hence dual) solution of $OSHF$, $\mathrm{VAL}_S = 0$, and $OSHF$ attains its optimum

Homogeneous Self-Dual (HSD) Model Embedding

Given initial values (x^0, y^0, z^0) satisfying $x^0 \in \mathrm{int} C, z^0 \in \mathrm{int} C^*$, $\tau^0 > 0, \kappa^0 > 0, \theta^0 > 0$, consider the homogeneous self-dual (HSD) embedding:

$$
\begin{array}{lllrrrrrrcr}
H: & \mathrm{VAL}_H := \min_{x,y,z,\tau,\kappa,\theta} & & & & & \bar{\alpha}\theta & & & & \\
 & \text{s.t.} & & & Ax & -b\tau & +\bar{b}\theta & & & = & 0 \\
 & & -A^T y & & & +c\tau & +\bar{c}\theta & -z & & = & 0 \\
 & & b^T y & & -c^T x & & +\bar{g}\theta & & -\kappa & = & 0 \\
 & & -\bar{b}^T y & & -\bar{c}^T x & -\bar{g}\tau & & & & = & -\bar{\alpha} \\
 & & & & x \in C & \tau \geq 0 & & z \in C^* & \kappa \geq 0 & &
\end{array}
$$

where:

$$
\begin{array}{rclcrcl}
\bar{b} & = & \frac{b\tau^0 - Ax^0}{\theta^0} & \quad & \bar{c} & = & \frac{A^T y^0 + z^0 - c\tau^0}{\theta^0} \\
\bar{g} & = & \frac{c^T x^0 - b^T y^0 + \kappa^0}{\theta^0} & \quad & \bar{\alpha} & = & \frac{(z^0)^T x^0 + \tau^0 \kappa^0}{\theta^0}
\end{array}
$$

Strategy for Reducing IPM Iterations, Recalled

Simplified Iteration Count:

$$T \lessapprox \frac{\ln\left(\text{RESID}^0\right) + |\ln(r_{\max})| + \ln\left(1 + \frac{\max\{1,\|b\|,\|c\|_*\}\left(\|x^*-x^0\|+\|y^*-y^0\|_*\right)}{(z^0)^T x^0+\kappa^0}\text{RESID}^0\right)}{|\ln(\beta)|}$$

Try to replace $(x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ with $(x^1, y^1, z^1, \tau^1, \kappa^1, \theta^1)$ so that

$$\text{RESID}^1 << \text{RESID}^0$$

$$\text{RESID}^1 = \left(\frac{\|b - Ax^1/\tau^1\|}{\max\{1, \|b\|\}} + \frac{\|A^T y^1/\tau^1 + z^1/\tau^1 - c\|_*}{\max\{1, \|c\|_*\}} + \frac{\left(c^T x^1/\tau^1 - b^T y^1/\tau^1 + \kappa^1/\tau^1\right)^+}{\max\{1, |\text{OPTVAL}|\}}\right)$$

Reducing RESID^0

The polar image set $\mathcal{H}^\circ$ corresponding to HSD model works out to be the feasible region of:

$$
\begin{array}{llrcl}
AUX: \quad +\infty = \max_{\check{y},\check{x},\check{\tau}} & & \tau^0 + \check{\tau} & & \\
& \text{s.t.} & A\check{x} - b\check{\tau} & = & 0 \\
& & z^0 - A^T\check{y} + c\check{\tau} & \in & C^* \\
& & \kappa^0 + b^T\check{y} - c^T\check{x} & \geq & 0 \\
& & x^0 + \check{x} & \in & C \\
& & \tau^0 + \check{\tau} & \geq & 0
\end{array}
$$

This optimization problem seeks to shoot out on a ray of $\mathcal{H}^\circ$, corresponding to an optimal solution of the original P/D conic system

Reducing RESID^0, continued

$$
\begin{array}{lllrcl}
AUX: & {\color{red}+\infty} & = \max_{\check{y},\check{x},\check{\tau}} & \tau^0 + \check{\tau} & & \\
 & & \text{s.t.} & A\check{x} - b\check{\tau} & = & 0 \\
 & & & z^0 - A^T\check{y} + c\check{\tau} & \in & C^* \\
 & & & \kappa^0 + b^T\check{y} - c^T\check{x} & \geq & 0 \\
 & & & x^0 + \check{x} & \in & C \\
 & & & \tau^0 + \check{\tau} & \geq & 0
\end{array}
$$

- $(\check{y}, \check{x}, \check{\tau}) := (0, 0, 0)$ is a strictly feasible solution of AUX, and
- The rays of the feasible region of AUX with strictly improving objective value are of the form $(y, x, \tau) = (y^*, x^*, 1)$ where x^*, y^* are optimal solutions of P and D.

