Shape from Moments
An Estimation Perspective

Michael Elad*, Peyman Milanfar, and Gene Golub***

SIAM 2002 Meeting

MS104 - Linear Algebra in Image Processing

July 12th, 2002

* The CS Department – SCCM Program Stanford University

** Peyman Milanfar is with the University of California Santa Cruz (UCSC).

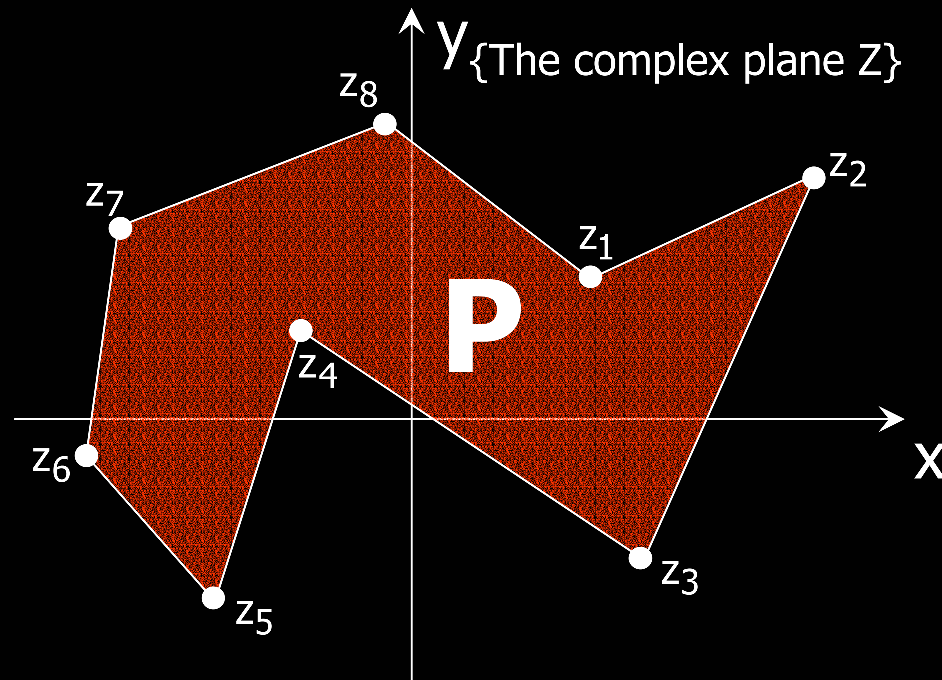
Chapter A

Background

A.1 Davis Theorem

***Theorem (Davis 1977)*:**

For any closed 2D polygon, and for any analytic function f(z) the following holds



$$\iint_P f''(z)\,dx\,dy = \sum_{n=1}^{N} \frac{i}{2}\left[\frac{\bar{z}_{n-1}-\bar{z}_n}{z_{n-1}-z_n} + \frac{\bar{z}_{n+1}-\bar{z}_n}{z_{n+1}-z_n}\right] f(z_n),$$

$= a_n$

A.2 Complex Moments

If we use the analytic function $f(z) = z^k$, we get from Davis Theorem that

$$\iint_P f''(z)\,dxdy = k(k-1)\iint_P z^{k-2}\,dxdy = \sum_{n=1}^{N} a_n z_n^k = \tau_k$$

$$= \mu_{k-2}$$

where we define

$$\tau_k = k(k-1)\iint_P z^{k-2}\,dxdy = k(k-1)\mu_{k-2}$$

A.3 Shape From Moments

$$\left\{ \tau_k = \sum_{n=1}^{N} \frac{i}{2} \left[\frac{\bar{z}_{n-1} - \bar{z}_n}{z_{n-1} - z_n} + \frac{\bar{z}_{n+1} - \bar{z}_n}{z_{n+1} - z_n} \right] z_n^k \right\}_{k=0}^{M}$$

- Can we compute the vertices from these equations ?
- How many moments are required for exact recovery ?

A.4 Previous Results

- Milanfar et. al. (1995):
  - (2N-1) moments are theoretically sufficient for computing the N vertices.
  - Prony's method is proposed.
- Golub et. al. (1999):
  - Pencil method replacing the Prony's - better numerical stability.
  - Sensitivity analysis.
- Prony's and the Pencil approaches:
  - Rely strongly on the linear algebra formulation of the problem.
  - Both are sensitive to perturbations in the moments.
  - Both will be presented briefly.

