Sparsity in Polynomial Optimization

IMA Annual Program Year Workshop
"Optimization and Control"
Minneapolis, January 16-20, 2007

Masakazu Kojima
Tokyo Institute of Technology

- Some joint works with S.Kim, K.Kobayashi, M.Muramatsu & H.Waki
- Applications to nonlinear PDEs (partial differential equations)
--- Ongoing joint work with M.Mevissen, J.Nie & N.Takayama

Contents

1. How do we formulate structured sparsity?
   1-1. Unconstrained cases.
   1-2. Constrained and linear objective function cases.
   (Recent Joint work with S.Kim & K.Kobayashi)
2. Sparse SDP relaxation of constrained POPs.

3. Applications to PDEs (partial differential equations).

4. Concluding remarks.

POP --- Polynomial Optimization Problem

Contents

1. How do we formulate structured sparsity?
    1-1. Unconstrained cases.
    1-2. Constrained and linear objective function cases.
    (Recent Joint work with S.Kim & K.Kobayashi)
2. Sparse SDP relaxation of constrained POPs.
3. Applications to PDEs (partial differential equations).
4. Concluding remarks.

POP --- Polynomial Optimization Problem

POP: min $f_0(x)$ sub.to $f_j(x) \geq 0$ $(j = 1, \ldots, m)$.

$x = (x_1, \ldots, x_n) \in \mathbb{R}^n$: **a vector variable.**

$f_j(x)$: **a multivariate polynomial in** $x \in \mathbb{R}^n$ $(j = 0, 1, \ldots, m)$.

How do we exploit sparsity in POP?

$\Downarrow$

The answer depends on which methods we use to solve POP.

POP

$\Downarrow$ SDP relaxation (Lasserre 2001)

SDP $\Leftarrow$ Primal-Dual IPM (Interior-Point Method)

We will assume a structured sparsity (correlative sparsity):

(a) The size of SDP gets smaller.

(b) SDP satisfies "a similar structured sparsity" under which Primal-Dual IPM works efficiently.

Contents

1. How do we formulate structured sparsity?
    1-1. Unconstrained cases.
    1-2. Constrained and linear objective function cases.

2. Sparse SDP relaxation of constrained POPs.

3. Applications to PDEs (partial differential equations).

4. Concluding remarks.

POP --- Polynomial Optimization Problem

Unconstrained POP: mininimize $f_0(x)$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Define $n \times n$ csp (correlative sparsity pattern) matrix R

$$R_{ij} = \begin{cases} \star \text{ (nonzero symbol)} & \text{if } i = j \text{ or if } \partial^2 f(x)/\partial x_i \partial x_j \not\equiv 0, \\ 0 & \text{otherwise.} \end{cases}$$

(The sparsity pattern of the Hessian matrix of $f_0(x)$ except the diagonal)

Unconstrained POP : c-sparse (correlatively sparse) $\Leftrightarrow$
R allows a sparse (symbolic) Cholesky factorization
(under an ordering like the min. degree ordering).

Example. $f(x) = x_1^4 + 2x_1^2 x_2 + x_2^4 - x_2 x_3 + x_3^4 - 3x_3 x_4^2 + x_4^4 - x_4 x_5 + x_5^6$.

$$R = \begin{pmatrix} \star & \star & 0 & 0 & 0 \\ \star & \star & \star & 0 & 0 \\ 0 & \star & \star & \star & 0 \\ 0 & 0 & \star & \star & \star \\ 0 & 0 & 0 & \star & \star \end{pmatrix} = LL^T, \text{ where } L = \begin{pmatrix} \star & 0 & 0 & 0 & 0 \\ \star & \star & 0 & 0 & 0 \\ 0 & \star & \star & 0 & 0 \\ 0 & 0 & \star & \star & 0 \\ 0 & 0 & 0 & \star & \star \end{pmatrix}.$$

No fill-in in the Cholesky factorization.

Unconstrained POP: mininimize $f_0(x)$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Define $n \times n$ csp (correlative sparsity pattern) matrix R

$$R_{ij} = \begin{cases} \star \text{ (nonzero symbol)} & \text{if } i = j \text{ or if } \partial^2 f(x)/\partial x_i \partial x_j \not\equiv 0, \\ 0 & \text{otherwise.} \end{cases}$$

(The sparsity pattern of the Hessian matrix of $f_0(x)$ except the diagonal)

Unconstrained POP : c-sparse (correlatively sparse) $\Leftrightarrow$
R allows a sparse (symbolic) Cholesky factorization
(under an ordering like the min. degree ordering).

Numerical results on a sparse SDP relaxation applied to
three nonconvex test problems with opt.values = 0 from globalib

	B. tridiagonal		C. Wood		G. Rosenbrock	
n	approx.opt.val	cpu	apprx.opt.val	cpu	apprx.opt.val	cpu
600	1.0e-7	9.3	1.4e-5	0.9	3.9e-7	3.4
800	2.2e-7	12.6	1.8e-5	1.3	2.1e-7	5.1
1000	2.6e-7	16.0	3.8e-5	1.6	4.5e-7	5.9

Broyden tridiagonal function

$$f(x) = \sum_{i=1}^{n} ((3 - 2x_i)x_i - x_{i-1} - 2x_{i+1} + 1)^2, \text{ where } x_0 = x_{n+1} = 0.$$

Contents

1. How do we formulate structured sparsity?
   1-1. Unconstrained cases.
   1-2. Constrained and linear objective function cases.

2. Sparse SDP relaxation of constrained POPs.

3. Applications to PDEs (partial differential equations).

4. Concluding remarks.

- We consider cases where objective functions are linear.
- LP, SOCP and SDP + Primal-Dual Interior-Point Method.

Opt.Problem: max. $\sum_{i \in N} a_i y_i$ s.t. $(y_i : i \in I_p) \in C_p \ (p \in M)$

$$M = \{1, \ldots, m\}, \ N = \{1, \ldots, n\}, \ I_p \subset N \ (p \in M)$$

$(y_i : i \in I_p)$: a subvector of $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ consisting of elements $y_i \ (i \in I_p)$,

C_p : a nonempty subset of the set of all $(y_i : i \in I_p)$.

Define the $n \times n$ csp (correlative sparsity pattern) matrix R by

$$R_{ij} = \begin{cases} \star \text{ (nonzero symbol)} & \text{if } i = j \text{ or if } i, \ j \in I_p \text{ for } \exists p \in M, \\ 0 & \text{otherwise.} \end{cases}$$

Opt.Problem: c-sparse (correlatively sparse) $\Leftrightarrow$ R allows a sparse (symbolic) Cholesky factorization.

Example

$$C_p = \{ \ (y_p, y_{p+1}, y_n) \in \mathbb{R}^n : \ 1 - y_p^2 - y_{p+1}^2 - y_n^2 \geq 0,$$
$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} b_p & c_p \\ c_p & d_p \end{pmatrix} y_p + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} y_p y_{p+1} + \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} y_{p+1} \succeq O,$$
$$\left(0.3(y_p^3 + y_n) + 1\right) - \|(y_p + \beta_p \, y_n)\| \geq 0 \ \} \qquad (p = 1, \ldots, n-1).$$

Here $a_i, b_p, d_p \in (-1, 0)$, $c_p, \beta_p \in (0, 1)$ denote random numbers.

