Total Variation Regularization of the Radiographic Inversion Problem: Follow-up Work to the SVD Number of Views Study

Kevin R. Vixie, Thomas J. Asaki, Patrick R. Campbell,
and David M. Strong

Los Alamos National Laboratory
and Pepperdine University

A Brief Outline:

Images: The Importance of Edges

Images and Edges: Key Features

- An example

- Edges contain the shape information
- Location and Magnitude important

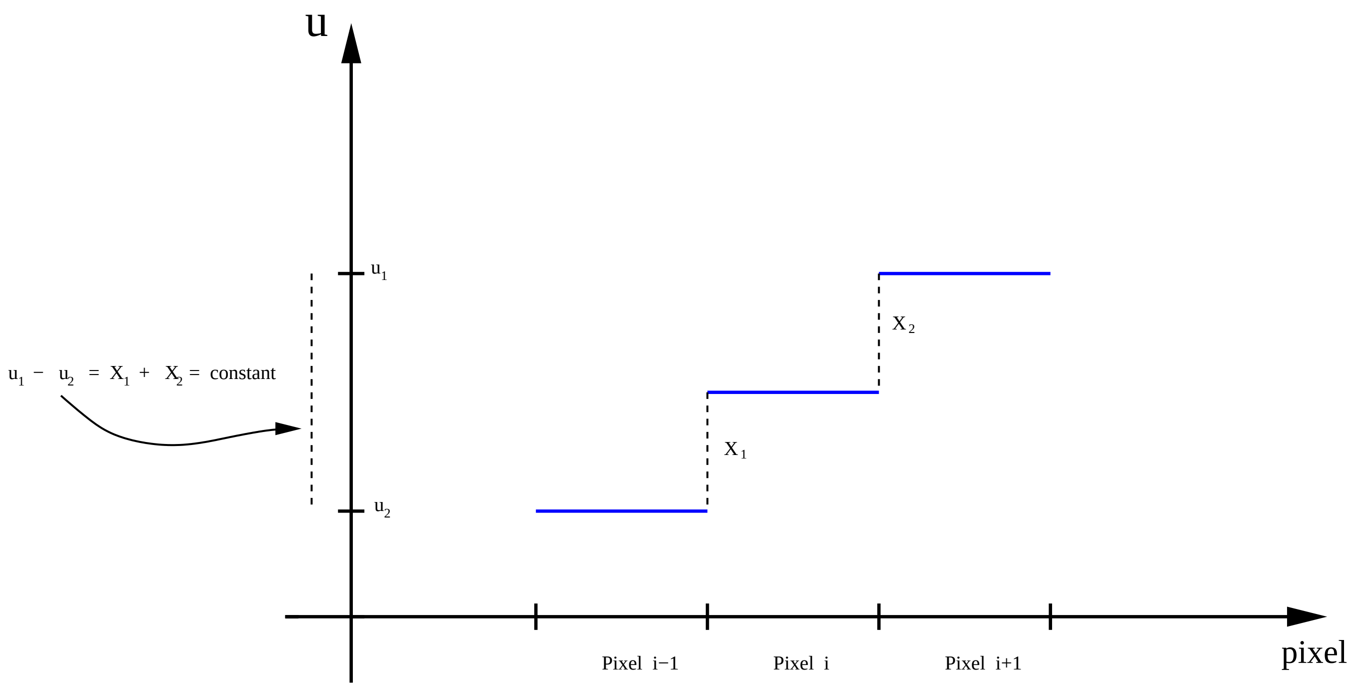
It is therefore natural to consider measures of images $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ of the form:

$$F(u) \equiv \int_\Omega \phi(|\nabla u|) d\mu$$

We will look at $\phi(x) = |x|^p \ p \in (0, 2]$.

(In images $F(u) = \int_\Omega |\nabla u|^p d\mu = \sum |\text{pixel differences}|^p$.)

