PageRank by Basis Pursuit

ICIAM 07

Zürich, Switzerland

July 16–20, 2007

Sou-Cheng Choi and Michael Saunders

iCME
Stanford University
Stanford, CA 94305-4026
scchoi@stanford.edu

Dept of Management Sci & Eng
Stanford University
Stanford, CA 94305-4026
saunders@stanford.edu

Abstract (slightly revised)

Many imaging and compressed sensing applications seek **sparse solutions to under-determined least-squares problems**:

$$Ax \approx b, \qquad x \text{ } \textbf{sparse}.$$

Abstract (slightly revised)

Many imaging and compressed sensing applications seek **sparse solutions to under-determined least-squares problems**:

$$Ax \approx b, \qquad x \text{ **sparse**}.$$

Basis Pursuit Denoising (BPDN) achieves sparsity by bounding the 1-norm of the solution (and there are now several computational algorithms):

$$\min_x \tfrac{1}{2}\|b - Ax\|^2 \; + \; \lambda\|x\|_1.$$

Abstract (slightly revised)

Many imaging and compressed sensing applications seek **sparse solutions to under-determined least-squares problems**:

$$Ax \approx b, \qquad x \textbf{ sparse}.$$

Basis Pursuit Denoising (BPDN) achieves sparsity by bounding the 1-norm of the solution (and there are now several computational algorithms):

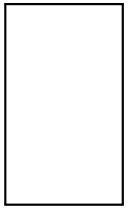
$$\min_x \tfrac{1}{2}\|b - Ax\|^2 \;+\; \lambda\|x\|_1.$$

Personalized PageRank eigenvectors are **sparse** in the sense that most elements are **exactly zero**, and some are very small.
BPDN **may be** a reasonable tool for finding the tiny proportion of significant nonzeros.

Sparse x

Lasso(ν)   Tibshirani 1996

$$\min_x \ \tfrac{1}{2}\|b - Ax\|^2 \quad \text{st} \quad \|x\|_1 \le \nu$$

$A =$ 

explicit

Sparse x

Lasso(ν)   Tibshirani 1996

$$\min_x \ \tfrac{1}{2}\|b - Ax\|^2 \quad \text{st} \quad \|x\|_1 \le \nu$$

$A =$ [ ]
explicit

Basis Pursuit   Chen, Donoho & Saunders 1998

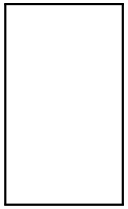
$$\min_x \ \|x\|_1 \quad \text{st} \quad Ax = b$$

$A =$ [ ]
fast operator

Sparse x

Lasso(ν)   Tibshirani 1996

$$\min_x \ \tfrac{1}{2}\|b - Ax\|^2 \quad \text{st} \quad \|x\|_1 \le \nu$$

$A =$ 

explicit

BPDN(λ)   Chen, Donoho & S 1998

$$\min_x \ \tfrac{1}{2}\|b - Ax\|^2 \ + \ \lambda\|x\|_1$$

$A =$ 

fast operator

x is likely to be **sparse**

($\|x\|_1$ is closest convex approximation to $\|x\|_0$)

BP and BPDN Algorithms

OMP	Davis, Mallat et al 1997	Greedy
BPDN-interior	Chen, Donoho & S, 1998, 2001	Interior, CG
PDSCO, PDCO	S 1997, 2002	Interior, LSQR
BCR	Sardy, Bruce & Tseng 2000	Orthogonal blocks
Homotopy	Osborne et al 2000	Active-set, all λ
LARS	Efron, Hastie et al 2004	Active-set, all λ
STOMP	Donoho, Tsaig et al 2006	Double greedy
l1_ls	Kim, Koh et al 2007	Primal barrier, PCG
GPSR	Figueiredo, Nowak & Wright 2007	Gradient-projection

Basis Pursuit Denoising (BPDN)

Pure LS

$$\min_{x,r} \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

Basis Pursuit Denoising (BPDN)

Pure LS

$$\min_{x,r} \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

If x not unique, need regularization

Basis Pursuit Denoising (BPDN)

Pure LS

$$\min_{x,r} \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

If x not unique, need regularization

Tikhonov(λ)

$$\min_{x,r} \lambda\|x\|^2 \; + \; \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

Basis Pursuit Denoising (BPDN)

Pure LS

$$\min_{x,r} \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

If x not unique, need regularization

Tikhonov(λ)

$$\min_{x,r} \lambda\|x\|^2 + \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

BPDN(λ)

$$\min_{x,r} \lambda\|x\|_1 + \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

Basis Pursuit Denoising (BPDN)