Opt.Problem: max. $\sum_{i \in N} a_i y_i$ s.t. $(y_i : i \in I_p) \in C_p \ (p \in M)$

$$M = \{1, \ldots, m\}, \ N = \{1, \ldots, n\}, \ I_p \subset N \ (p \in M)$$

$(y_i : i \in I_p)$: a subvector of $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ consisting of elements $y_i \ (i \in I_p)$,

C_p : a nonempty subset of the set of all $(y_i : i \in I_p)$.

Define the $n \times n$ csp (correlative sparsity pattern) matrix R by

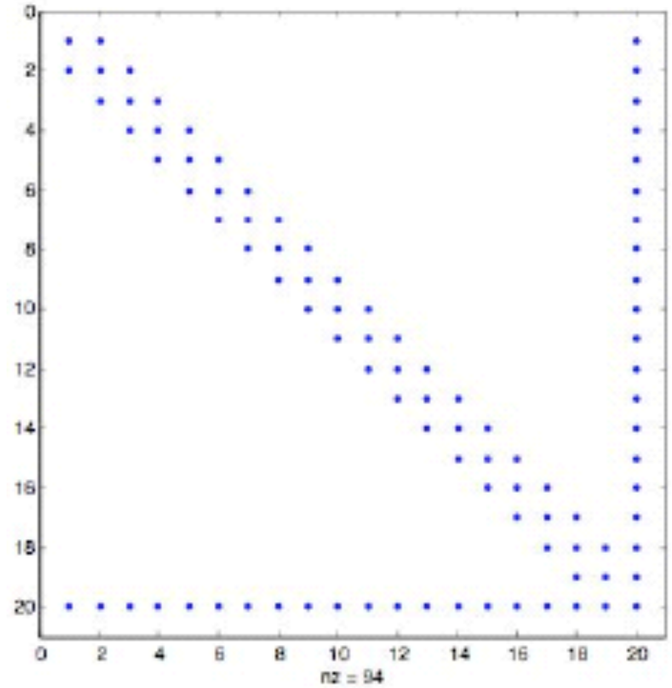
$$R_{ij} = \begin{cases} \star \text{ (nonzero symbol)} & \text{if } i = j \text{ or if } i, \ j \in I_p \text{ for } \exists p \in M, \\ 0 & \text{otherwise.} \end{cases}$$

Opt.Problem: c-sparse (correlatively sparse) $\Leftrightarrow$ R allows a sparse (symbolic) Cholesky factorization.

Example

csp matrix R =

(n=20)



Opt.Problem: max. $\sum_{i \in N} a_i y_i$ s.t. $(y_i : i \in I_p) \in C_p \ (p \in M)$

$$M = \{1, \ldots, m\}, \ N = \{1, \ldots, n\}, \ I_p \subset N \ (p \in M)$$

$(y_i : i \in I_p)$: a subvector of $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ consisting of elements y_i $(i \in I_p)$,

C_p : a nonempty subset of the set of all $(y_i : i \in I_p)$.

Define the $n \times n$ csp (correlative sparsity pattern) matrix R by

$$R_{ij} = \begin{cases} \star \text{ (nonzero symbol)} & \text{if } i = j \text{ or if } i, \ j \in I_p \text{ for } \exists p \in M, \\ 0 & \text{otherwise.} \end{cases}$$

Opt.Problem: c-sparse (correlatively sparse) ⇔ R allows a sparse (symbolic) Cholesky factorization.

Example

Numerical results on the sparse SDP relaxation

n	cpu sec.	ϵ_{obj}	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros in A
600	25.7	4.0e-12	0.0	11,974 × 113,022	235,612
800	34.8	3.2e-12	0.0	15,974 × 150,822	314,412
1000	44.5	1.6e-12	0.0	19,974 × 188,622	393,212

Opt.Problem: max. $\sum_{i \in N} a_i y_i$ s.t. $(y_i : i \in I_p) \in C_p$ $(p \in M)$

$$M = \{1, \ldots, m\},\ N = \{1, \ldots, n\},\ I_p \subset N\ (p \in M)$$

$(y_i : i \in I_p)$: a subvector of $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ consisting of elements y_i $(i \in I_p)$,

C_p : a nonempty subset of the set of all $(y_i : i \in I_p)$.

Define the $n \times n$ csp (correlative sparsity pattern) matrix R by

$$R_{ij} = \begin{cases} \star \text{ (nonzero symbol)} & \text{if } i = j \text{ or if } i,\ j \in I_p \text{ for } \exists p \in M, \\ 0 & \text{otherwise.} \end{cases}$$

Opt.Problem: c-sparse (correlatively sparse) $\Leftrightarrow$ R allows a sparse (symbolic) Cholesky factorization.

(a) $\forall C_p$ is described by poly. (matrix or second-order cone) inequalites.
$\Rightarrow$ A sparse SDP relaxation whose csp matrix R' is of "a similar sparsity pattern" to R; the size of $R' \geq$ the size of R.

(b) $\forall C_p$ is described by linear matrix inequalites (SDP)
$\Rightarrow$ The coef. matrix B of the Schur complement eq. $Bdy = r$, which is the most time consuming in Primal-dual IPMs, for a search direction dy has the same pattern as the csp matrix R' of SDP.

Contents

1. How do we formulate structured sparsity?
   1-1. Unconstrained cases.
   1-2. Constrained and linear objective function cases.

2. Sparse SDP relaxation of constrained POPs.
   (Joint work with S.Kim, M.Muramatsu & H.Waki)
3. Applications to PDEs (partial differential equations).

4. Concluding remarks.

Sections 1-1 + 1-2 ==> Section 2

Sparse SDP relaxation = Modification of Lasserre’s relaxation

POP: max. $f_0(x)$ s.t. $(x_i : i \in I_p) \in C_p$ $(p \in M)$

$M = \{1, \ldots, m\}$, $N = \{1, \ldots, n\}$, $I_p \subset N$ $(p \in M)$

$C_p \subset$ the set of all $(x_i : i \in I_p)$, described as poly. inequalites.

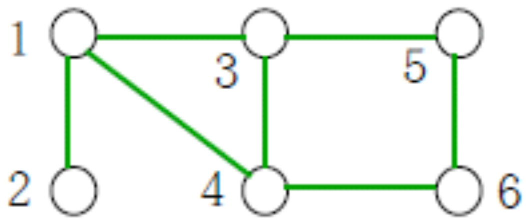
$$R_{ij} = \begin{cases} \star \text{ (nonzero symbol)} & \text{if } i = j,\ \partial^2 f_0(x)/\partial x_i \partial x_j \not\equiv 0, \\ & \text{or } i,\ j \in I_p \text{ for } \exists p \in M, \\ 0 & \text{otherwise.} \end{cases}$$

POP : c-sparse (correlatively sparse) $\Leftrightarrow$
The $n \times n$ csp matrix $R = (R_{ij})$ allows a sparse Cholesky factorization.