A.5 To Recap

Known values

$$\begin{bmatrix} \tau_0 \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_M \end{bmatrix} = \begin{bmatrix} z_1^0 & z_2^0 & \cdots & z_N^0 \\ z_1^1 & z_2^1 & \cdots & z_N^1 \\ z_1^2 & z_2^2 & \cdots & z_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^M & z_2^M & \cdots & z_N^M \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}$$

Function of the vertices (and their order)

Estimate $\{z_n\}_{n=1}^N$!!!

A.6 Our Focus

- **Noisy measurements:** What if the moments are contaminated by additive noise ? How can re-pose our problem as an estimation task and solve it using traditional stochastic estimation tools ?

- **More measurements:** What if there are M>2N-1 moments ? How can we exploit them to robustify the computation of the vertices ?

A.7 Related Problems

- It appears that there are several very different applications where the same formulation is obtained
  - Identifying an auto-regressive system from its output,
  - Decomposing of a linear mixture of complex cissoids,
  - Estimating the Direction Of Arrival (DOA) in array processing,
  - and more ...
- Major difference – $\{a_k\}_{k=1}^{N}$ are not functions of the unknowns but rather free parameters.
- Nevertheless, existing algorithms can be of use.

Chapter B

Prony and Pencil Based Methods

B.1 Prony's Relation

$$\left\{ \tau_k = \sum_{n=1}^{N} a_n z_n^k \right\}_{k=0}^{M} \Rightarrow P(z) = \prod_{n=1}^{N} (z - z_n) = z^N + \sum_{n=1}^{N} p_n z^{N-n}$$

$$-\begin{bmatrix} \tau_N \\ \tau_{N+1} \\ \tau_{N+2} \\ \vdots \\ \tau_M \end{bmatrix} = \begin{bmatrix} \tau_0 & \tau_1 & \cdots & \tau_{N-1} \\ \tau_1 & \tau_2 & \cdots & \tau_N \\ \tau_2 & \tau_3 & \cdots & \tau_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{M-N} & \tau_{M-N+1} & \cdots & \tau_{M-1} \end{bmatrix} \begin{bmatrix} p_N \\ p_{N-1} \\ \vdots \\ p_1 \end{bmatrix}$$

B.2 Prony's Methods

$$\begin{bmatrix} \tau_0 & \tau_1 & \cdots & \tau_{N-1} \\ \tau_1 & \tau_2 & \cdots & \tau_N \\ \tau_2 & \tau_3 & \cdots & \tau_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{M-N} & \tau_{M-N+1} & \cdots & \tau_{M-1} \end{bmatrix} \begin{bmatrix} p_N \\ p_{N-1} \\ \vdots \\ p_1 \end{bmatrix} = - \begin{bmatrix} \tau_N \\ \tau_{N+1} \\ \tau_{N+2} \\ \vdots \\ \tau_M \end{bmatrix}$$

a. Regular Least-Squares, followed by root-finding,

b. Total-Least-Squares, followed by root-finding,

c. Hankel Constrained SVD, followed by root-finding,

d. IQML, Structuted-TLS, Modified Prony, and more.

B.3 Pencil Relation

Define

$$T_0 = \begin{bmatrix} \tau_0 & \tau_1 & \cdots & \tau_{N-1} \\ \tau_1 & \tau_2 & \cdots & \tau_N \\ \tau_2 & \tau_3 & \cdots & \tau_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{M-N} & \tau_{M-N+1} & \cdots & \tau_{M-1} \end{bmatrix} \qquad T_1 = \begin{bmatrix} \tau_1 & \tau_2 & \cdots & \tau_N \\ \tau_2 & \tau_3 & \cdots & \tau_{N+1} \\ \tau_3 & \tau_4 & \cdots & \tau_{N+2} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{M-N+1} & \tau_{M-N+2} & \cdots & \tau_M \end{bmatrix}$$

After some (non-trivial) manipulation we obtain

$$\left[T_1 - z_n \cdot T_0\right] \underline{V}_n = 0$$

For some non zero vectors $\underline{V}_n$.