Pure LS

$$\min_{x,r} \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

If x not unique, need regularization

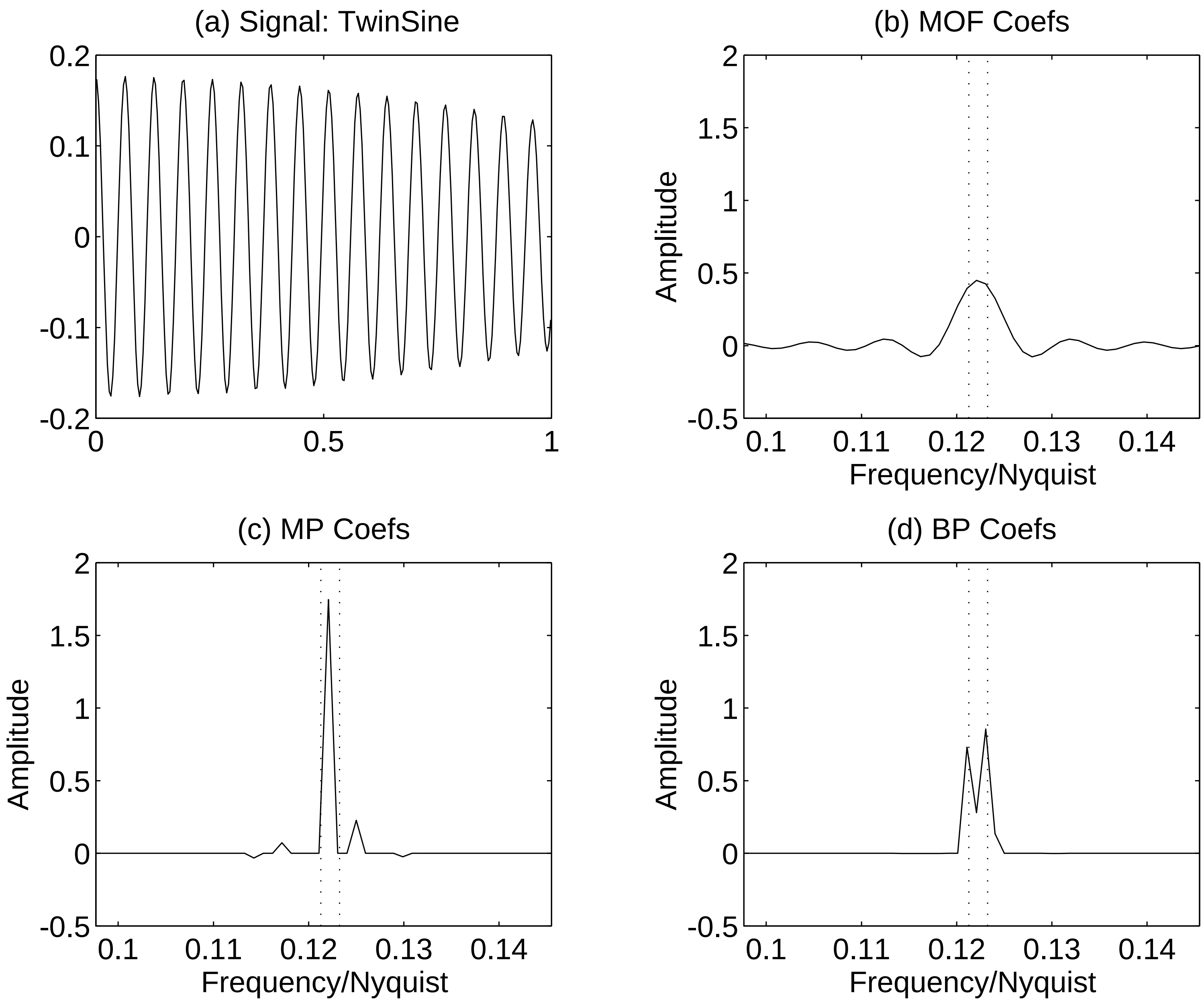
Tikhonov(λ)

$$\min_{x,r} \lambda\|x\|^2 \ + \ \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

BPDN(λ)

$$\min_{x,r} \lambda\|x\|_1 \ + \ \tfrac{1}{2}\|r\|^2 \qquad \text{st} \qquad r = b - Ax$$

Smaller $\|x\|$, bigger $\|r\|$


(a) Signal: TwinSine
(b) MOF Coefs
(c) MP Coefs
(d) BP Coefs
Amplitude
Frequency/Nyquist
0.2
0.1
0
-0.1
-0.2
0.5
1
2
1.5
-0.5
0.11
0.12
0.13
0.14

Let $r = \lambda y$

BPprimal(λ) $\quad \min_{x,y} \; \|x\|_1 + \frac{1}{2}\lambda\|y\|^2 \quad \text{st} \quad Ax + \lambda y = b$

BPdual(λ) $\quad \min_{y} \; -b^T y + \frac{1}{2}\lambda\|y\|^2 \quad \text{st} \quad \|A^T y\|_\infty \le 1$

Let $r = \lambda y$

BPprimal(λ) $\quad \min_{x,y} \; \|x\|_1 + \frac{1}{2}\lambda\|y\|^2 \quad \text{st} \quad Ax + \lambda y = b$

BPdual(λ) $\quad \min_{y} \; -b^T y + \frac{1}{2}\lambda\|y\|^2 \quad \text{st} \quad \|A^T y\|_\infty \le 1$

Suggests regularized LP problems:

LPprimal(λ) $\quad \min_{x,y} \; c^T x + \frac{1}{2}\lambda\|y\|^2 \quad \text{st} \quad Ax + \lambda y = b, \;\; x \ge 0$

LPdual(λ) $\quad \min_{y} \; -b^T y + \frac{1}{2}\lambda\|y\|^2 \quad \text{st} \quad A^T y \le c$

Friedlander and S 2007

New active-set MATLAB codes for each problem

PageRank

Langville and Meyer 2006

$$H = \begin{pmatrix} H_{11} & H_{12} \\ 0 & 0 \end{pmatrix}$$ Hyperlink matrix $n \times n$

$$S = \begin{pmatrix} H_{11} & H_{12} \\ \frac{1}{n}ee^T & \frac{1}{n}ee^T \end{pmatrix}$$ Stochastic matrix $Se = e$

$$G = \alpha S + (1-\alpha)ev^T$$ Google matrix $Ge = e$

$\alpha = 0.99$ say

$v = e_i$ say Personalization vector

eigenvector	$\equiv$	linear system
$G^T x = x$	$\equiv$	$(I - \alpha H_{11}^T)x_1 = v$

Sparse PageRank

Regard $(I - \alpha H_{11}^T)x_1 = v$ as $Ax = v$

x can be sparse if v is sparse

Apply active-set solver to

$$\min_{x,y}\ e^T x + \tfrac{1}{2}\lambda\|y\|^2 \quad \text{st} \quad Ax + \lambda y = v, \quad x \geq 0$$

LPprimal seems to proceed in a **greedy way**:

Starts with $x = 0$, chooses one x_j at a time

s nonzero $x_j \quad \Rightarrow \quad s$ iterations

LPprimal solver

$$\min_{x,y} \; e^T x + \tfrac{1}{2}\lambda\|y\|^2 \quad \text{st} \quad Ax + \lambda y = v, \quad x \geq 0$$

$S =$ columns of A corresponding to $x_j > 0$

Main work per itn:

Form $z = A^T y$
Solve $S^T S\, dx = c_S - S^T y$
Form $dy = S\, dx$

Numerical results

$$H = 9914 \times 9914 \text{ Stanford CS web matrix}$$

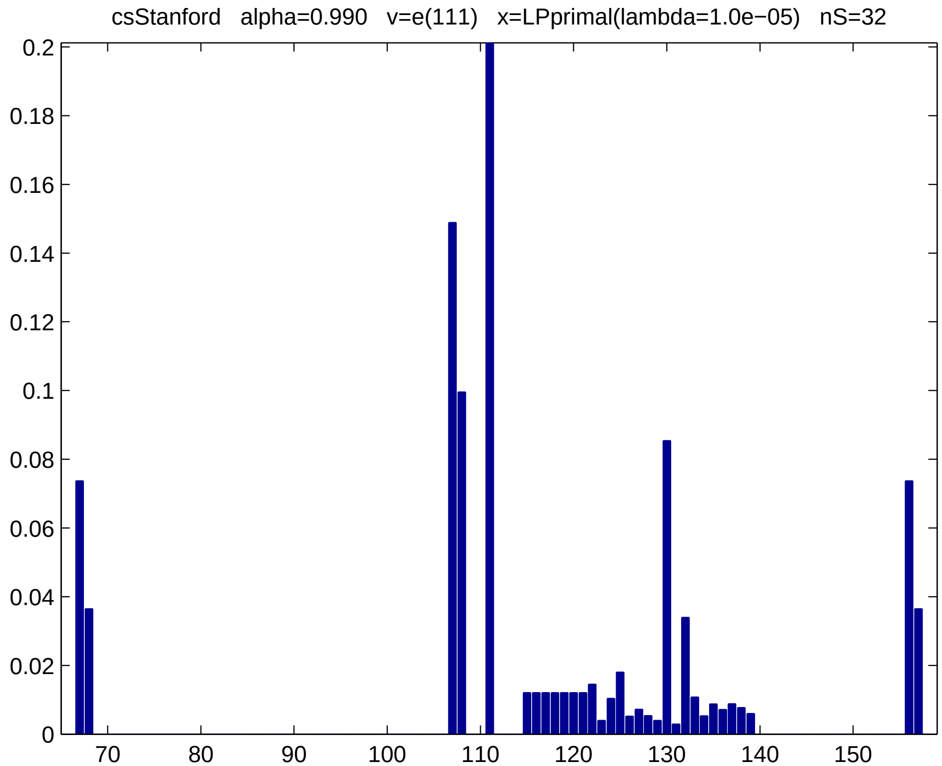
$$H_{11} = 7053 \times 7053$$

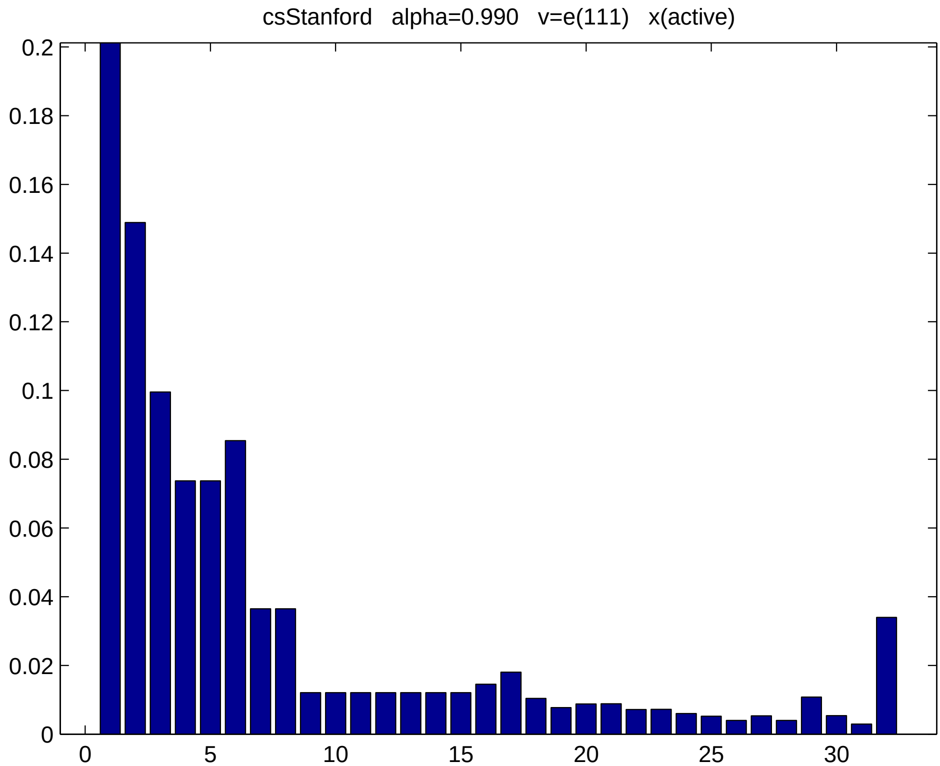
$$\alpha = 0.99$$

$$v = \text{various } e_i$$

Plot nonzero x_j and the order in which they came in

csStanford $v = e_{111}$

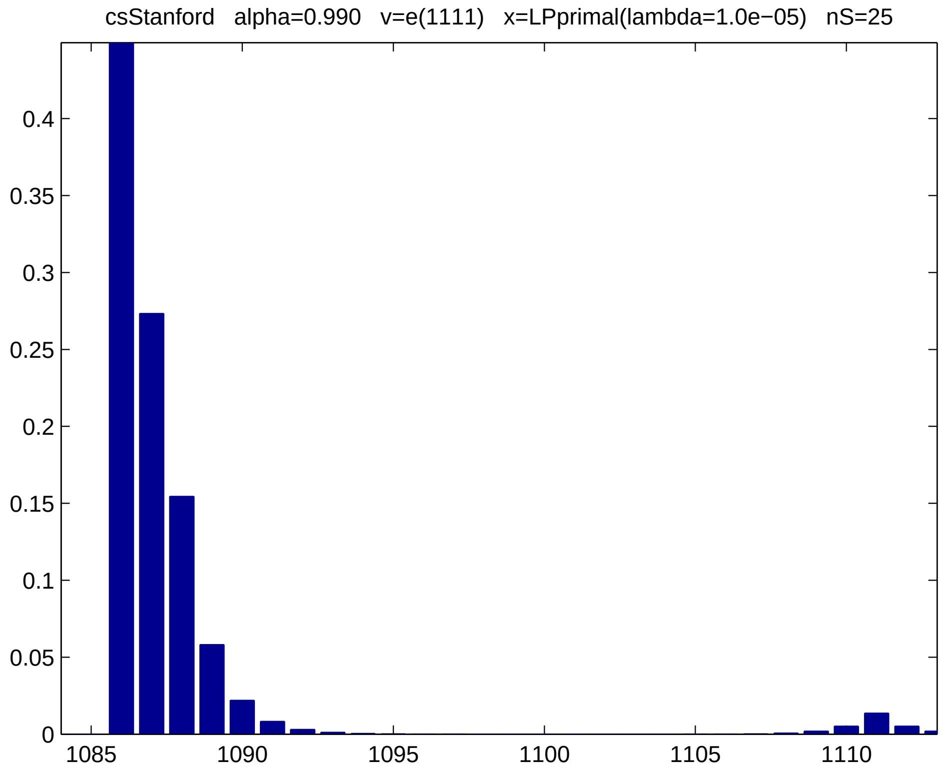