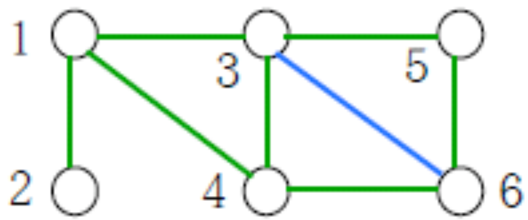
$E = \{\{i, j\} \in N \times N : R_{ij} = \star,\ i \neq j\}$ $\Updownarrow$

POP : c-sparse (correlatively sparse) $\Leftrightarrow$
The csp graph $G(N, E)$ has a sparse chordal extension $G(N, \bar{E})$; $E \subseteq \bar{E}$.

$G(N, E)$: not chordal



$G(N, \bar{E})$: chordal



∀ cycle having more than 3 edges has a chord.

- The added edge $\{3, 6\}$ is corresponding to a fill-in.
- The maximal cliques $= \{1, 2\}, \{1, 3, 4\}, \{3, 4, 6\}, \{3, 5, 6\}$.

POP: max. $f_0(x)$ s.t. $(x_i : i \in I_p) \in C_p \ (p \in M)$

$M = \{1, \ldots, m\}$, $N = \{1, \ldots, n\}$, $I_p \subset N \ (p \in M)$

$C_p \subset$ the set of all $(x_i : i \in I_p)$, described as poly. inequalites.

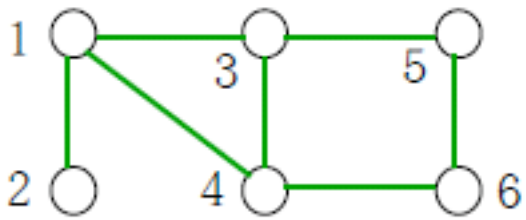
POP : c-sparse (correlatively sparse) $\Leftrightarrow$
The csp graph $G(N, E)$ has a sparse chordal extension $G(N, \bar{E})$; $E \subseteq \bar{E}$.

Two steps to derive a sparse SDP relaxation of POP

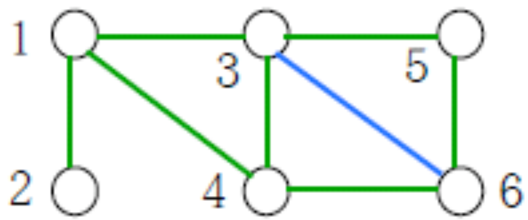
(a) Using the max. cliques $J_q \ (q \in L)$ of $G(N, \bar{E})$, we convert POP into an equivalent poly.SDP with the csp graph $G(N, \bar{E})$.

(b) Linearize poly.SDP $\Rightarrow$ SDP with a similar sparsity to poly.SDP.

$G(N, E)$: not chordal



$G(N, \bar{E})$: chordal



$\forall$ cycle having more than 3 edges has a chord.

- The added edge $\{3, 6\}$ is corresponding to a fill-in.
- The maximal cliques $= \{1, 2\}, \{1, 3, 4\}, \{3, 4, 6\}, \{3, 5, 6\}$.

POP: max. $f_0(x)$ s.t. $(x_i : i \in I_p) \in C_p$ $(p \in M)$

$M = \{1,\ldots,m\}$, $N = \{1,\ldots,n\}$, $I_p \subset N$ $(p \in M)$

$C_p \subset$ the set of all $(x_i : i \in I_p)$, described as poly. inequalites.

POP : c-sparse (correlatively sparse) $\Leftrightarrow$
The csp graph $G(N,E)$ has a sparse chordal extension $G(N,\bar{E})$; $E \subseteq \bar{E}$.

Two steps to derive a sparse SDP relaxation of POP

(a) Using the max. cliques J_q $(q \in L)$ of $G(N,\bar{E})$, we convert POP into an equivalent poly.SDP with the csp graph $G(N,\bar{E})$.

(b) Linearize poly.SDP $\Rightarrow$ SDP with a similar sparsity to poly.SDP.

Notation: For every nonnegative integer s, let $u_s(x_i : i \in J_q)$ denote the column vector of monomials with degree at most s in variables x_i $(i \in J_q)$.

Example: Let $J_q = \{1,4\}$. Then

$$s = 0 \Rightarrow u_0(x_i : i \in J_q) = 1,$$
$$s = 1 \Rightarrow u_1(x_i : i \in J_q) = (1, x_1, x_4)^T,$$
$$s = 3 \Rightarrow u_3(x_i : i \in J_q) = (1, x_1, x_4, x_1^2, x_1x_4, x_4^2, x_1^3, x_1^2x_4, x_1x_4^2, x_4^3)^T,$$

$$s = 1 \Rightarrow u_1(x_i : i \in J_q)u_1(x_i : i \in J_q)^T = \begin{pmatrix} 1 & x_1 & x_4 \\ x_1 & x_1^2 & x_1x_4 \\ x_4 & x_1x_4 & x_4^2 \end{pmatrix}$$

POP: max. $f_0(x)$ s.t. $(x_i : i \in I_p) \in C_p$ $(p \in M)$

$M = \{1, \ldots, m\}$, $N = \{1, \ldots, n\}$, $I_p \subset N$ $(p \in M)$
$C_p \subset$ the set of all $(x_i : i \in I_p)$, described as poly. inequalites.

POP : c-sparse (correlatively sparse) $\Leftrightarrow$
The csp graph $G(N, E)$ has a sparse chordal extension $G(N, \bar{E})$; $E \subseteq \bar{E}$.

Two steps to derive a sparse SDP relaxation of POP

(a) Using the max. cliques J_q $(q \in L)$ of $G(N, \bar{E})$, we convert POP into an equivalent poly.SDP with the csp graph $G(N, \bar{E})$.

(a-1) Let $r_0 = \lceil \deg(\text{POP})/2 \rceil \equiv \lceil$"the max.deg. of the poly. in POP"$/2\rceil$.

(a-2) Choose $r \geq r_0$; a sequence of poly.SDPs depending on $r \geq r_0$.

r : the relaxation order of the sparse SDP relaxation of POP;
$r = \lceil \deg(\text{poly.SDP})/2 \rceil$

(a-3) Replace each $f(x_i : i \in I_p) \geq 0$ involved in C_p by an equivalent

$$f(x_i : i \in I_p) u_s(x_i : i \in J_q) u_s(x_i : i \in J_q)^T \succeq O,$$

where $s = r - \lceil$"the degree of $f(x_i : i \in I_p)$"$/2\rceil$ and $I_p \subseteq J_q$.

(a-4) Add (redundant) $u_r(x_i : i \in J_q) u_r(x_i : i \in J_q)^T \succeq O$ $(q \in L)$ to POP.

An equiv.poly.SDP with the csp graph $G(N, \bar{E})$ of the form

$$\max.\ f_0(x) \text{ s.t. } P_j(x) \succeq O \ (j = 1, \ldots, \ell).$$

Here $P_j(x)$: a poly. with sym. mat. coefficients.

POP: max. $f_0(x)$ s.t. $(x_i : i \in I_p) \in C_p$ $(p \in M)$

$$M = \{1, \ldots, m\},\ N = \{1, \ldots, n\},\ I_p \subset N\ (p \in M)$$

$C_p \subset$ the set of all $(x_i : i \in I_p)$, described as poly. inequalites.

POP : c-sparse (correlatively sparse) $\Leftrightarrow$
The csp graph $G(N, E)$ has a sparse chordal extension $G(N, \bar{E})$; $E \subseteq \bar{E}$.