B.4 Non-Square Pencil

$$\left(T_1 - z_n T_0 \right) \underline{v}_n = 0$$

B.5 Pencil Methods

$$\left[T_1 - z_n \cdot T_0\right]\underline{v}_n = 0$$

a. Take square portions, solve for the eigenvalues, and cluster the results,

b. Square by left multiplication with T_0^H (closely related to LS-Prony),

c. Hua-Sarkar approach: different squaring methods which is more robust and related to ESPRIT.

Chapter C

ML and MAP Approaches

C.1 What are we Missing ?

- We have seen a set of simple methods that give reasonable yet inaccurate results.
- In our specific problem we do not exploit the fact that $\{a_k\}_{k=1}^{N}$ are vertices-dependent.
- In all the existing methods there is no mechanism for introducing prior-knowledge about the unknowns.

C.2 Recall ...

We have the following system of equations

$$\underline{t} = \begin{bmatrix} \tau_0 \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_M \end{bmatrix} = \begin{bmatrix} z_1^0 & z_2^0 & \cdots & z_N^0 \\ z_1^1 & z_2^1 & \cdots & z_N^1 \\ z_1^2 & z_2^2 & \cdots & z_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^M & z_2^M & \cdots & z_N^M \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = \mathbf{V}(\underline{z}) \cdot \underline{a}(\underline{z})$$

Measured

Function of the unknowns

C.3 Our Suggestion

- If we assume that the moments are contaminated by zero-mean white Gaussian noise, Direct-Maximum-Likelihood (DML) solution is given by

$$\underset{\underline{z}}{\text{Minimize}} \quad \left\| \underline{t} - \mathbf{V}(\underline{z}) \cdot \underline{a}(\underline{z}) \right\|_2^2$$

- Direct minimization is hard to workout, BUT
- We can use one of the above methods to obtain an initial solution, and then iterate to minimize the above function until getting to a local minima.

C.4 Things to Consider

- Even (complex) coordinate descent with effective line-search can be useful and successful (in order to avoid derivatives).
- Per each candidate solution we HAVE TO solve the ordering problem !!!! Treatment of this problem is discussed in Durocher (2001).
- If the initial guess is relatively good, the ordering problem becomes easier, and the chances of the algorithm to yield improvement are increased.

C.5 Relation to VarPro

- VarPro (Golub & Pereyra 1973)
  - Proposed for minimizing $\left\| \underline{t} - \mathbf{V}(\underline{z}) \cdot \underline{a} \right\|_2^2$
  - The basic idea: Represents the $\underline{a}$ as $\underline{a} = \mathbf{V}^{+}(\underline{z})\underline{t}$ and use derivatives of the Pseudo-Inverse matrix.
- Later work (1978) by Kaufman and Pereyra covered the case where $\underline{a}=\underline{a}(\underline{z})$ (linear constraints).
- We propose to exploit this or similar method, and choose a good initial solution for our iterative procedure.

C.6 Regularization

- Since we are minimizing (numerically) the DML function, we can add a regularization – a penalty term for directing the solution towards desired properties.

$$\underset{\underline{z}}{\text{Minimize}} \quad \left\| \underline{t} - \mathbf{V}(\underline{z}) \cdot \underline{a}(\underline{z}) \right\|_2^2 + g\{\underline{z}\}$$

- The minimization process is just as easy.
- This concept is actually an application of the Maximum A-posteriori-Probability (MAP) estimator.