Let's look at the very simple discrete picture



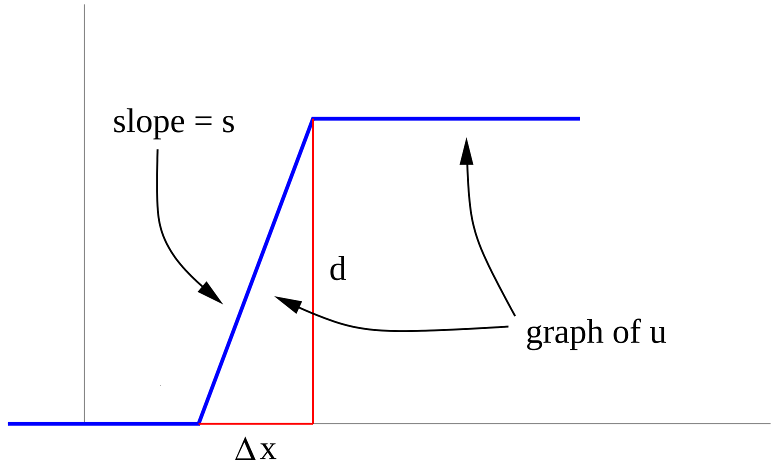
Suppose we want to find the "picture" with minimum $F(u)$ comprised of the three pixels above with the constraint that $u_2 - u_1 =$ a constant, call it $\hat{U}$.

- cost of jump $x_j = |x_j|^p$
- define cost per length, C_L to be $|x_j|^p / |x_j|$
- we want to minimize

$$
\begin{aligned}
|x_1|^p + |x_2|^p &= |x_1| * C_L(x_1) + |x_2| * C_L(x_2) \\
&= \hat{U} * \left(\frac{|x_1|}{\hat{U}} * C_L(x_1) + \frac{|x_2|}{\hat{U}} * C_L(x_2) \right) \\
&= \hat{U} * (\text{average cost per length})
\end{aligned}
$$

We consider the (1-dimensional) continuous case:

we compute $F(u) \equiv \int_\Omega |\nabla u|^p dx$



$$\textbf{(}p > 1\textbf{)} \quad F(u) = s^p(\Delta x) = \frac{(s\Delta x)^p}{(\Delta x)^{p-1}} = \frac{d^p}{(\Delta x)^{p-1}} \underset{\Delta x \to 0}{\longrightarrow} \infty$$

$$\textbf{(}p = 1\textbf{)} \quad F(u) = s\Delta x = d$$

$$\textbf{(}p < 1\textbf{)} \quad F(u) = (s\Delta x)^p(\Delta x)^{1-p} = d^p(\Delta x)^{1-p} \underset{\Delta x \to 0}{\longrightarrow} 0$$

Moral of the Story:

- For $p > 1$ discontinuities are avoided ... smooth u preferred,
- For $p < 1$ discontinuities cost nothing ... step u preferred,
- BUT for $p = 1$ only the variation or jump magnitude "counts", no bias towards either smooth or step!