Two steps to derive a sparse SDP relaxation of POP

(a) Using the max. cliques J_q $(q \in L)$ of $G(N, \bar{E})$, we convert POP into an equivalent poly.SDP with the csp graph $G(N, \bar{E})$.

An equiv.poly.SDP with the csp graph $G(N, \bar{E})$ of the form

$$\max.\ f_0(x) \text{ s.t. } P_j(x) \succeq O\ (j = 1, \ldots, \ell).$$

Here $P_j(x)$: a poly. with sym. mat. coefficients.

Represent poly.SDP as

$$\max.\ \sum_{\alpha \in \mathcal{A}_0} g_0(\alpha) x^{\alpha} \text{ s.t. } \sum_{\alpha \in \mathcal{A}_j} G_j(\alpha) x^{\alpha} \succeq O\ (j = 1, \ldots, \ell).$$

$\Downarrow$ (b) Linearize by replacing each x^{α} by an indep. var. y_{α}

SDP: max. $\sum_{\alpha \in \mathcal{A}_0} g_0(\alpha) y_{\alpha}$ s.t. $\sum_{\alpha \in \mathcal{A}_j} G_j(\alpha) y_{\alpha} \succeq O\ (j = 1, \ldots, \ell)$,

which forms a sparse SDP relaxation of POP.

- poly.SDP dep.on $r \geq r_0 = \lceil \deg(\text{POP})/2 \rceil \Rightarrow$ a seq.of SDPs dep.on $r \geq r_0$.
- Under an assump., opt.val.SDP $\to$ opt.val.POP as $r \to \infty$ (Lasserre '05).

Example

POP: min. $\sum_{i=1}^{3}(-x_i^3)$ s.t. $-i \times x_i^2 - x_4^2 + 1 \geq 0 \ (i = 1,2,3)$.

$\Updownarrow$ (a) with the relaxation order $r = 2 \geq r_0 = \lceil 3/2 \rceil = 2$

poly.SDP

$$\begin{array}{ll} \text{min.} & \sum_{i=1}^{3}(-x_i^3) \\ \text{s.t.} & (-i \times x_i^2 - x_4^2 + 1)(1, x_i, x_4)^T(1, x_i, x_4) \succeq O \ (i = 1,2,3), \\ & (1, x_i, x_4, x_i^2, x_i x_4, x_4^2)^T(1, x_i, x_4, x_i^2, x_i x_4, x_4^2) \succeq O \ (i = 1,2,3). \end{array}$$

Represent poly.SDP as

$$\text{min.} \sum_{\alpha \in \mathcal{A}_0} g_0(\alpha) x^{\alpha} \ \text{s.t.} \sum_{\alpha \in \mathcal{A}_j} G_j(\alpha) x^{\alpha} \succeq O \ (j = 1, \ldots, 6),$$

where $\mathcal{A}_j \subset \mathbb{Z}_+^4$ and $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} x_4^{\alpha_4}$; $x^{(1,2,1,0)} = x_1 x_2^2 x_3$.

$\Downarrow$ (b) Linearize by replacing each x^{α} by an indep. var. y_{α}; x^0 by 1

$$\text{SDP min.} \sum_{\alpha \in \mathcal{A}_0} g_0(\alpha) y_{\alpha} \ \text{s.t.} \sum_{\alpha \in \mathcal{A}_j} G_j(\alpha) y_{\alpha} \succeq O \ (j = 1, \ldots, 6),$$

which forms an SDP relaxation of POP.

Contents

1. How do we formulate structured sparsity?
   1-1. Unconstrained cases.
   1-2. Constrained and linear objective function cases.

2. Sparse SDP relaxation of constrained POPs.

3. Applications to PDEs (partial differential equations).
   (Ongoing joint work with M.Mevissen, J.Nie & N.Takayama)

4. Concluding remarks.

- **Various numerical methods have been developed for (nonlinear) PDEs.**

- **Is SDP relaxation of POPs useful in solving PDEs?**

- We are not sure how far we can go; so far only small size PDEs with at most two independent variables and two unknown functions.

- Challenge to PDEs using SDP relaxation of POPs.

Basic idea of solving a PDE by using SDP relaxation of POPs.

PDE with some boundary conditions such as
Dirichlet, Neumann and periodic conditions

Assump. PDE is described as "a mult. poly. equation." in unknown functions and their derivatives for each fixed independent variables.

Example 1 (A nonlinear elliptic equation with an inhomogeneous term):

$$u_{xx}(x,y) + u_{yy}(x,y) + 22u(x,y)(1 - u(x,y)^2) + 5\sin(\pi x)\sin(2\pi y) = 0,$$

$$u(0,y) = u(1,y) = u(x,0) = u(x,1) = 0, \ \forall (x,y) \in [0,1] \times [0,1].$$

Example 2 — A nonlinear wave equation with a periodic condition.

Examples 3 & 4
— 2 unknown cases with Dirichlet and Neumann conditions, respectively (modifications of the Ginzburg-Landau equation for superconductivity).

We will show some numerical results on these examples later.

Basic idea of solving a PDE by using SDP relaxation of POPs.

PDE with some boundary conditions such as Dirichlet, Neumann and periodic conditions

$\Downarrow$ discretize on finite grid points; approximate partial derivatives by finite differences

A system of polynomial equations

$\Downarrow$ add an objective function and/or polynomial inequality constraints

A POP (Polynomial Optimization Problem)

$\Downarrow$ apply SDP relaxation with $\exists$ relaxation order r

A discretized solution of PDE

Advantage

(a) We can add an objective function and/or polynomial inequality constraints to pick up a specific solution which we want to compute.

(b) The system of polynomial equations induced from PDE satisfies the correlative sparsity.