Reducing RESID^0, continued

$$
\begin{array}{llrcl}
AUX: & +\infty = \max_{\breve{y},\breve{x},\breve{\tau}} & \tau^0 + \breve{\tau} & & \\
& \text{s.t.} & A\breve{x} - b\breve{\tau} & = & 0 \\
& & z^0 - A^T\breve{y} + c\breve{\tau} & \in & C^* \\
& & \kappa^0 + b^T\breve{y} - c^T\breve{x} & \geq & 0 \\
& & x^0 + \breve{x} & \in & C \\
& & \tau^0 + \breve{\tau} & \geq & 0
\end{array}
$$

- If $(\breve{y}, \breve{x}, \breve{\tau})$ is feasible for AUX, then under the assignment:

$$
\begin{pmatrix} x^1 \\ y^1 \\ z^1 \\ \tau^1 \\ \kappa^1 \\ \theta^1 \end{pmatrix} = \begin{pmatrix} (x^0 + \breve{x})/(\tau^0 + \breve{\tau}) \\ (y^0 + \breve{y})/(\tau^0 + \breve{\tau}) \\ (z^0 - A^T\breve{y} + c\breve{\tau})/(\tau^0 + \breve{\tau}) \\ 1 \\ (\kappa^0 + b^T\breve{y} - c^T\breve{x})/(\tau^0 + \breve{\tau}) \\ 1 \end{pmatrix},
$$

we have

$$
\text{RESID}^1 = \left(\frac{\tau^0}{\tau^0 + \breve{\tau}}\right) \text{RESID}^0
$$

Strategy for Reducing RESID

Starting at $(\check{y}, \check{x}, \check{\tau}) = (0, 0, 0)$, perform a random walk on the feasible region of AUX to improve the objective function to some (pre-set) goal value U.

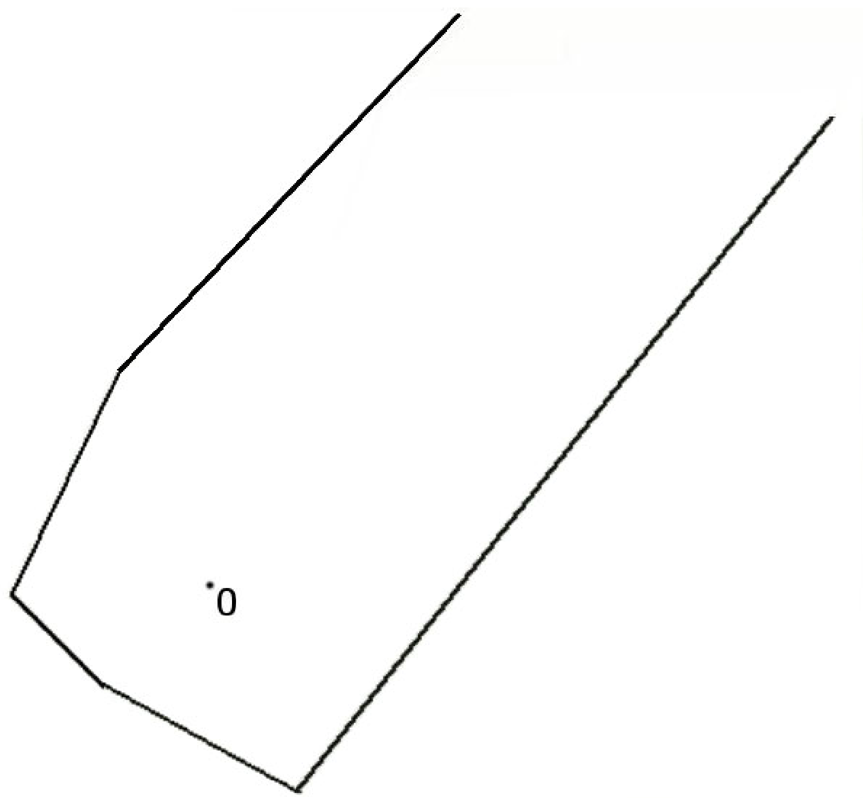
Output final value of $(\check{y}, \check{x}, \check{\tau})$ and set