Aside: Lagrange Multiplier Picture

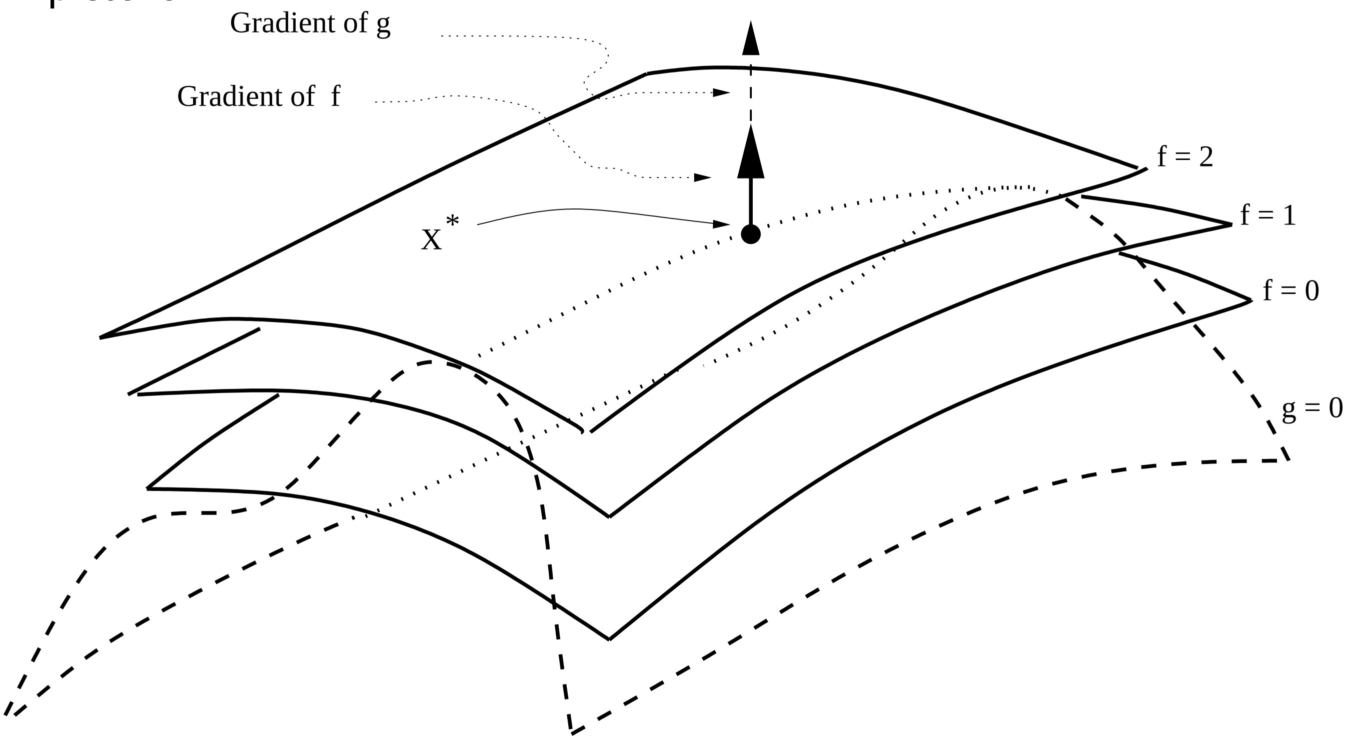
Problem Maximize $f(x)$ subject to $g(x) = c$.

Multiplier method Find stationary points of

$$L(x, \lambda) \equiv f(x) + \lambda g(x)$$

Do ... solve $DL(x, \lambda) = Df(x) + \lambda Dg(x) = 0$

A picture:



Aside: BV Functions

*W*e can define a space of functions whose norm is based on the measure we introduced to look at edges, $F(f) \equiv \int_\Omega |\nabla f| d\mu$.

The space will be those $f : \mathbb{R}^N \to \mathbb{R}$ such that

$$f \in L^1(\Omega) \text{ and } \int_\Omega |Df| < \infty$$

where we define $\int_\Omega |Df|$ when $f \notin C^1(\Omega)$ by

$$\int_\Omega |Df| \equiv \sup \left\{ \int_\Omega f \mathrm{div}(g) dx : g \in C_c^1(\Omega, \mathbb{R}^N), |g(x)| \leq 1 \ \ \forall x \in \Omega \right\}$$

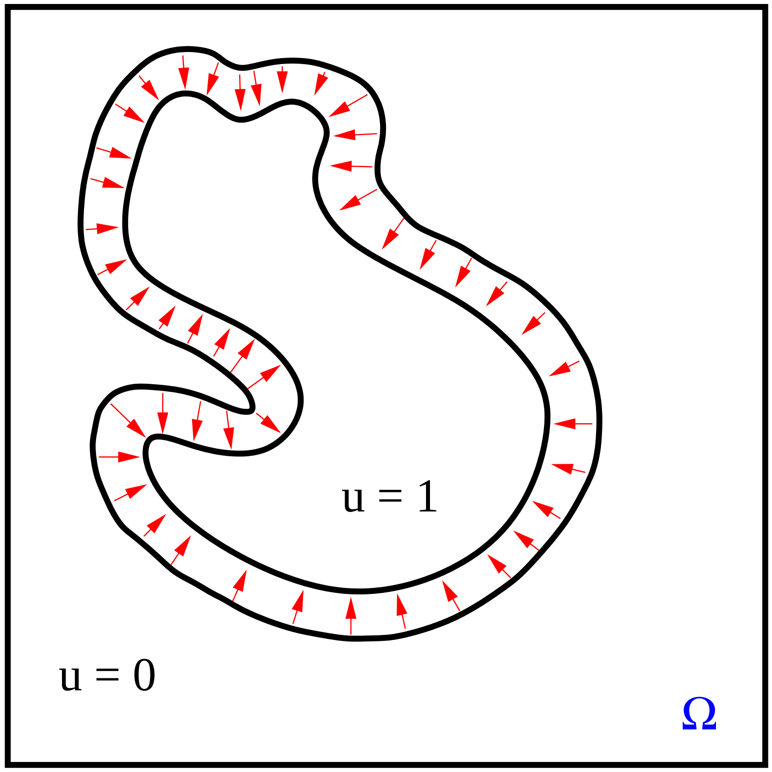
In this case we define the BV norm by:

$$||f||_{BV} \equiv ||f||_{L^1} + \int_\Omega |Df|$$

*T*he space of functions f such that $||f||_{BV} < \infty$ together with the BV norm is a Banach space of great usefulness and versatility.

$TV(f) \equiv \int_\Omega |Df|$ $(= \int_\Omega |\nabla f|$ when $f \in C^1)$ is the *Total Variation* of f.

... a bit more: level sets and feature scale dependence



$TV(u)$ can be also computed as an integral over level sets:

$$TV(u) = \int_{\Omega} |Du| dxdy = \int_{\mathbb{R}} L(u, r) dr$$

where

$L(u, r) =$ length of the boundary of the r-level set.

Total Variation Regularization

Now we consider the image recovery problem and the role that *optimization* *and* *total variation* *play in regularized reconstruction from projections.*

A common regularization of the radiographic image reconstruction problem is the use of the L^2 norm to regularize the inverse problem.

$$u_{\text{optimal}} \equiv \arg\min_u F(u) = \arg\min_u \left(||Pu - d||^2 + ||u||_2 \right)$$

where P is the radiographic projection operator and d is the radiographic data. Another regularization is given by the minimization:

$$u_{\text{optimal}} \equiv \arg\min_u F(u) = \arg\min_u \left(||Pu - d||^2 + \int_\Omega |\nabla u|^2 \right)$$

But, as noted above, the $|\nabla u|^2$ is biased against edges, while $|\nabla u|$ is biased neither for or against edges. This leads us to consideration of:

$$u_{\text{optimal}} \equiv \arg\min_u F(u) = \arg\min_u \left(||Pu - d||^2 + \int_\Omega |\nabla u| \right)$$

Looking at the regularization more carefully:

noise $||Pu - d||^2 = \sigma^2$

optimization We use Lagrange Multipliers

Discretization We work on discrete images

So we end up with the following continuous functional and it's discrete counterpart:

$$F(u) = \lambda ||Pu - d||^2 + \int_\Omega |\nabla u|$$

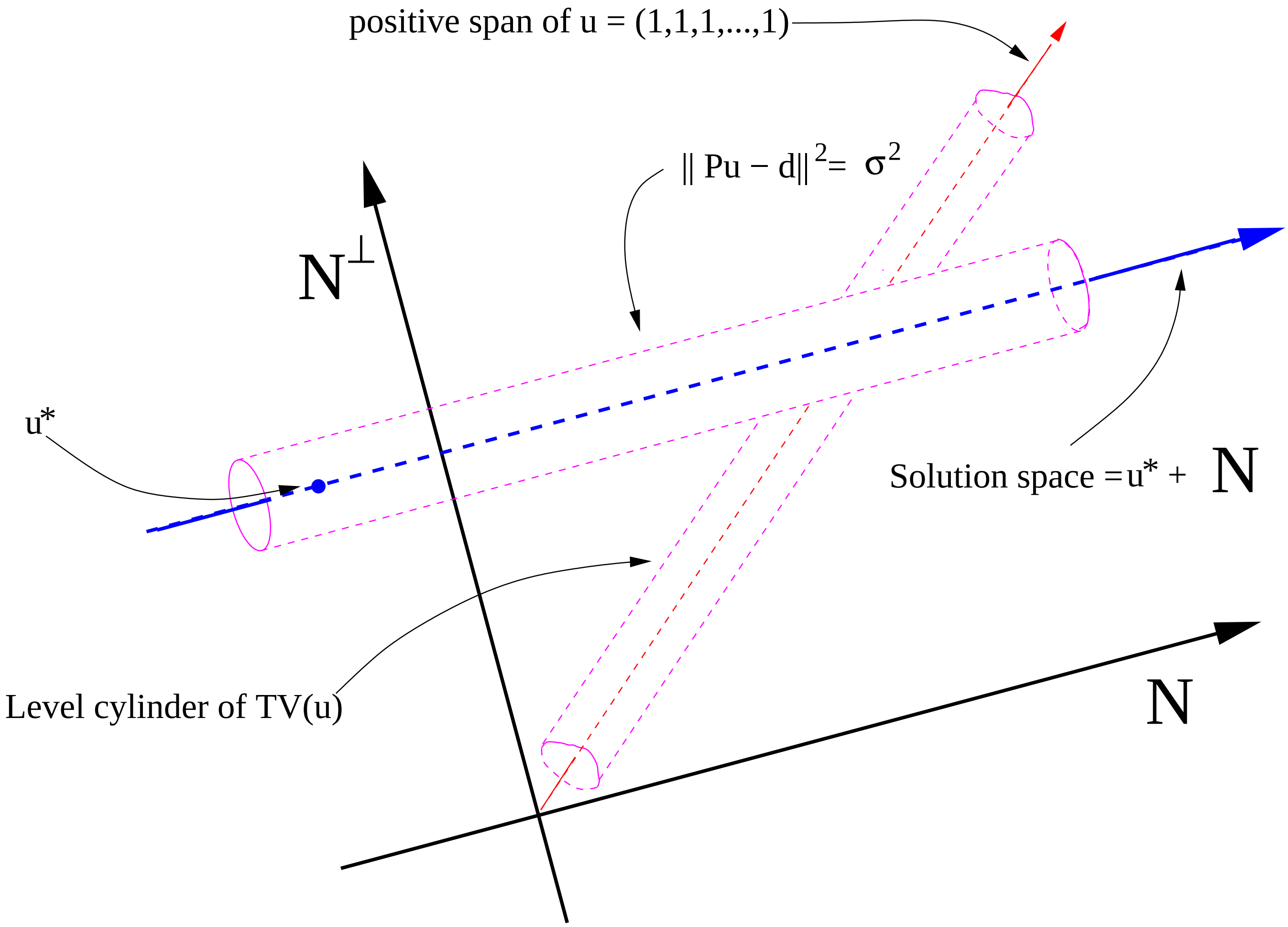
$$\hat{F}(u) = \lambda ||Pu - d||^2 + \sum_{i,j} |\nabla_{i,j} u|$$

where $\nabla_{i,j}$ is the discretized gradient. For specific choices of λ we implicitly seek solutions constrained by $||Pu - d||^2 = \sigma_\lambda^2$).