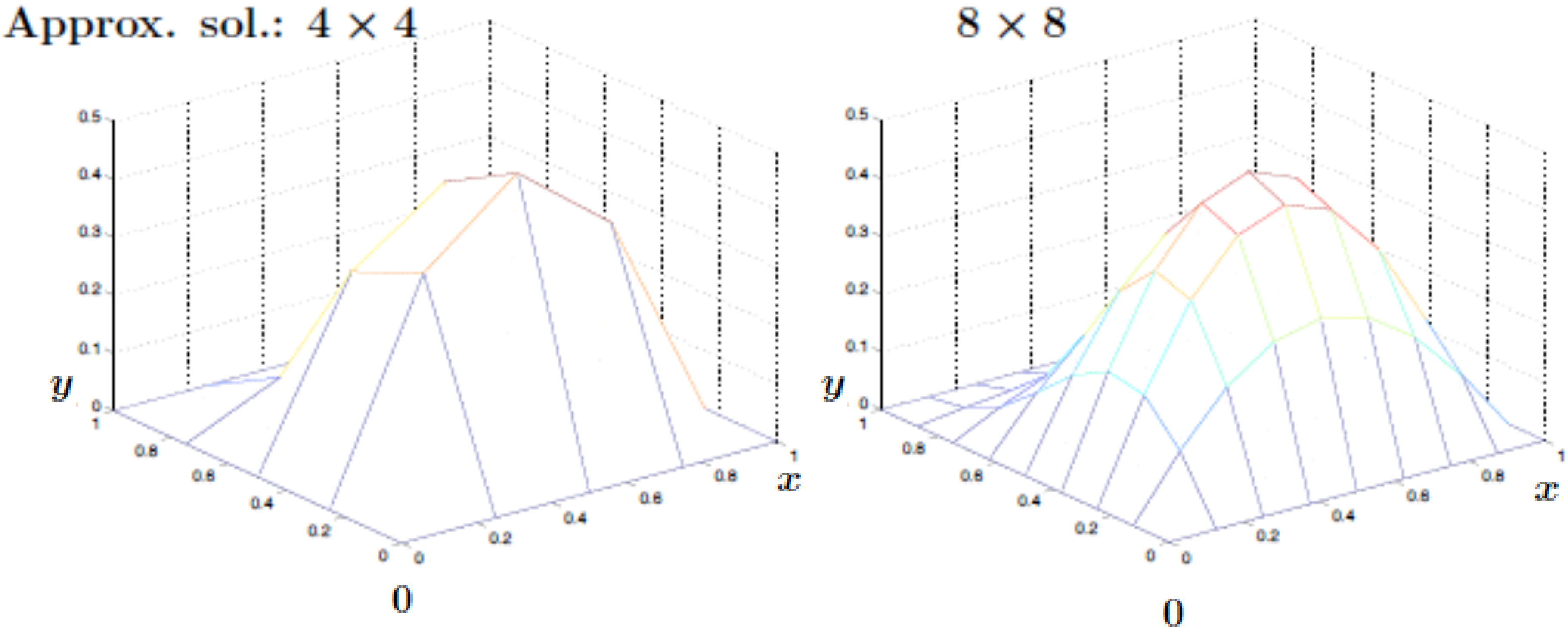
But (c) Expensive, depending on a relaxation order r unknown in advance.

Example 1 (A nonlinear elliptic equation with an inhomogeneous term):

$$u_{xx}(x,y) + u_{yy}(x,y) + 22u(x,y)(1 - u(x,y)^2) + 5\sin(\pi x)\sin(2\pi y) = 0,$$

$$u(0,y) = u(1,y) = u(x,0) = u(x,1) = 0,\ \forall (x,y) \in [0,1] \times [0,1].$$

grid size	# of var.	cpu sec.	relax. order r	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros in A
4×4	9	0.92	2	8.4e-11	$183 \times 1{,}506$	2013
8×4	21	1.7	2	4.7e-10	$544 \times 4{,}807$	6,380
8×8	49	33.1	2	1.5e-10	$3{,}642 \times 31{,}907$	42,425



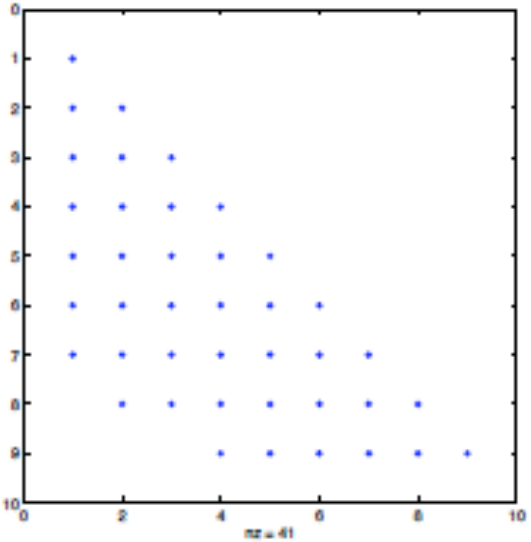
Example 1 (A nonlinear elliptic equation with an inhomogeneous term):

$$u_{xx}(x,y) + u_{yy}(x,y) + 22u(x,y)(1 - u(x,y)^2) + 5\sin(\pi x)\sin(2\pi y) = 0,$$
$$u(0,y) = u(1,y) = u(x,0) = u(x,1) = 0,\ \forall (x,y) \in [0,1] \times [0,1].$$

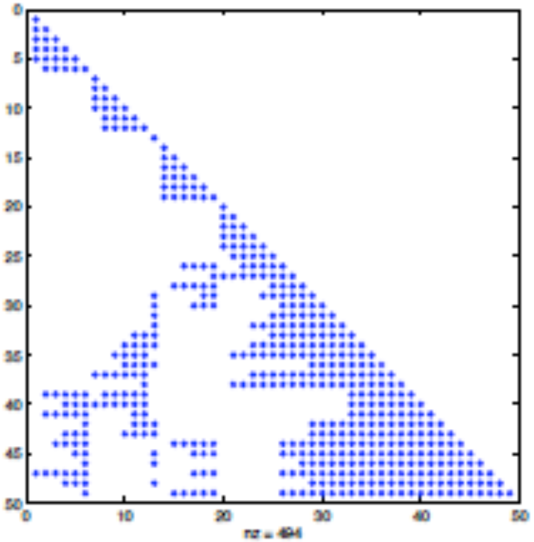
grid size	# of var.	cpu sec.	relax. order r	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros in A
4×4	9	0.92	2	8.4e-11	$183 \times 1{,}506$	2013
8×4	21	1.7	2	4.7e-10	$544 \times 4{,}807$	6,380
8×8	49	33.1	2	1.5e-10	$3{,}642 \times 31{,}907$	42,425

A sparse Cholesky factorization of the CSP matrix under a symmetic minimum degree ordering:

4×4
$\Downarrow$
41/45
nonzeros



8×8
$\Downarrow$
494/1225
nonzeros



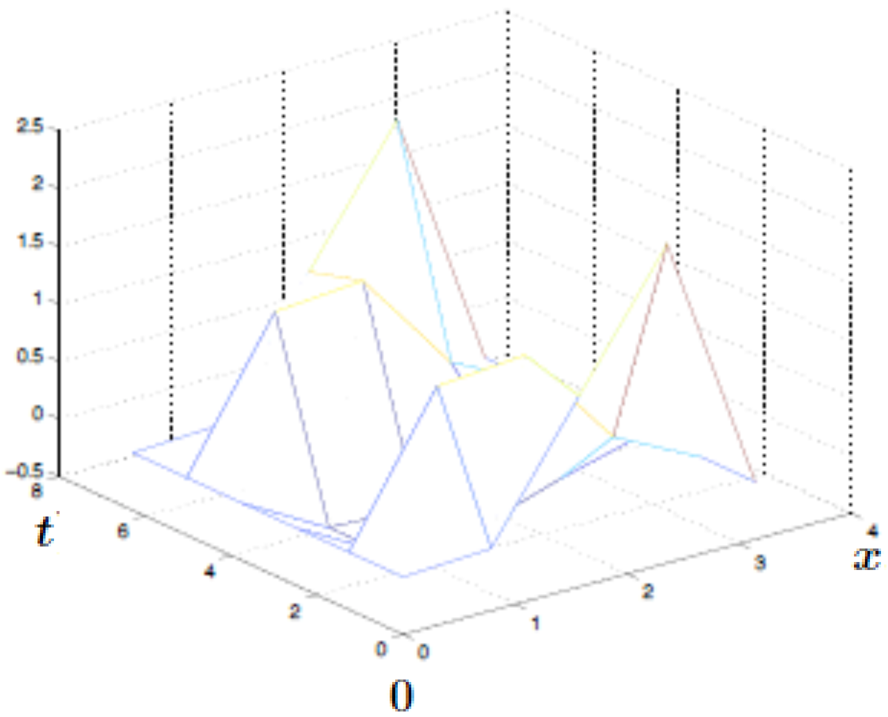
Example 2 (A nonlinear wave equation on $[0,\pi] \times [0, 2\pi]$):