C.7 MAP Possibilities

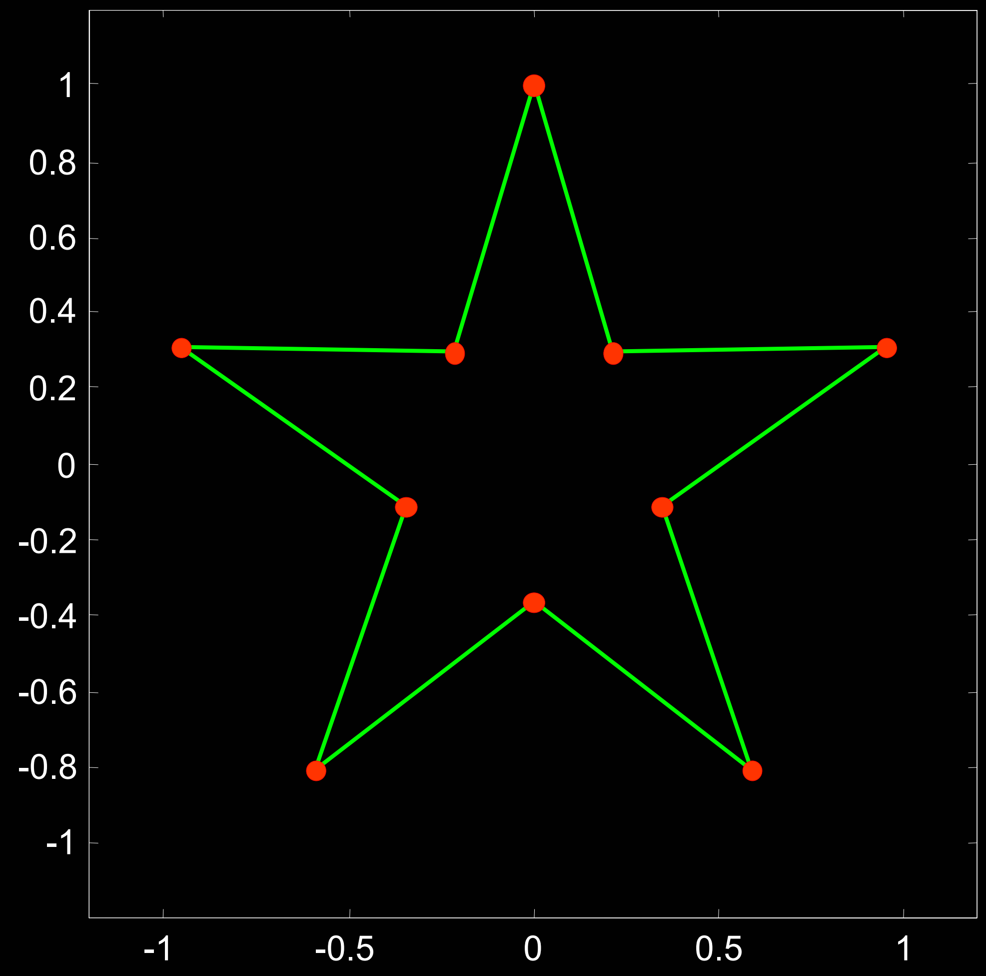
Kind of Prior	Expression for g(*)
Promoting 90° angles	$\sum_{n=1}^{N}\left(\frac{1}{2}\left\lvert\frac{\bar{z}_{n-1}-\bar{z}_n}{z_{n-1}-z_n}+\frac{\bar{z}_{n+1}-\bar{z}_n}{z_{n+1}-z_n}\right\rvert-1\right)^2$
Promoting smoothness (1)	$\sum_{n=1}^{N}\lvert z_{n-1}-z_n\rvert^2$
Promoting smoothness (2)	$\sum_{n=1}^{N}\lvert z_{n-1}+z_{n+1}-2z_n\rvert^2$
Promoting regularity	$\left(\sum_{n=1}^{N}\lvert z_{n-1}-z_n\rvert\right)^2 \Big/ \mathrm{Im}\left(\sum_{n=1}^{N} z_{n+1}\bar{z}_n\right)$
Promoting small area	$\mathrm{Im}\left(\sum_{n=1}^{N} z_{n+1}\bar{z}_n\right)$