$$\begin{pmatrix} x^1 \\ y^1 \\ z^1 \\ \tau^1 \\ \kappa^1 \\ \theta^1 \end{pmatrix} = \begin{pmatrix} (x^0 + \check{x})/(\tau^0 + \check{\tau}) \\ (y^0 + \check{y})/(\tau^0 + \check{\tau}) \\ (z^0 - A^T\check{y} + c\check{\tau})/(\tau^0 + \check{\tau}) \\ 1 \\ (\kappa^0 + b^T\check{y} - c^T\check{x})/(\tau^0 + \check{\tau}) \\ 1 \end{pmatrix} .$$

Use $(x^1, y^1, z^1, \tau^1, \kappa^1, \theta^1)$ as the new given initial values in the HSD model

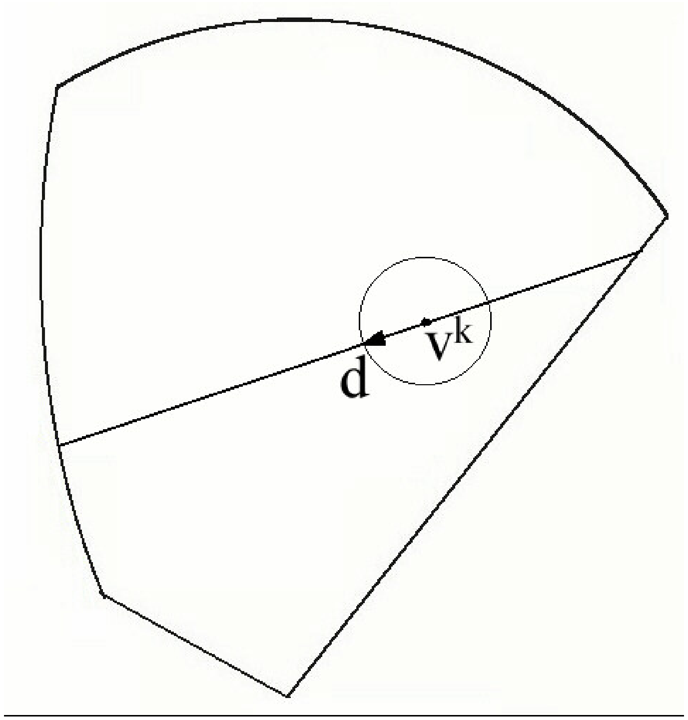
Distribution on a Convex Set

$X \sim f(\cdot)$ on S



The Hit-and-Run Algorithm

$X \sim f(\cdot)$ on S



The Hit-and-Run Algorithm

$X \sim f(\cdot)$ on S

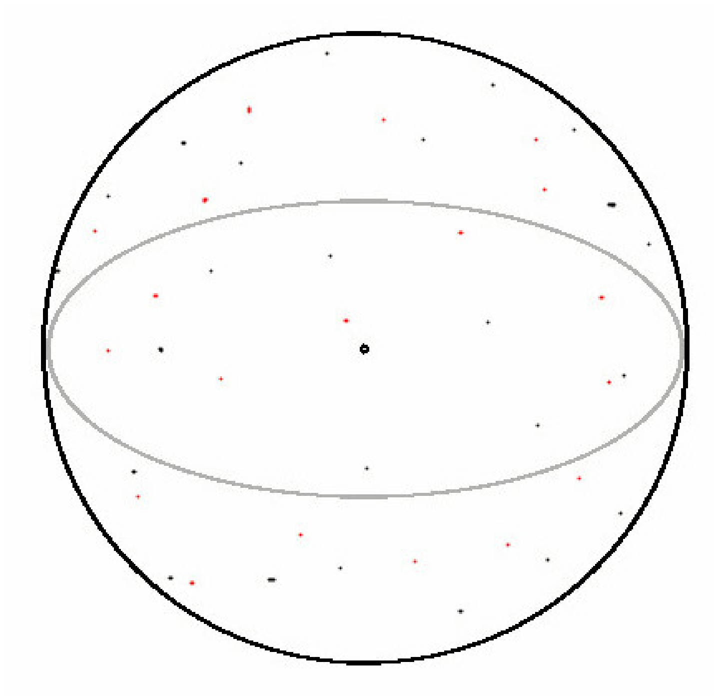
Let $v^0 \in \text{int}S$ be given

v^k is current point in hit-and-run algorithm

- choose $d \sim U(S^{n-1})$, the $(n-1)$-sphere in $\Re^n$
- v^{k+1} is chosen according to the marginal distribution of $f(\cdot)$ on