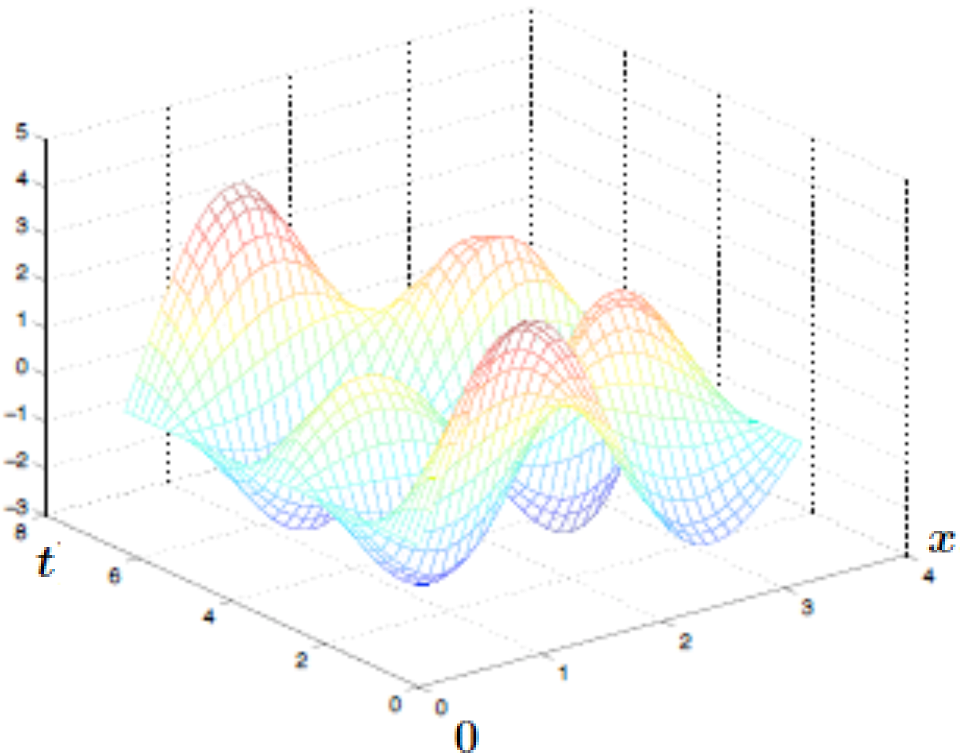
$$-u_{xx}(x,t) + u_{tt}(x,t) + u(x,t)(1 - u(x,t)) + 0.2\sin(x) = 0,$$
$$u(0,t) = u(\pi,t) = 0,\ \forall t \in [0,2\pi],\ \ u(x,0) = u(x,2\pi),\ \forall x \in [0,\pi].$$

grid size	# of var.	cpu sec.	relax. order r	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros in A
4×5	15	159.5	2	9.4e-10	2,616 $\times$ 36,029	43,689

"Multigrid technique"

Approx. sol.: 4×5 $\Longrightarrow$ 32×40





Example 2 (A nonlinear wave equation on $[0, \pi] \times [0, 2\pi]$):

$$-u_{xx}(x,t) + u_{tt}(x,t) + u(x,t)(1 - u(x,t)) + 0.2\sin(x) = 0,$$

$$u(0,t) = u(\pi,t) = 0,\ \forall t \in [0, 2\pi],\ u(x,0) = u(x,2\pi),\ \forall x \in [0,\pi].$$

grid size	# of var.	cpu sec.	relax. order r	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros in A
4×5	15	159.5	2	9.4e-10	2,616 × 36,029	43,689

"Multigrid technique"

$\Longrightarrow$

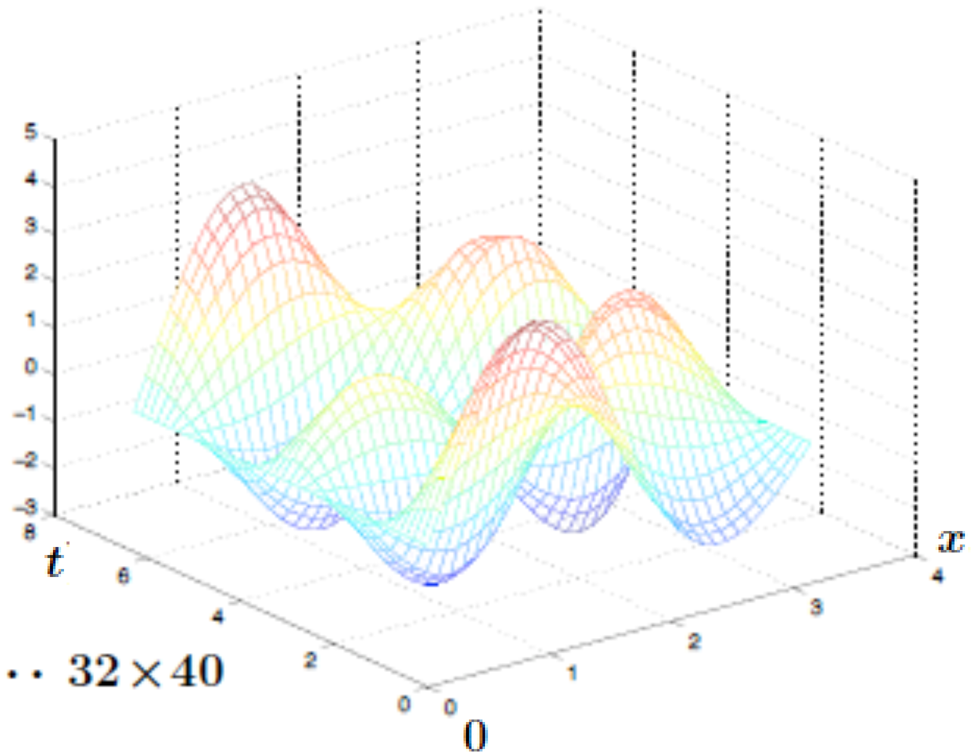
32×40

1. A rough approx. sol. u^0 for 8×5 case by interpolation to the solution of 4×5 case.

2-a. Sparse SDP relax. to 8×5 case with obj.funct. $\|u - u^0\|^2 \downarrow$, $u^0_k - \epsilon \leq u_k \leq u^0 + \epsilon$, $\forall k$ ($\epsilon = 0.5$), and $r = 1$, or

2-b. Newton meth. to 8×5 case with the init. pt. u^0.

(2-a is more expensive, but robust(?))