Chapter D

Results

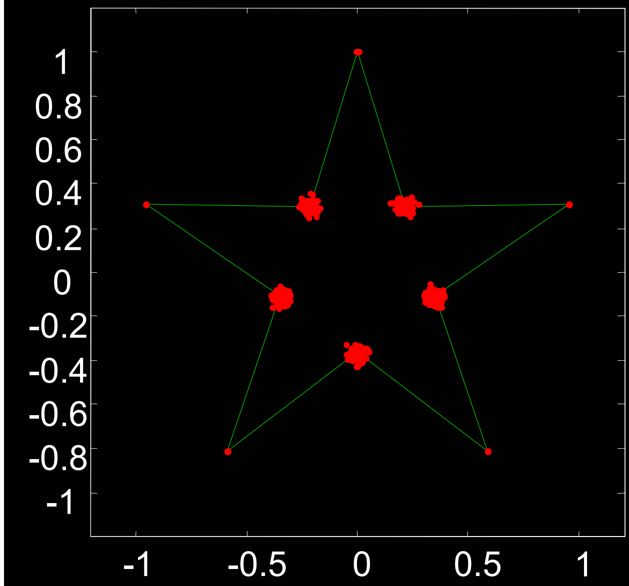
D.1 Experiment #1

- Compose the following star-shaped polygon (N=10 vertices),
- Compute its exact moments (M=100),
- add noise (σ=1e-4),
- Estimate the vertices using various methods.