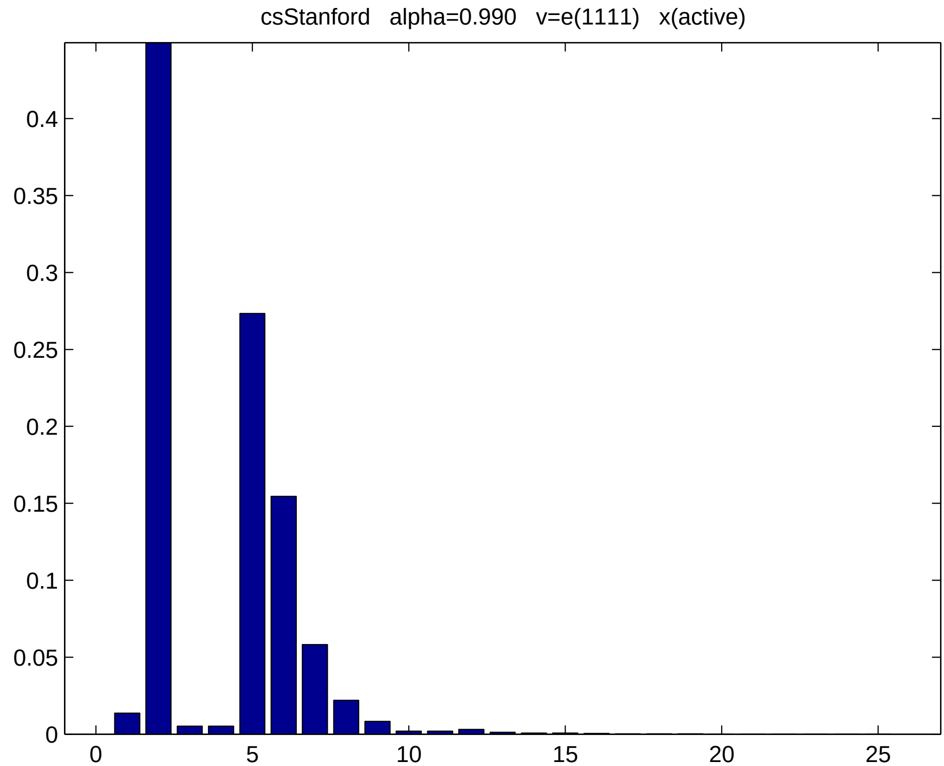


nonzero x_j

chosen in this order

csStanford $v = e_{1111}$

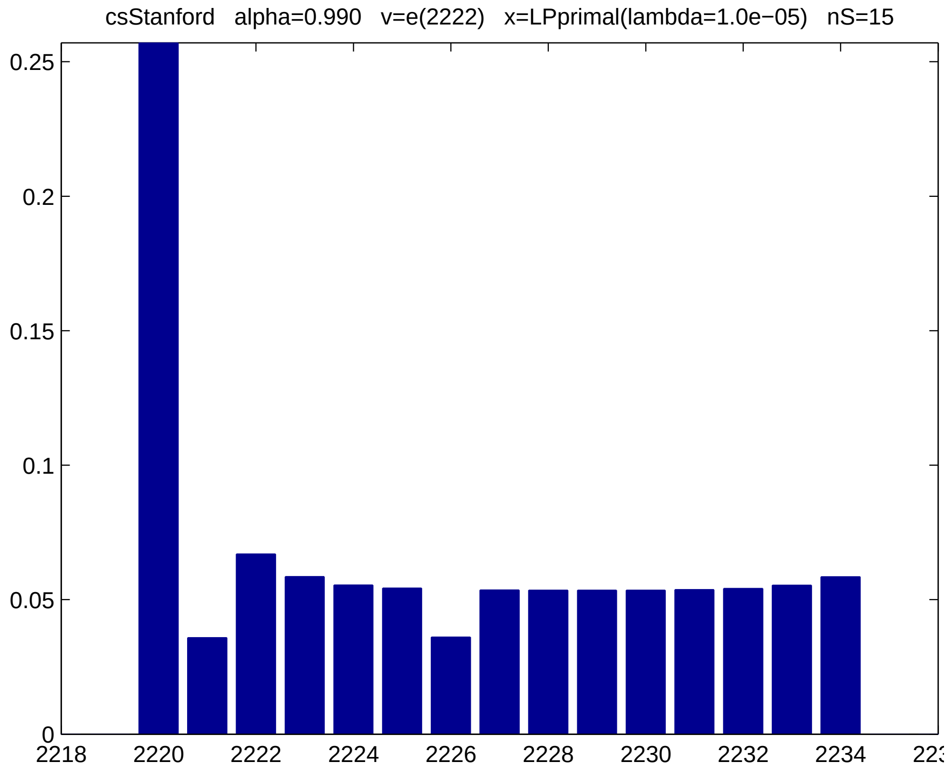


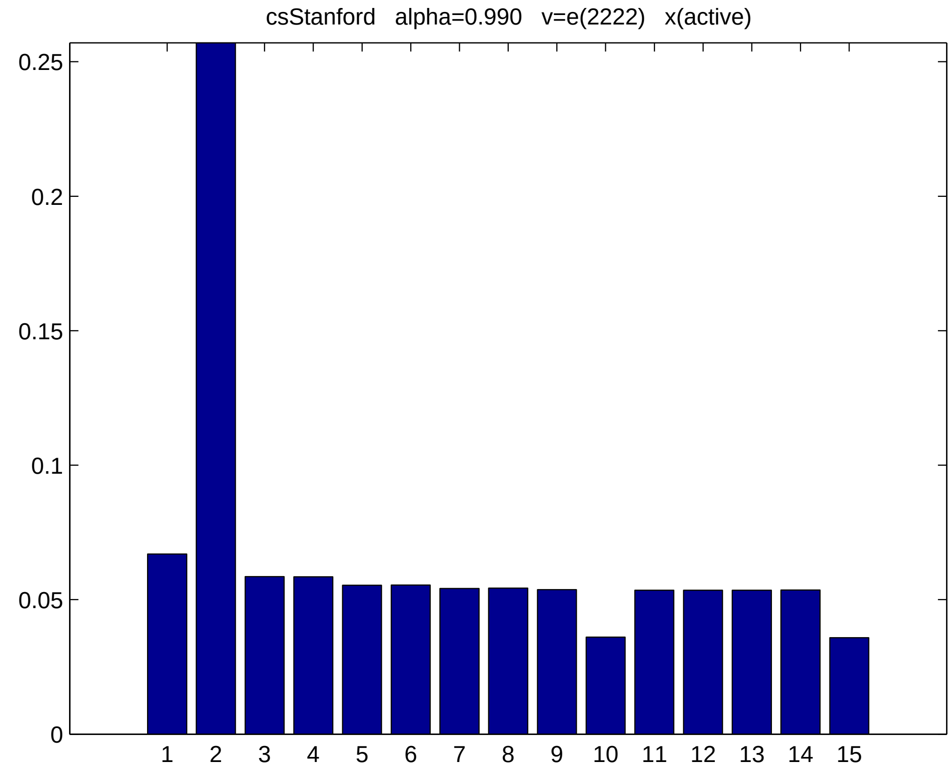


nonzero x_j

chosen in this order

csStanford $v = e_{2222}$

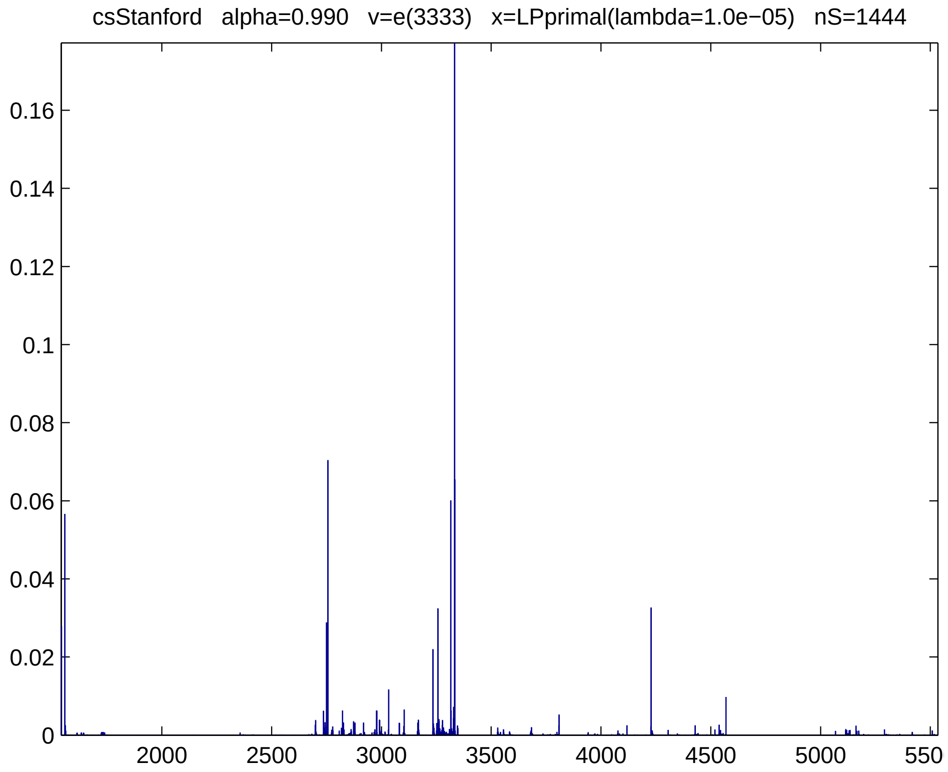


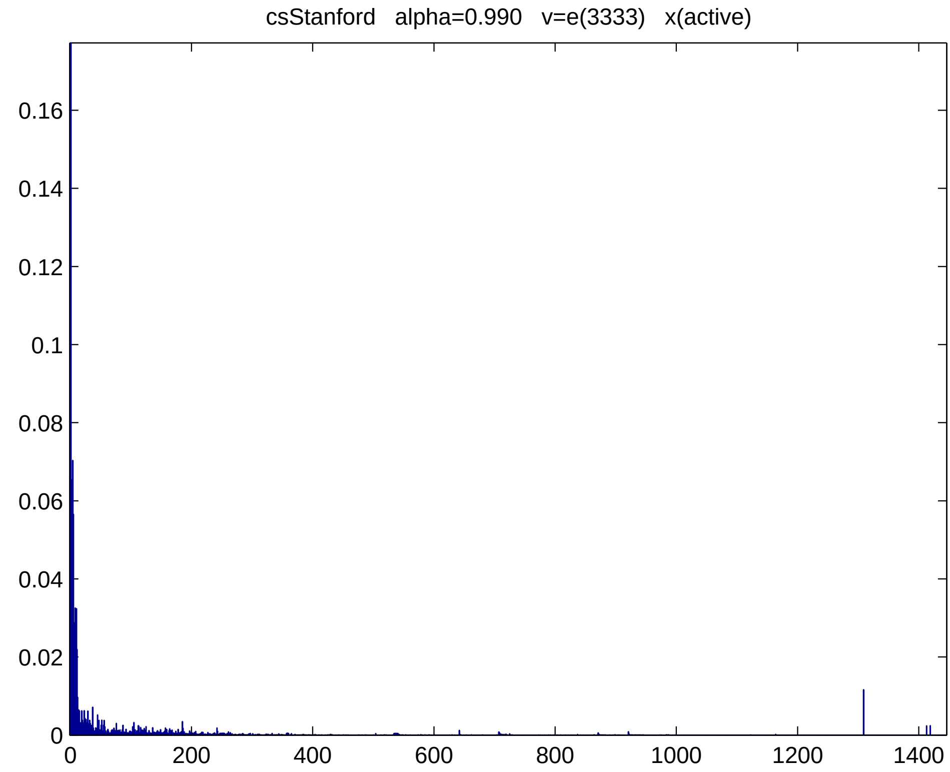


nonzero x_j

chosen in this order

csStanford $v = e_{3333}$

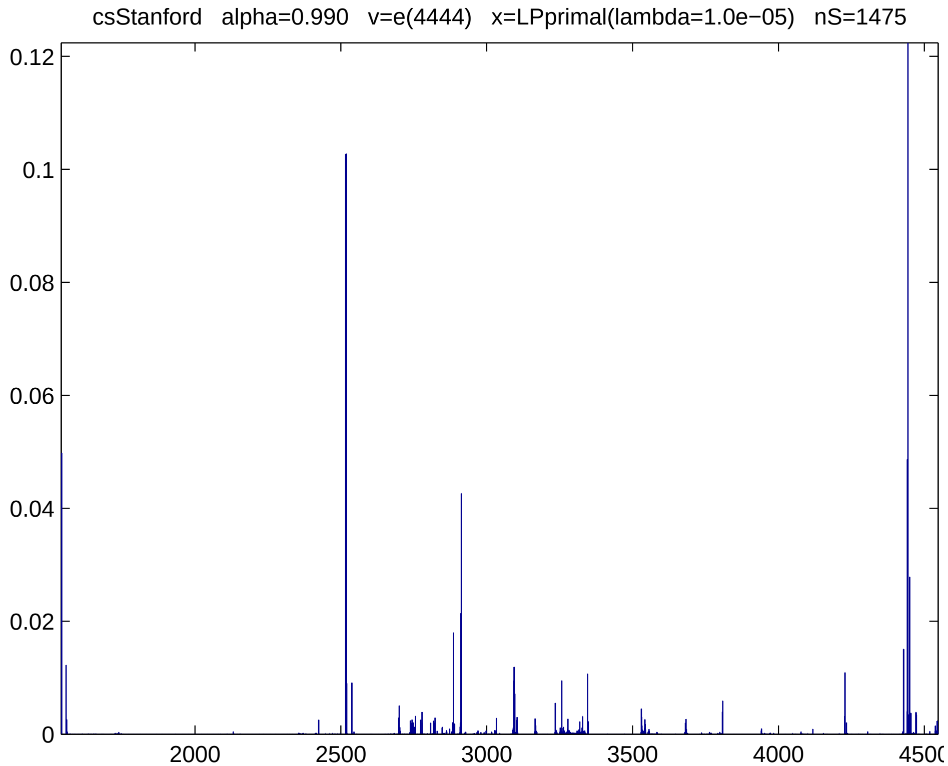