$$S \cap \{v^k + \alpha d : \alpha \in \Re\}$$

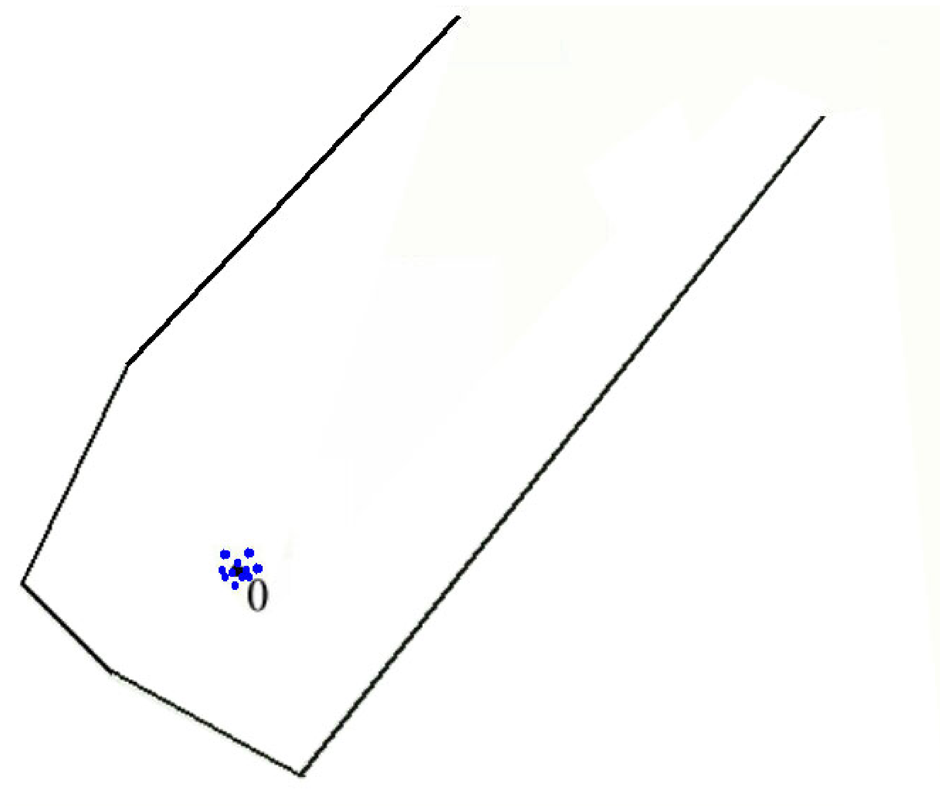
Uniform Vector on the Sphere



Random Sampling when S is Unbounded

$f(x) \sim e^{-t\|x\|}$

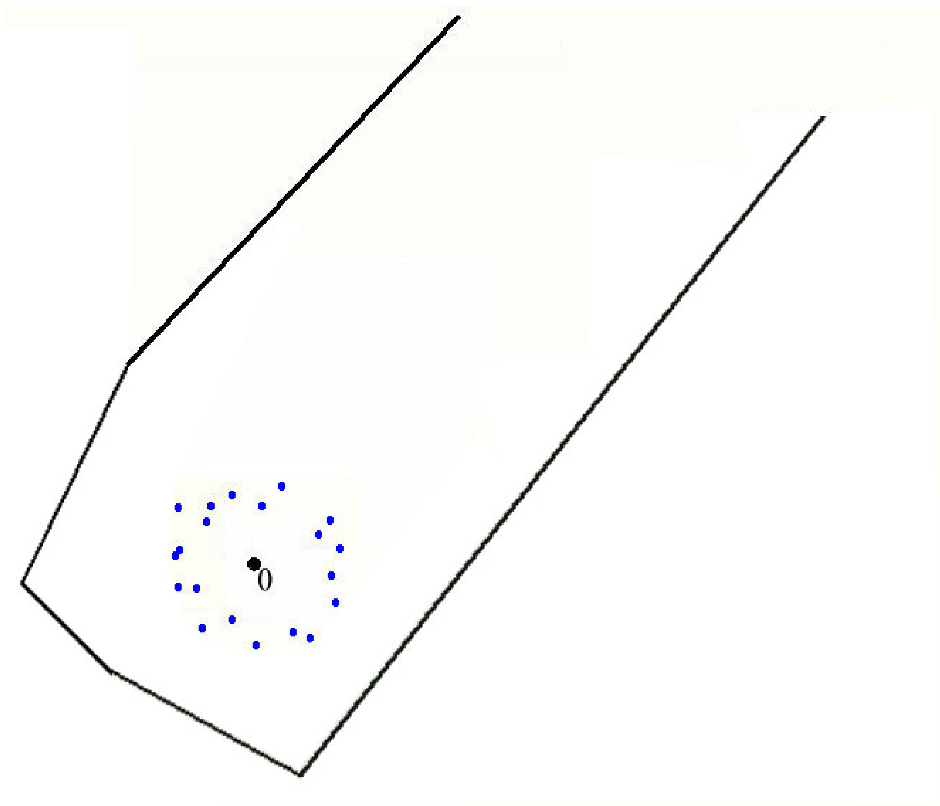
$t >> 0$



Random Sampling when S is Unbounded

$f(x) \sim e^{-t\|x\|}$

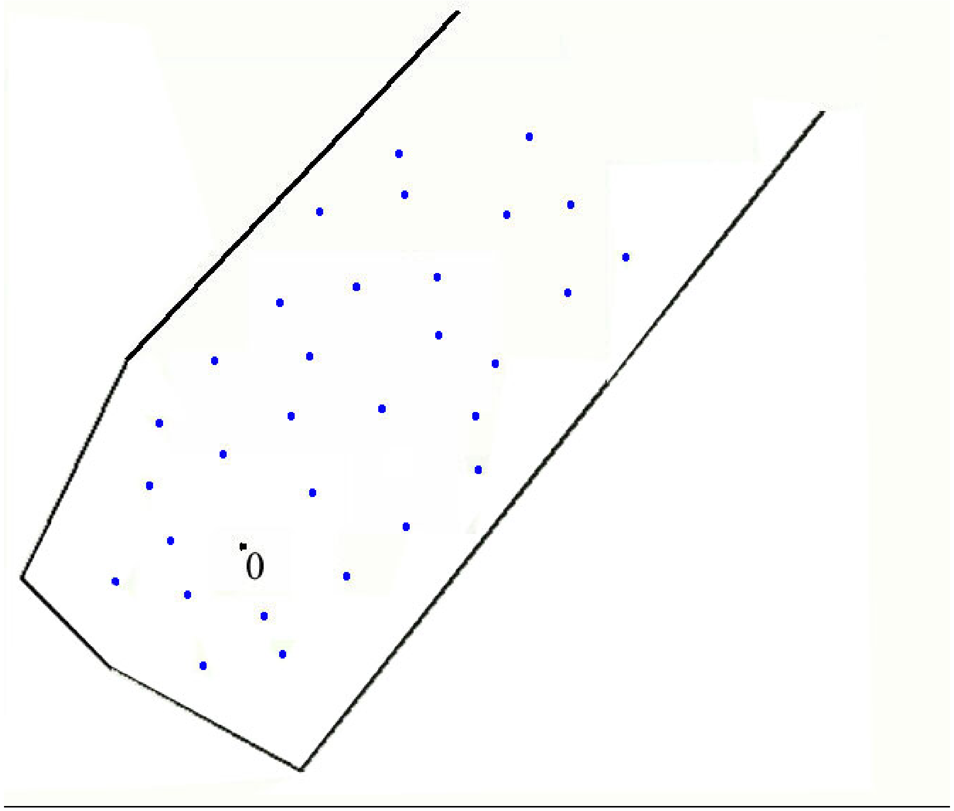
$t > 0$



Random Sampling when S is Unbounded

$f(x) \sim e^{-t\|x\|}$

$t \approx 0$



Previous Computational Experience

Computations on 50 “random” problems for each m, n pairing

Dimensions		SDPT3-HSD IPM Iterations	
m	n	Original	After Re-Initialization
20	100	19.42	18.78
100	500	21.30	18.16
200	1000	19.74	16.14
100	5000	32.86	20.00
200	5000	31.78	18.18

Problems were pre-designed to hopefully be poorly conditioned

Used SDPT3-HSD software

Numbers in table are arithmetic averages

Previous Computational Experience, continued

Dimensions m	n	SDPT3-HSD Running Time (seconds) Original	After Re-Initialization	Random Walk Running Time (seconds)
20	100	0.99	0.98	0.90
100	500	3.62	3.08	4.38
200	1000	14.89	12.20	14.55
100	5000	44.43	27.66	308.84
200	5000	125.21	73.28	319.19

Could the random walk methodology be made significantly more efficient?