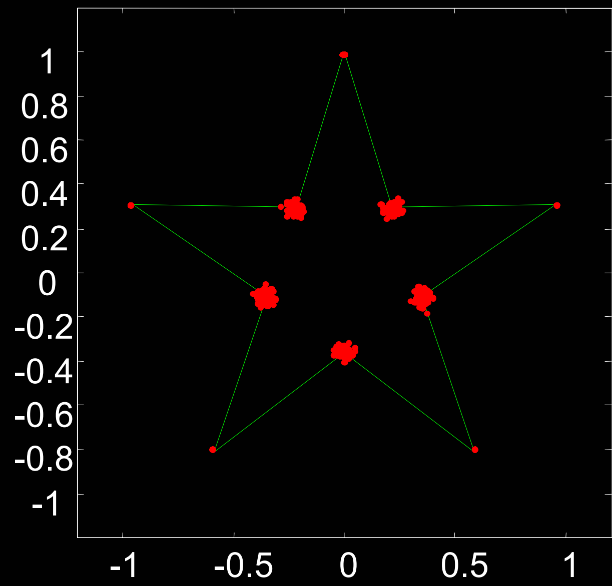
D.1 Experiment #1

Mean Squared Error averaged over 100 trials



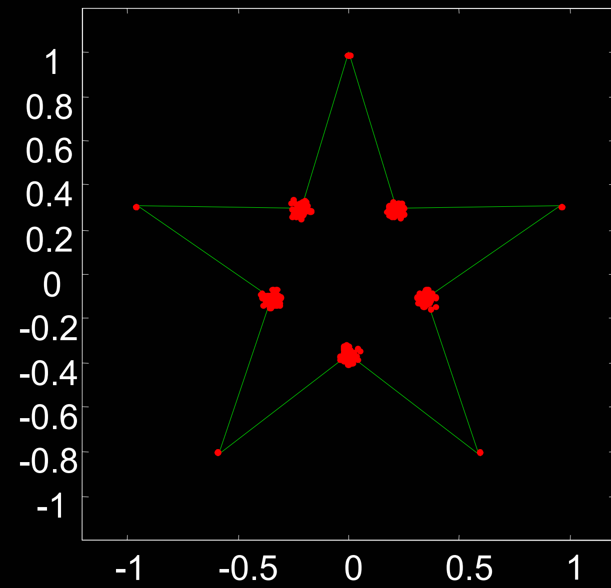
LS-Prony method

0.0201



Squared Pencil method

0.0196

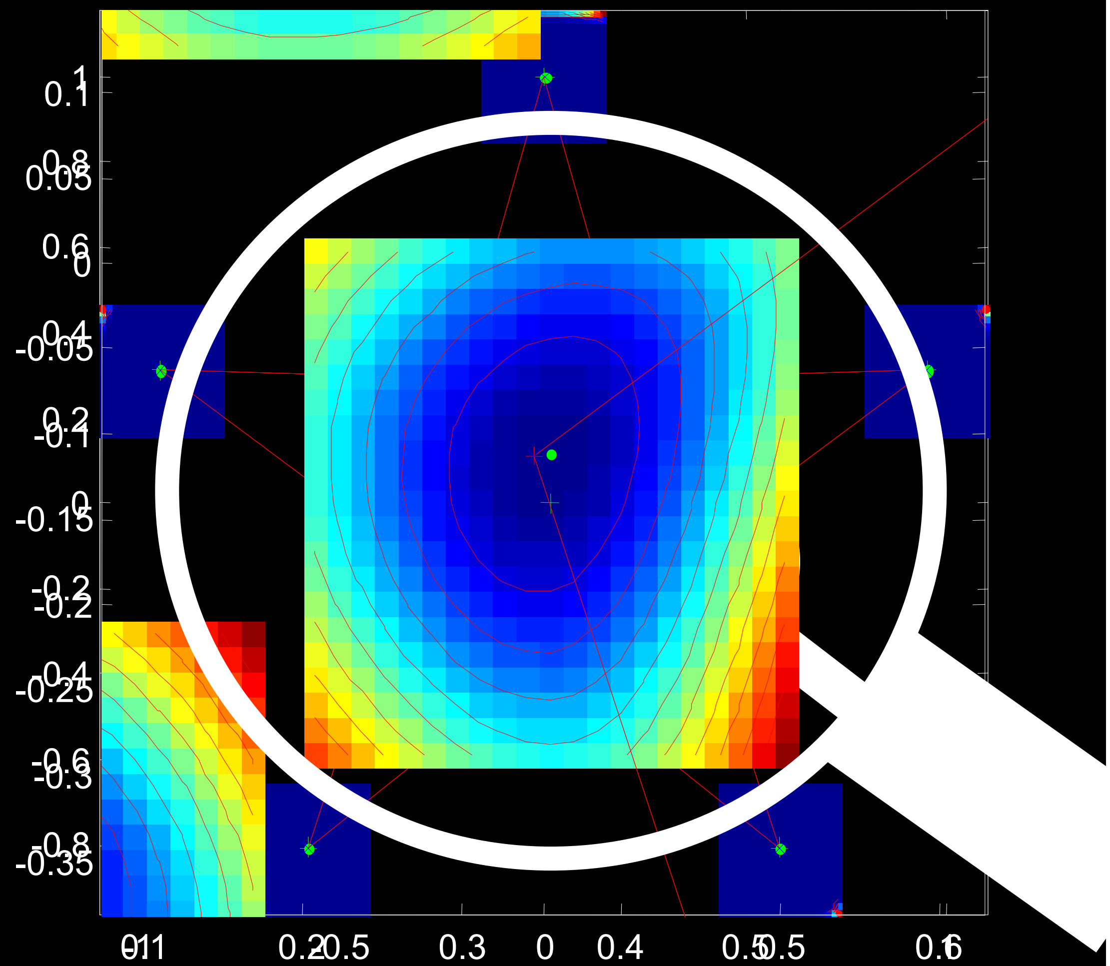


Hua-Sarkar method

0.0174

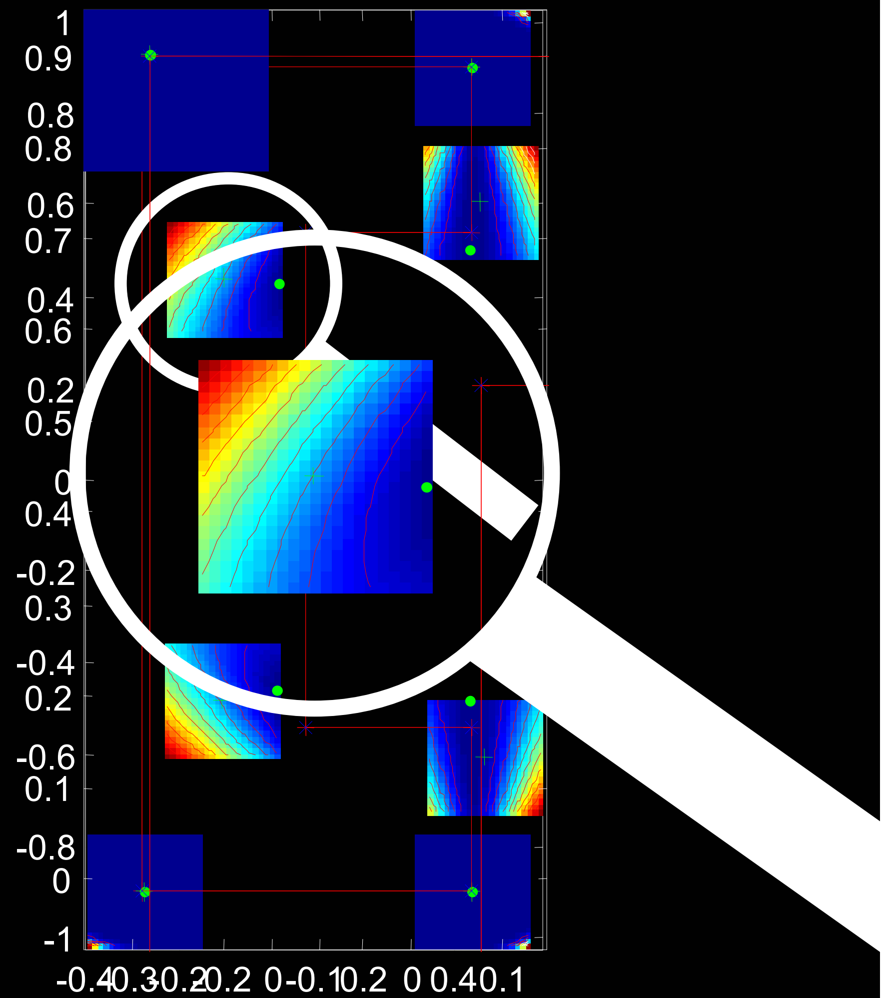
D.2 Experiment #2

- For the star-shape polygon with noise variance σ=1e-4, initialize using Hua-Sarkar algorithm.