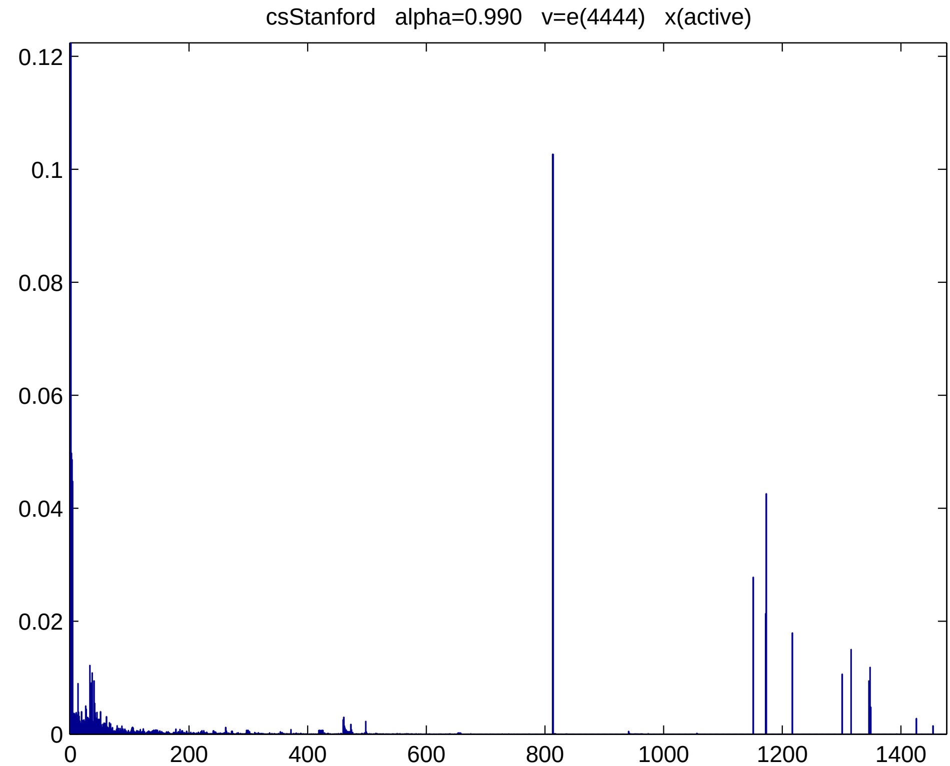


nonzero x_j

chosen in this order

csStanford $v = e_{4444}$

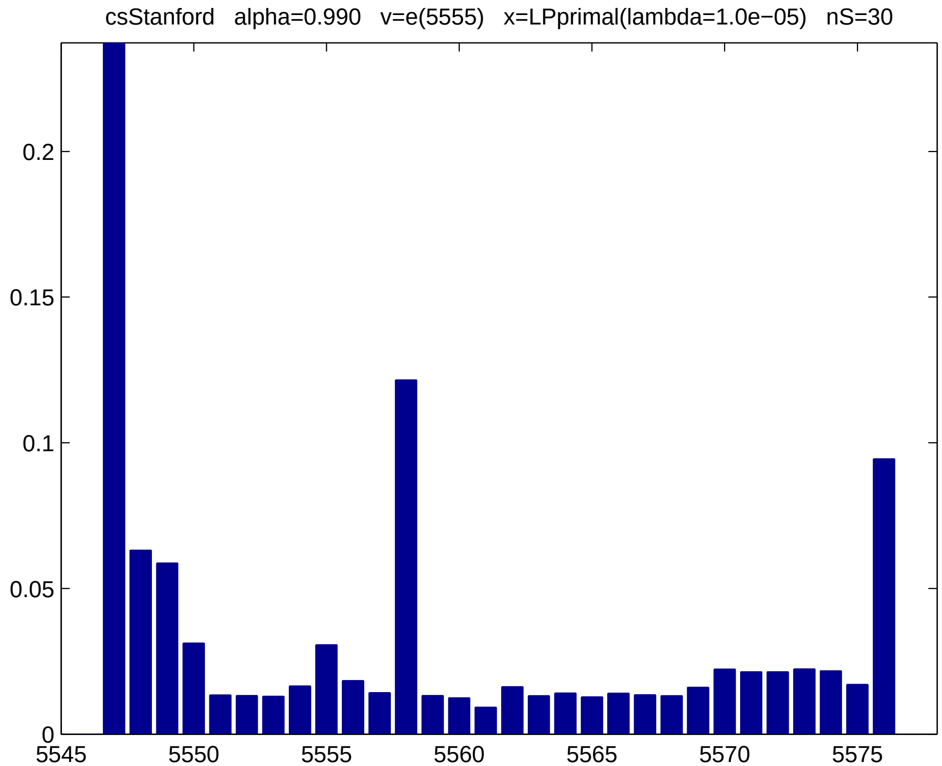


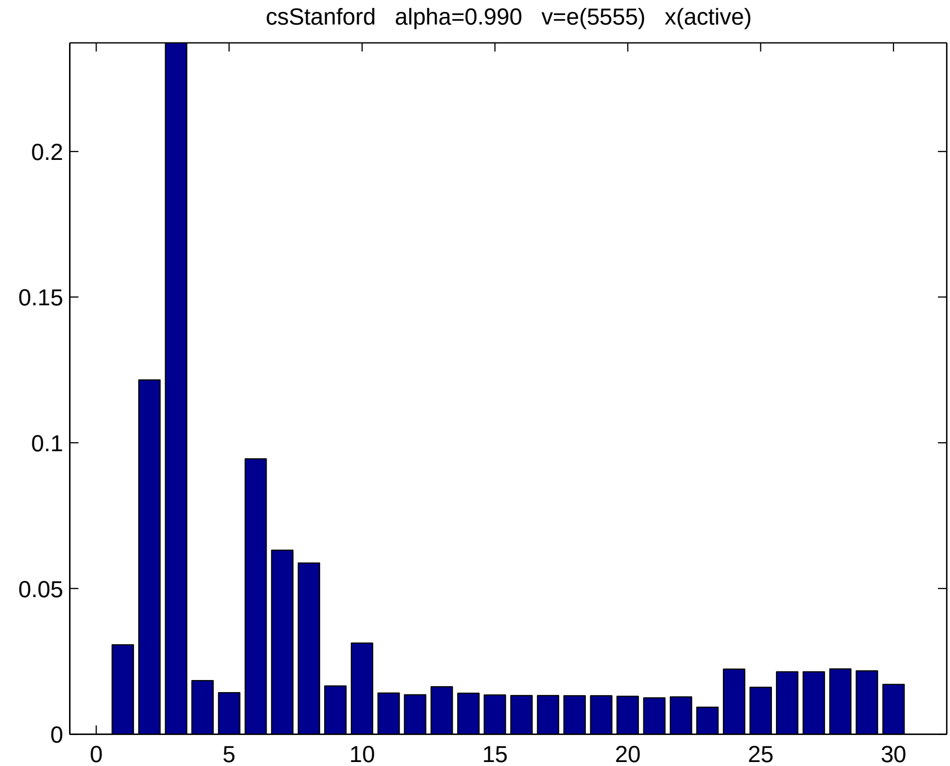


nonzero x_j

chosen in this order

csStanford $v = e_{5555}$

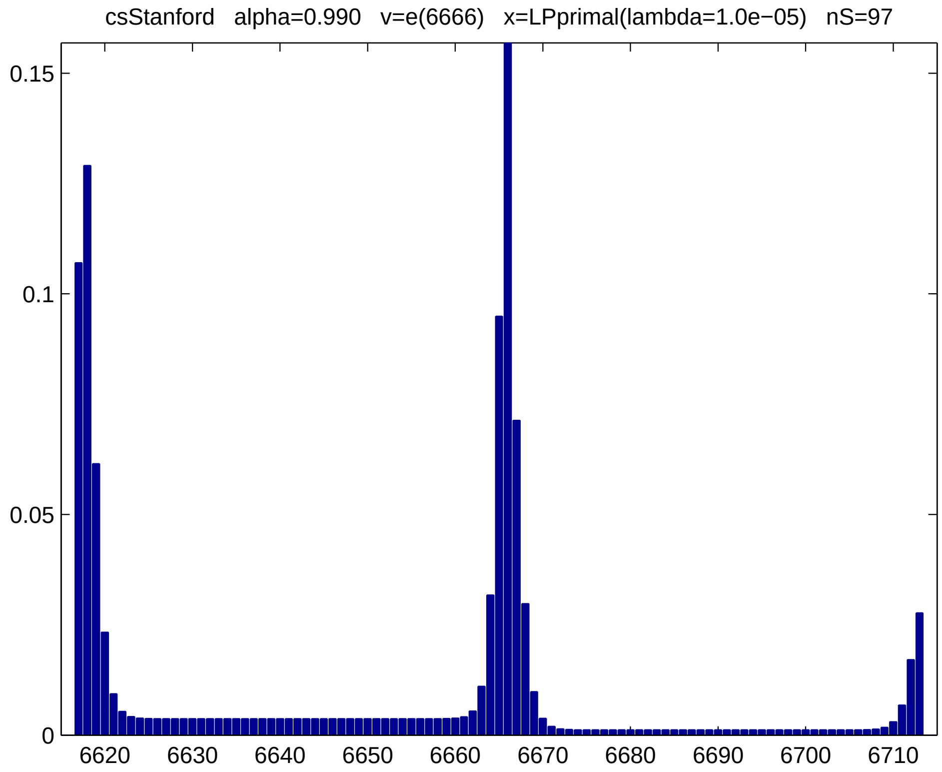


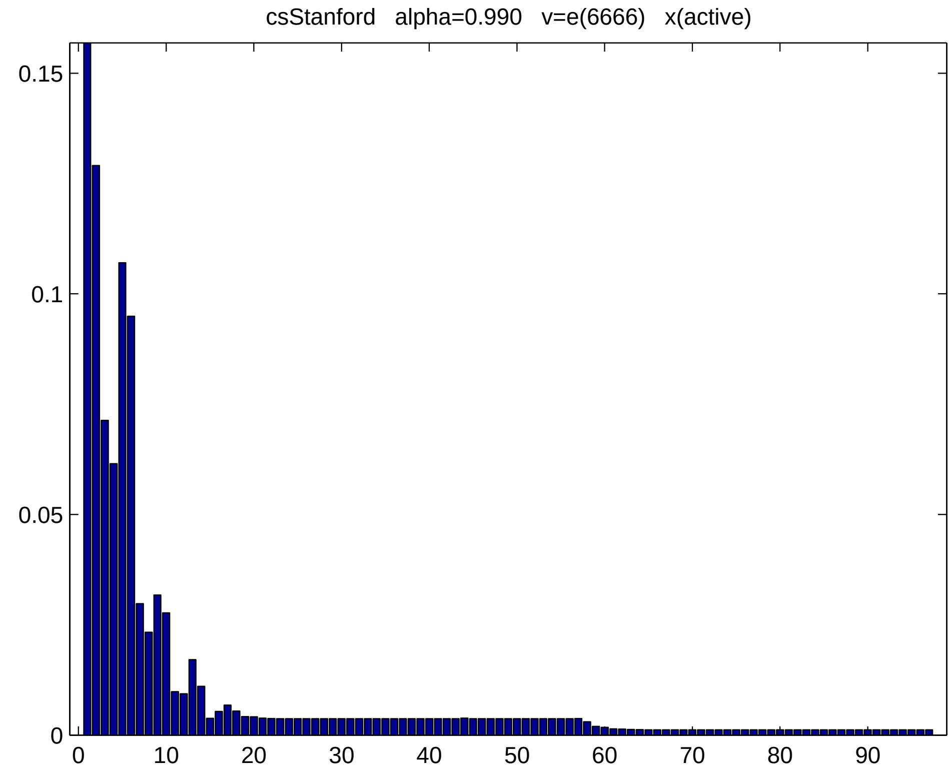


nonzero x_j

chosen in this order

csStanford $v = e_{6666}$