Current Computational Experience

More efficient random walk methodology

20 randomly generated dense problems of larger sizes for each m, n pairing

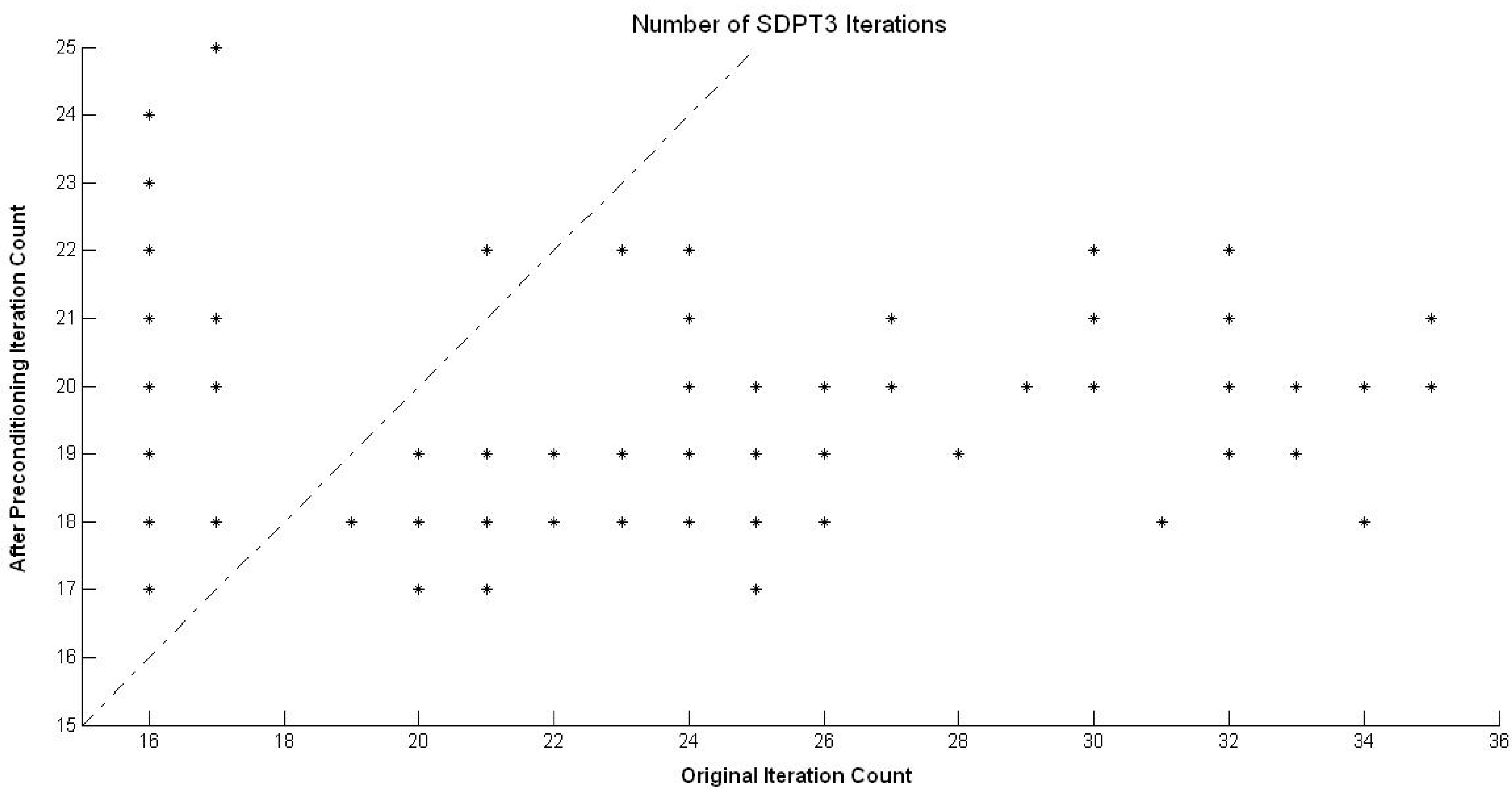
100 problems in all

Problems are mix of easy and difficult problems (for IPMs)

		AVERAGE RUNNING TIMES (seconds)		