- Then, show the DML function per each vertex, assuming all other vertices fixed.
  - + Hua-Sarkar result
  - ● New local minimum



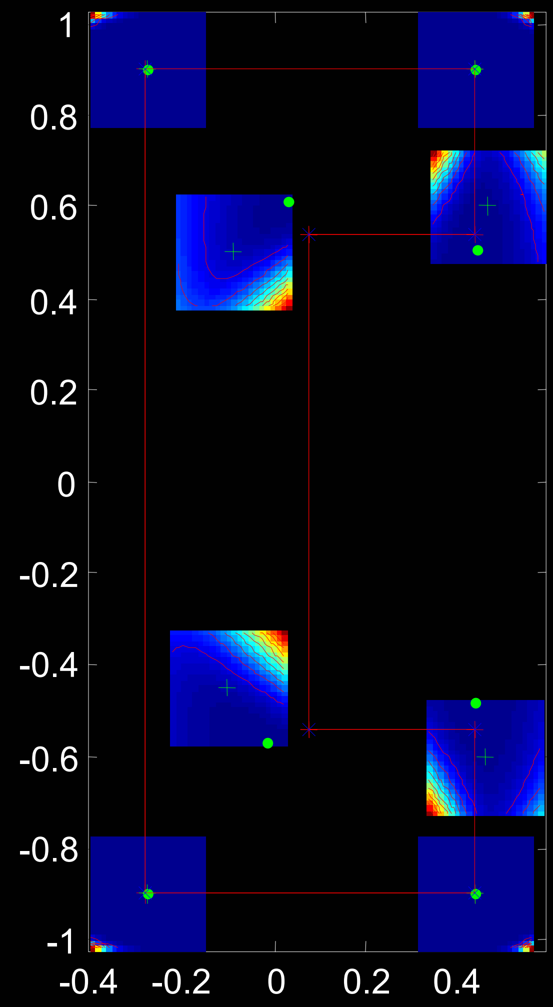
D.3 Experiment #3

- For the E-shape polygon with noise variance σ=1e-3, initialize using LS-Prony algorithm.
- Then, show the DML function per each vertex, assuming all other vertices fixed.
  - + LS-Prony result
  - ● New local minimum