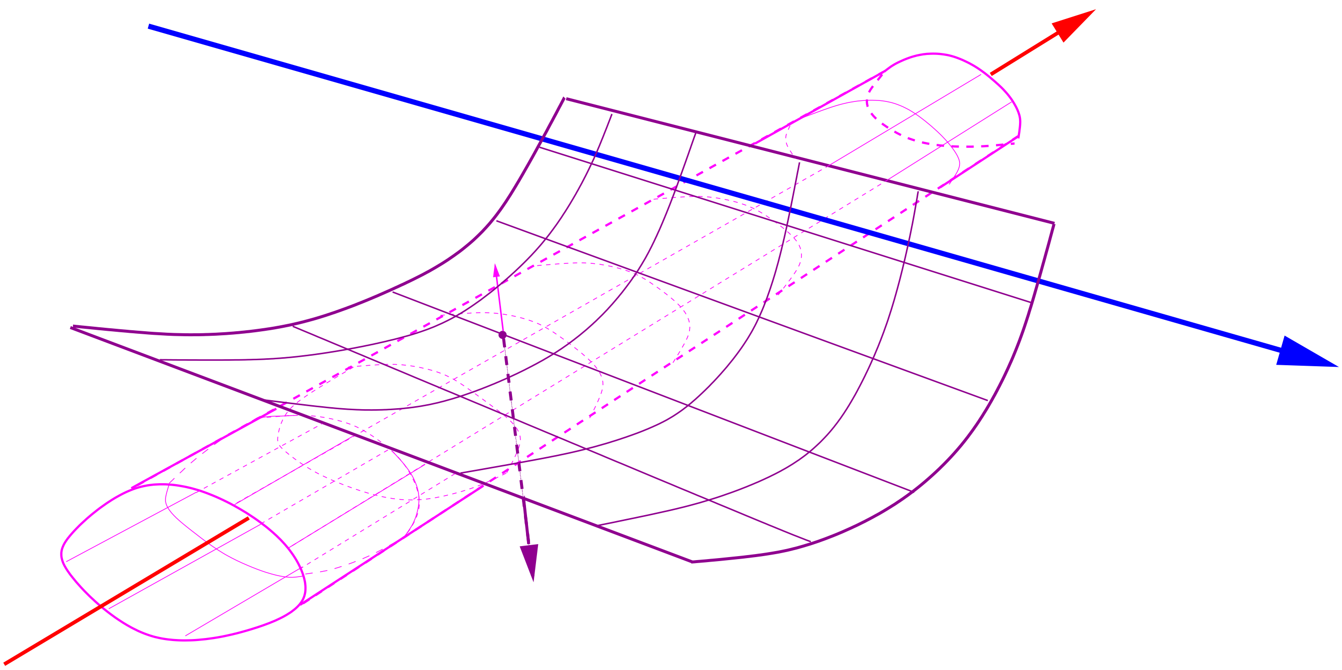
*T*echnically speaking, we need the discretized functional to be strictly convex and coercive for guaranteed existence and uniqueness. It also permits us to prove linear convergence.

A Picture: two cylinders

Recalling our pictures of the Lagrange Multipliers and specializing them to our functional $\hat{F}(u)$:



... a closer look ...



A Fixed point method

If we compute the derivative of the functional

$$F(u) = \alpha \sum_{i,j} |\nabla_{i,j} u| + \frac{1}{2} ||Pu - d||^2$$

(where $\alpha = \frac{2}{\lambda}$) and set it to zero, we get

$$\alpha \sum_{i,j} \nabla_{i,j}^T \left(\frac{\nabla_{i,j} u}{|\nabla_{i,j} u|} \right) + P^T P u - P^T d = 0$$

We turn this into an iterative method that (we can prove) converges to a unique fixed point. The iterative method is given by

$$\alpha \sum_{i,j} \nabla_{i,j}^T \left(\frac{\nabla_{i,j} u_{k+1}}{|\nabla_{i,j} u_k|} \right) + P^T P u_{k+1} - P^T d = 0$$

At each step we solve for u_{k+1} using a conjugate gradient method. A last modification to remove the singularity in derivative of the TV term is done by noticing that

$$|\nabla u| = \sqrt{|\nabla u|^2} \approx \sqrt{|\nabla u|^2 + \beta}$$

for small β.

Convergence results and rates

In Chan and Mulet's paper on the convergence of the fixed point method they prove that

$$|F(u_{k+1}) - F(u^*)| \leq \gamma |F(u_k) - F(u^*)|$$

where

$$0 \leq \gamma < 1.$$

Furthermore the u_k's converge with a linearly convergence rate of at most $\sqrt{\gamma}$.

While small β's tend to slow the convergence, it appears from numerical experiments that $\gamma \to 1 - \epsilon$ for $\epsilon > 0$ as $\beta \to 0$.