Dimensions		Single	Random Walk	Ratio
m	n	IPM Iteration	(250 steps)	IPM Iteration/Random Walk
200	1000	.45	.36	.78
200	2000	.95	.62	.65
200	5000	2.47	1.46	.59
500	5000	12.99	2.91	.22
1000	10000	91.60	10.62	.12

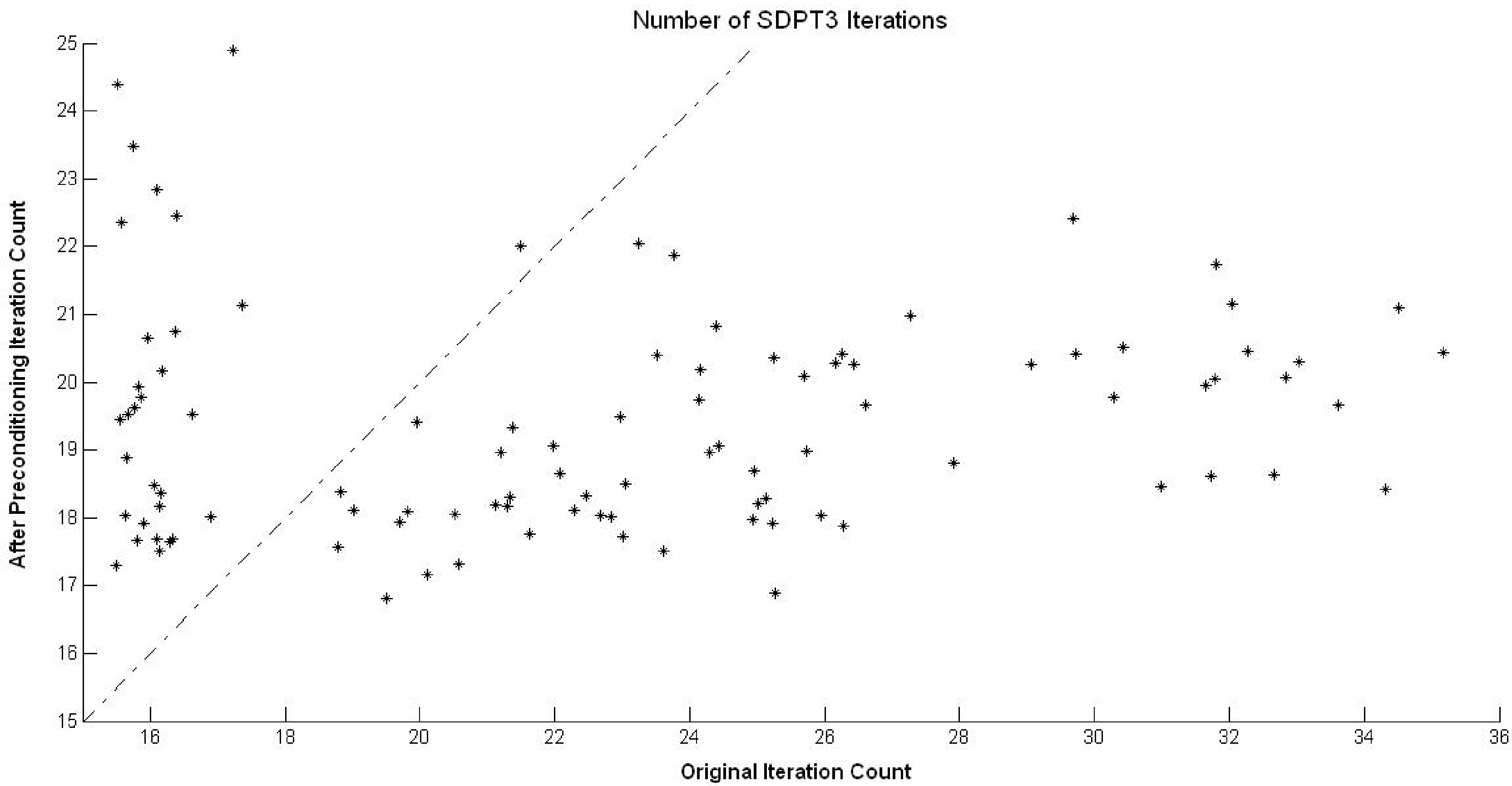
Numbers in table are arithmetic averages