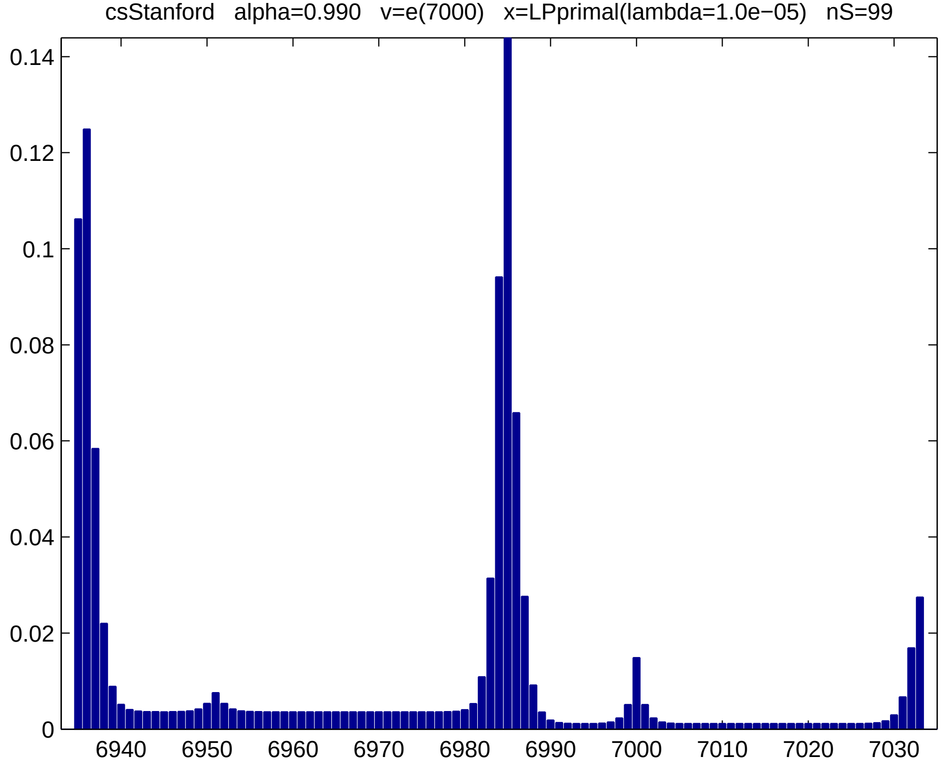


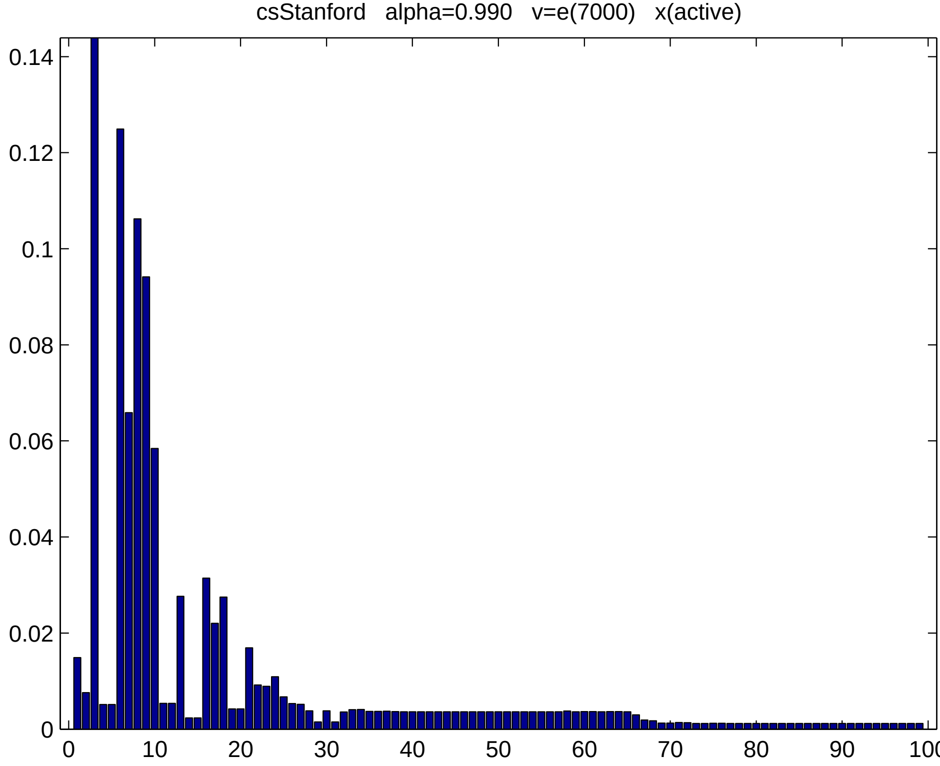


nonzero x_j

chosen in this order

csStanford $v = e_{7000}$





nonzero x_j

chosen in this order

LPdual solver

$$\min_y \; -v^T y + \tfrac{1}{2}\lambda\|y\|^2 \quad \text{st} \quad A^T y \le e$$

S = columns of A corresponding to $S^T y = e$ (active constraints)

Main work per itn:

Solve $\min \|Sx - g\|$
Form $dy = (g - Sx)/\lambda$
Form $dz = A^T dy$

Selects S in the same **greedy** manner as LPprimal, in almost the same order

$$(I - \alpha H^T)x = v$$

Comments

Sparse v **sometimes** gives sparse x

LPprimal and LPdual work in greedy manner

Greedy solvers guarantee sparse x (unlike power method)

Not sensitive to α

100, 1000, . . . times faster than power method for some v

$$(I - \alpha H^T)x = v$$

Comments

Sparse v **sometimes** gives sparse x

LPprimal and LPdual work in greedy manner

Greedy solvers guarantee sparse x (unlike power method)

Not sensitive to α

100, 1000, ... times faster than power method for some v

Questions

Run much bigger examples

Which properties of $I - \alpha H^T$ contribute to success of “greedy”

Thanks

Sou-Cheng Choi

Lek-Heng Lim

Michael Friedlander

Amy Langville and Carl Meyer (splendid book)

David Gleich

ICIAM 07 organizers