Aside: Other Methods

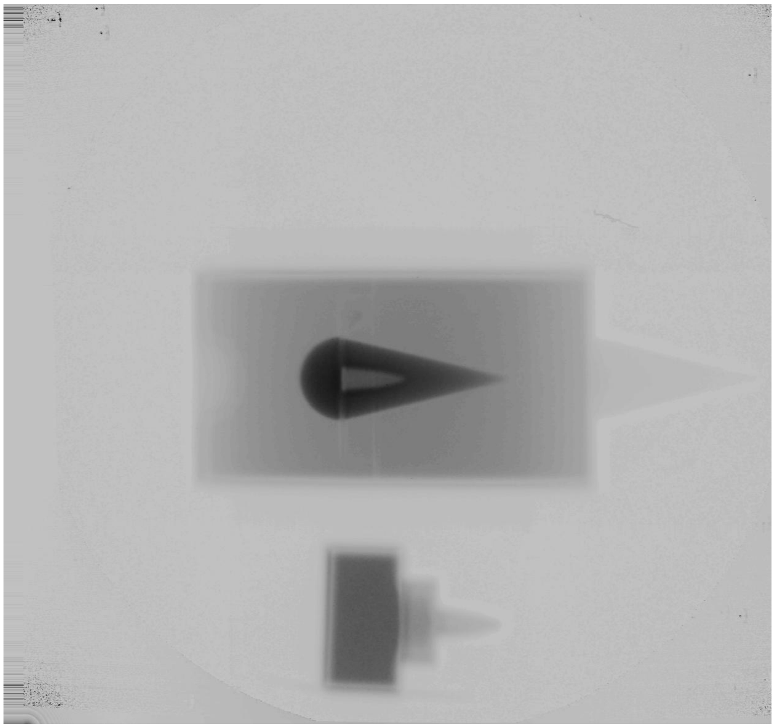
Time Marching This is the method used by Rudin, Osher, and Fatemi. Essentially one simply evolves u by treating the gradient of F as the time right hand side of a differential equation. In other words one solves

$$u_t = -DF(u).$$

Primal-Dual In this method one uses the Fenchel Transform together with results from variational theory to reformulate the problem. One can obtain much better convergence results.