Current Computational Experience Graphic



100 instances, but only 55 dots in figure

Current Computational Experience Graphic, continued



ε-perturbation of all dots, now we see all 100 instance

Iterations and Running Times

Original IPM Iteration Range	Percentage of Instances	Average IPM Iterations		Average Total Running Time		Net Savings (%)
		Original	After Random Walk	Original	After Random Walk	
16-18	29%	16.1	19.8	916.6	1167.8	-27%
19-21	16%	20.3	18.3	9.9	9.3	5%
22-24	19%	23.2	19.2	497.2	442.5	11%
25-27	16%	25.7	19.0	102.5	76.6	25%
28-30	6%	29.5	20.3	65.6	47.5	28%
31-35	14%	32.9	19.9	81.1	51.5	36%

Averages in table are arithmetic averages

Re-Initialization and Numerical Conditioning

Re-initialization yields better-conditioning for Sherman-Morrison-Woodbury (SMW) updates in our 100 dense instances:

Number of Instances with Ill-Conditioning in SMW Update

Original	After Re-Initialization
45/100	0/100

- When SMW is ill-conditioned, SDPT3 uses LU factorization on a related system
- When the instances are dense, the cost savings is minimal
- When the instances are structured and sparse, the cost savings has the potential to be large

Conclusions and Caveats

Conclusions:

- Re-Initialization is effective on problems that otherwise take at least 19 IPM iterations
- Effectiveness grows with original IPM iterations
- Re-Initialization decreases the variance in IPM iterations

Caveats:

- Our dense problem instances were not as random as we would like
- Larger instances were better behaved on average, in spite of our design efforts

Next Steps

- Testing on NETLIB, SDPLIB, other large collections of problems that are more representative of application environment

- Complexity theory of the random walk to approach a ray (we have some results on this already)

- Heuristics to further improve random walk efficiency

Back-up Slides to Follow

Homogenized Conic Optimization as SSFP

Solve for (x, y, z, τ, κ):

$$
\begin{array}{rrrrrrcl}
HCOP: & & Ax & -b\tau & & & = & 0 \\
& -A^T y & & +c\tau & -z & & = & 0 \\
& b^T y & -c^T x & & & -\kappa & = & 0 \\
& y \in \Re^m & x \in C & \tau > 0 & z \in C^* & \kappa \geq 0 & &
\end{array}
$$