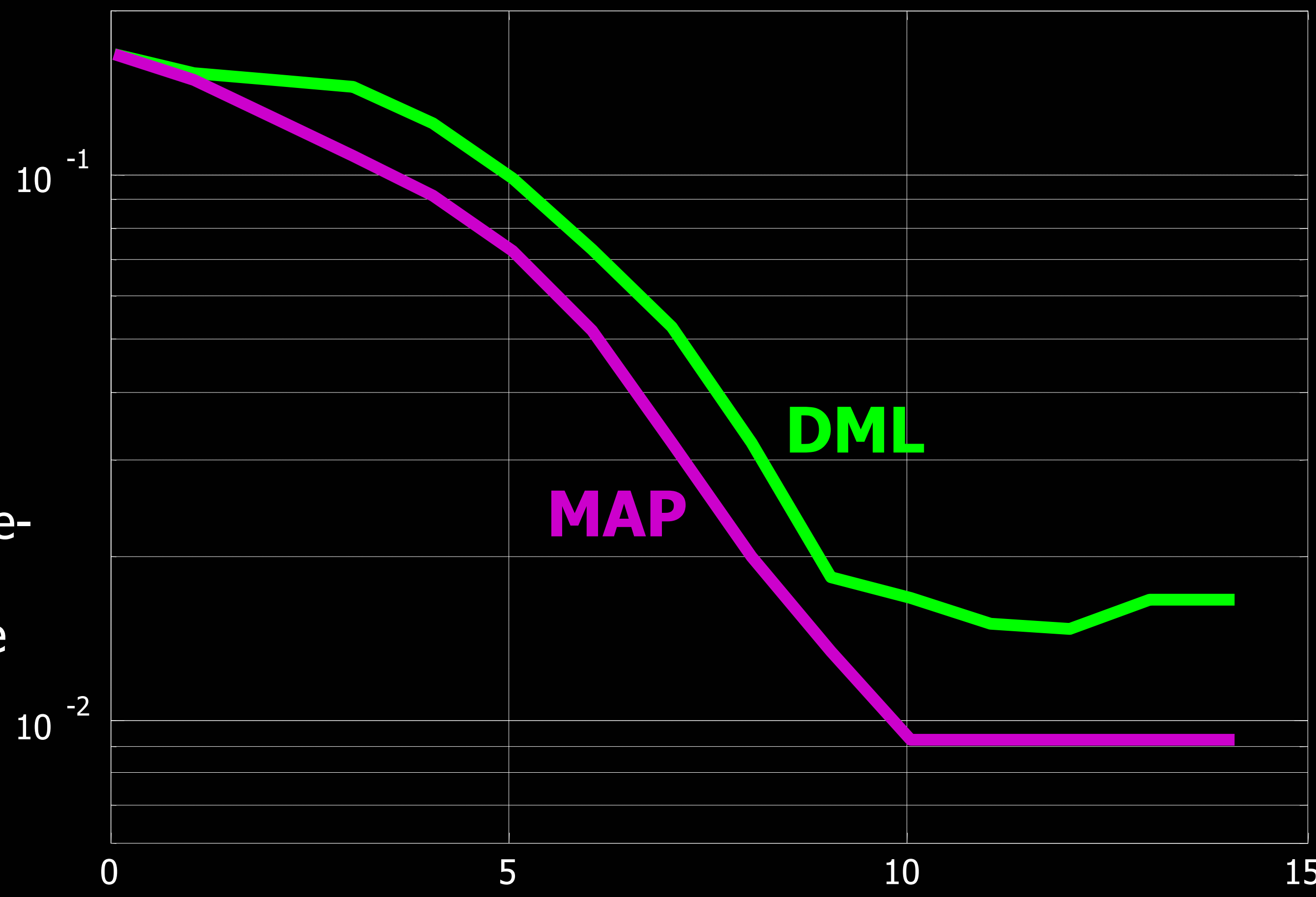
D.4 Experiment #4

- For the E-shape polygon with noise variance σ=1e-3, initialize using LS-Prony algorithm.
- Then, show the MAP function per each vertex, assuming all other vertices fixed.
- Regularization – promote 90° angles.
  - + LS-Prony result
  - ● New local minimum



D.5 Experiment #5

Error as a function of the iteration number*



* Using a derivative free coordinate-descent procedure

D.6 To Conclude

- The shape-from-moments problem is formulated, showing a close resemblance to other problems in array processing, signal processing, and antenna theory.
- The existing literature offers many algorithms for estimating the “vertices” – some of them are relatively simple but also quite sensitive.
- In this work we propose methods to use these simple algorithms as initialization, followed by a refining stage based on the Direct Maximum Likelihood and the MAP estimator.