• $4 \times 5 \Rightarrow 8 \times 5 \Rightarrow 8 \times 10 \cdots 32 \times 40$

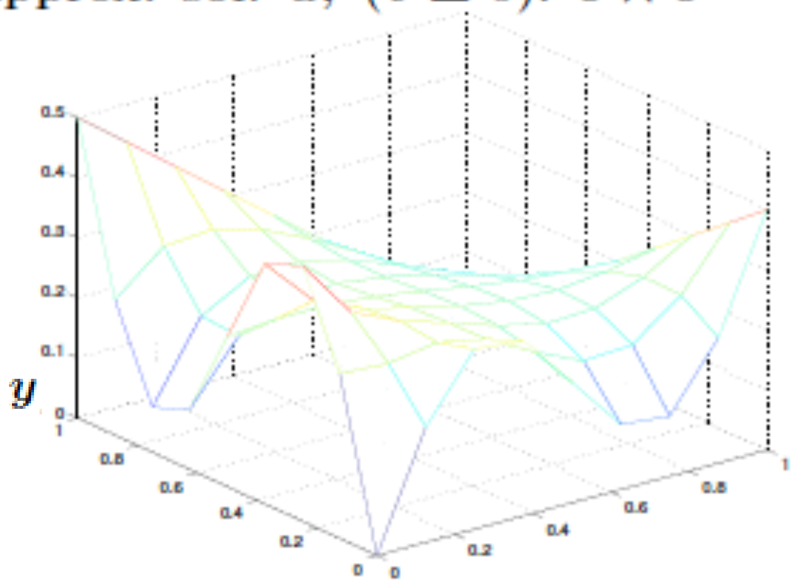
Example 3 (2 unknown case on $[0,1] \times [0,1]$, Dirichlet condition):

$$u_{xx}(x,y) + u_{yy}(x,y) + u(x,y)(1 - u(x,y)^{\rho} - v(x,y)^{\rho}) = 0,$$
$$v_{xx}(x,y) + v_{yy}(x,y) + v(x,y)(1 - u(x,y)^{\rho} - v(x,y)^{\rho}) = 0,$$
$$u(0,y) = 0.5y + 0.3\sin(2\pi y),\ u(1,y) = 0.4 - 0.4y,\ \forall y \in [0,1],$$
$$u(x,0) = 0.4x + 0.2\sin(2\pi x),\ u(x,1) = 0.5 - 0.5x,\ \forall x \in [0,1],$$
$$v(x,0) = v(x,1) = v(0,y) = v(1,y) = 0,\ \forall x \in [0,1],\ \forall y \in [0,1].$$

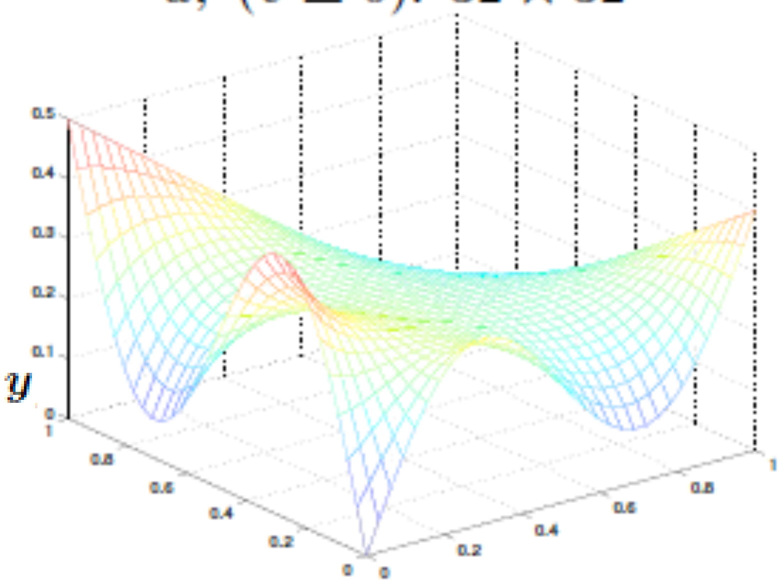
ρ	grid size	# of var.	cpu sec.	relax. order r	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros in A
1	8×8	98	19.0	1	6.2e-07	1,999 $\times$ 21,377	21,865
2	8×8	98	10,959	2	9.4e-07	25,699 $\times$ 235,471	319,306

$\rho = 2$ case, Sparse SDP relaxation + 2.b (Newton Method)

Approx. sol. u, $(v \equiv 0)$: 8×8



u, $(v \equiv 0)$: 32×32



0

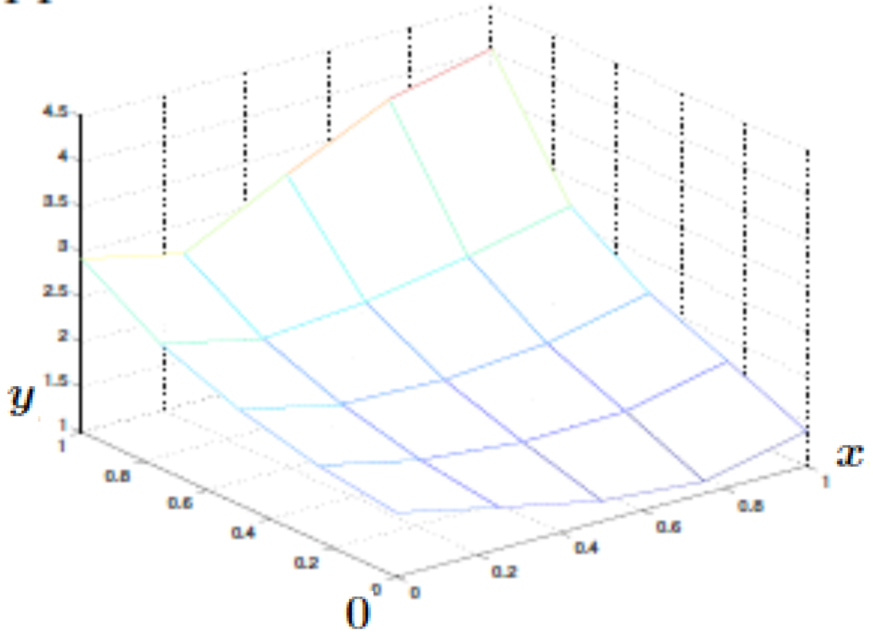
0

Example 4 (2 unknown case on $[0,1] \times [0,1]$, Neumann condition):

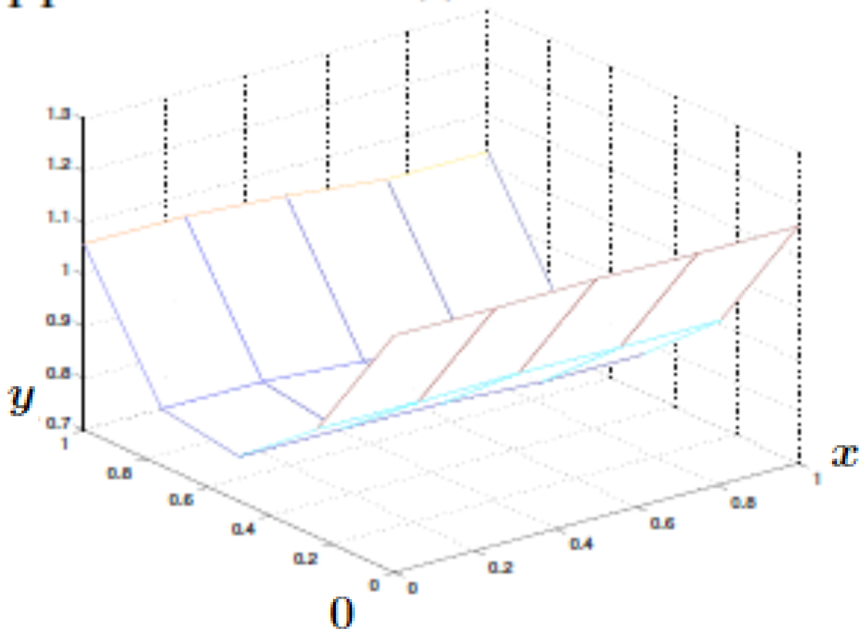
$$\begin{aligned}
&u_{xx}(x,y) + u_{yy}(x,y) + u(x,y)(1 - u(x,y)^2 - v(x,y)^2) = 0,\\
&v_{xx}(x,y) + v_{yy}(x,y) + v(x,y)(1 - u(x,y)^2 - v(x,y)^2) = 0,\\
&u_x(0,y) = -1,\ u_x(1,y) = 1,\ \forall y \in [0,1],\\
&u_y(x,0) = 2x,\ u_y(x,1) = x + 5\sin(\pi x/2),\ \forall x \in [0,1],\\
&v_x(0,y) = 0,\ v_x(1,y) = 0,\ \forall y \in [0,1],\\
&v_y(x,0) = -1,\ v_y(x,1) = 1,\ \forall x \in [0,1].
\end{aligned}$$

grid size	# of var.	cpu sec.	relax. order r	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros in A
4×4	18	6.9	2	1.7e-10	979 × 9,165	12,598

Approx. sol. u: 4×4



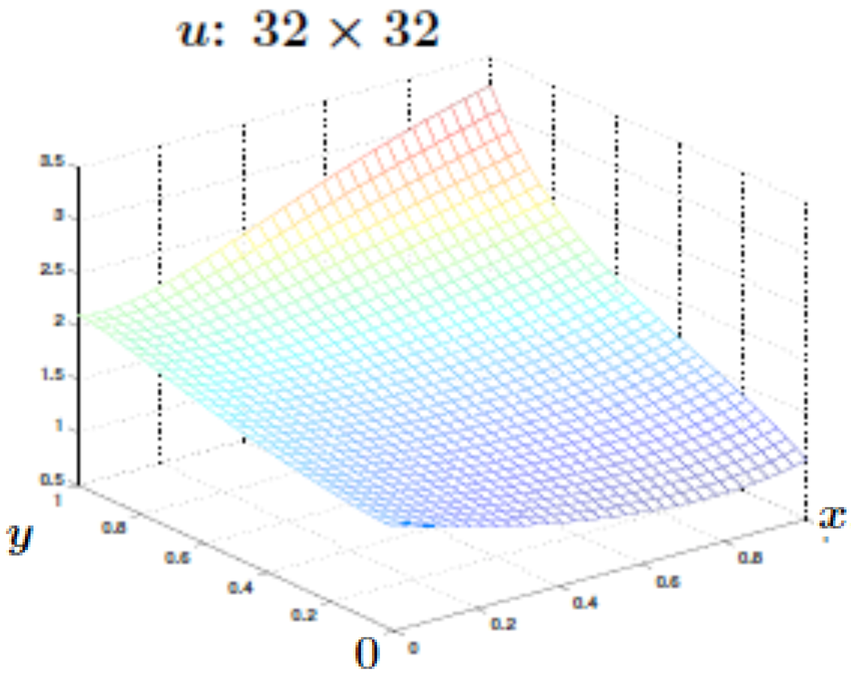
Approx. sol. v: 4×4

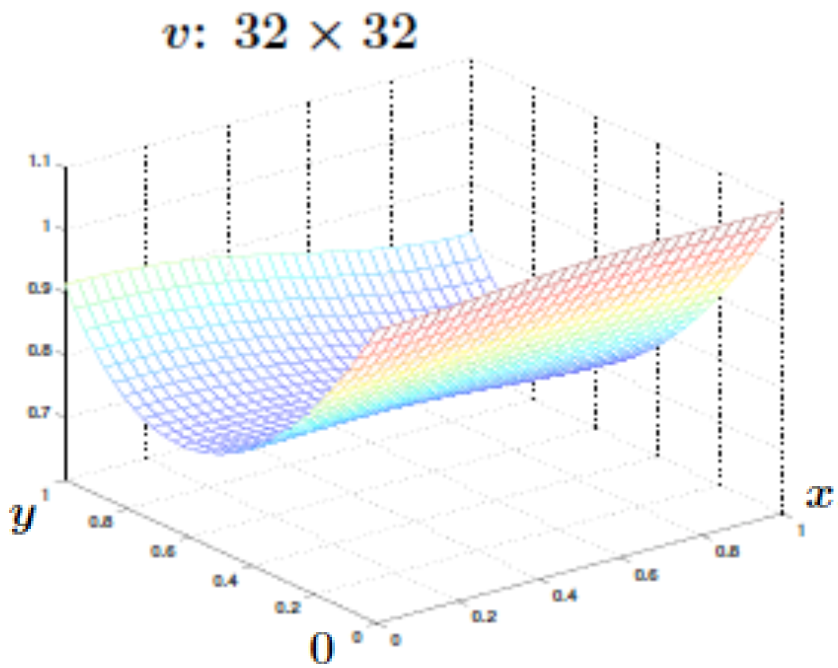


Example 4 (2 unknown case on $[0,1]\times[0,1]$, Neumann condition):

$$u_{xx}(x,y)+u_{yy}(x,y)+u(x,y)(1-u(x,y)^2-v(x,y)^2)=0,$$
$$v_{xx}(x,y)+v_{yy}(x,y)+v(x,y)(1-u(x,y)^2-v(x,y)^2)=0,$$
$$u_x(0,y)=-1,\ u_x(1,y)=1,\ \forall y\in[0,1],$$
$$u_y(x,0)=2x,\ u_y(x,1)=x+5\sin(\pi x/2),\ \forall x\in[0,1],$$
$$v_x(0,y)=0,\ v_x(1,y)=0,\ \forall y\in[0,1],$$
$$v_y(x,0)=-1,\ v_y(x,1)=1,\ \forall x\in[0,1].$$

(Sparse SDP relaxation + 2.b (Newton Method))





Contents

1. How do we formulate structured sparsity?
   1-1. Unconstrained cases.
   1-2. Constrained and linear objective function cases.

2. Sparse SDP relaxation of constrained POPs.

3. Applications to PDEs (partial differential equations).

4. Concluding remarks.

Some difficulties in SDP relaxation of POPs

(a) Sparse SDP relaxation problems of a POP are sometimes difficult to solve accurately (by the primal-dual interior-point method).

(b) The efficiency of the (sparse) SDP relaxation of a POP depends on the relaxation order r which is required to get an accurate optimal solution but is unknown in advance.

A difficulty in application of the sparse SDP relaxation to PDEs

(c) A polynomial system induced from a PDE is not c-sparse enough to process finer grid discretization.

$\Downarrow$

- More powerful and stable software to solve SDPs.
- Some additional techniques.

Thank you!