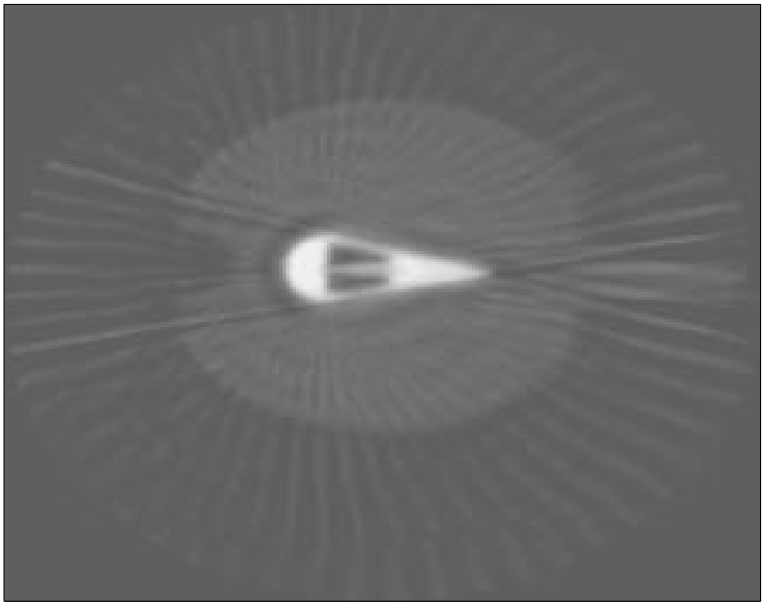
Initial Results: BCO4

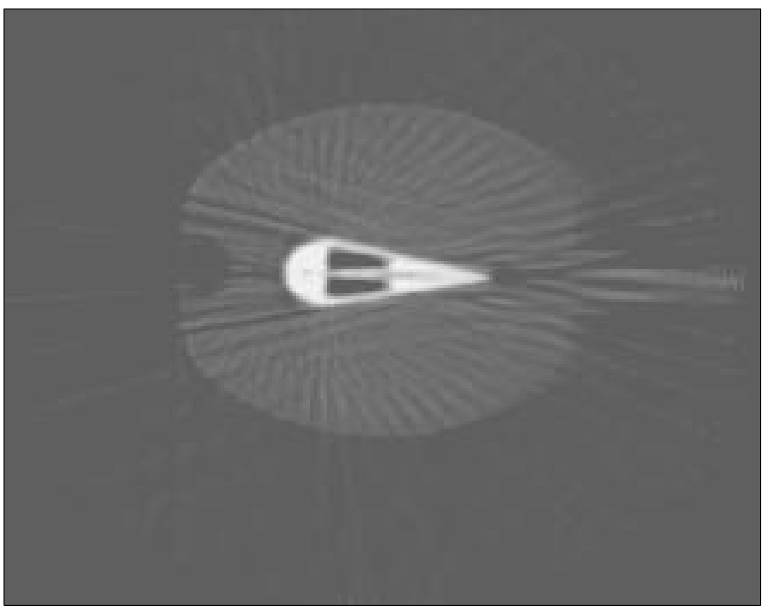
This data was taken at LANSCE.



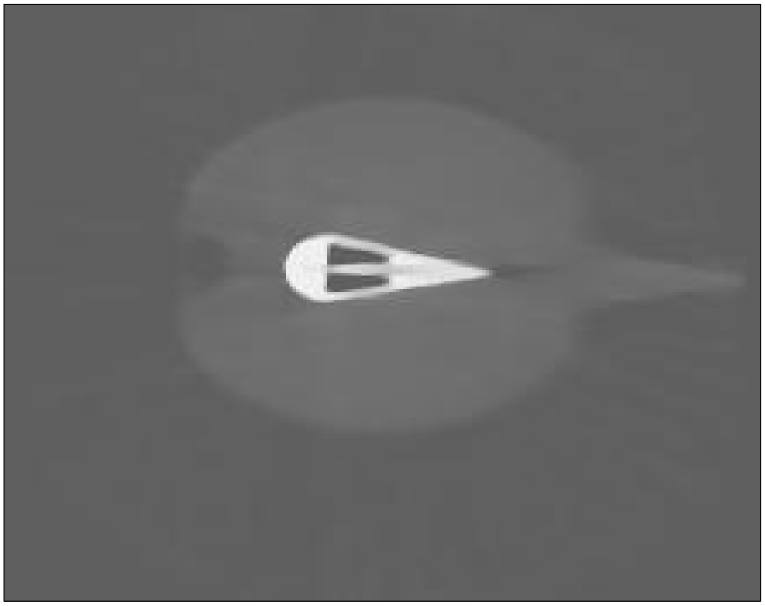
Now we show preliminary results obtained by total variation minimization using the fixed point method. The results were obtained using our own hacks of Pep Mulet's ImageTool, a matlab package.

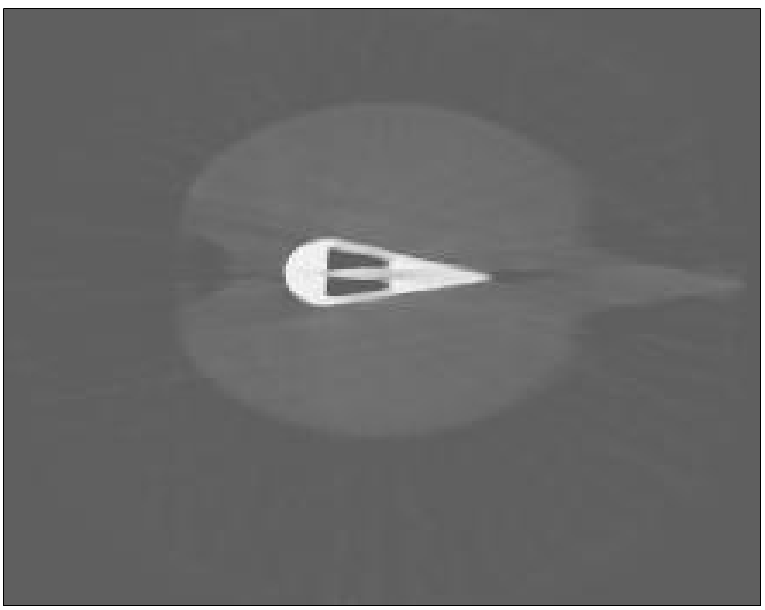
BCO4 Results: continued





BCO4 Results: continued

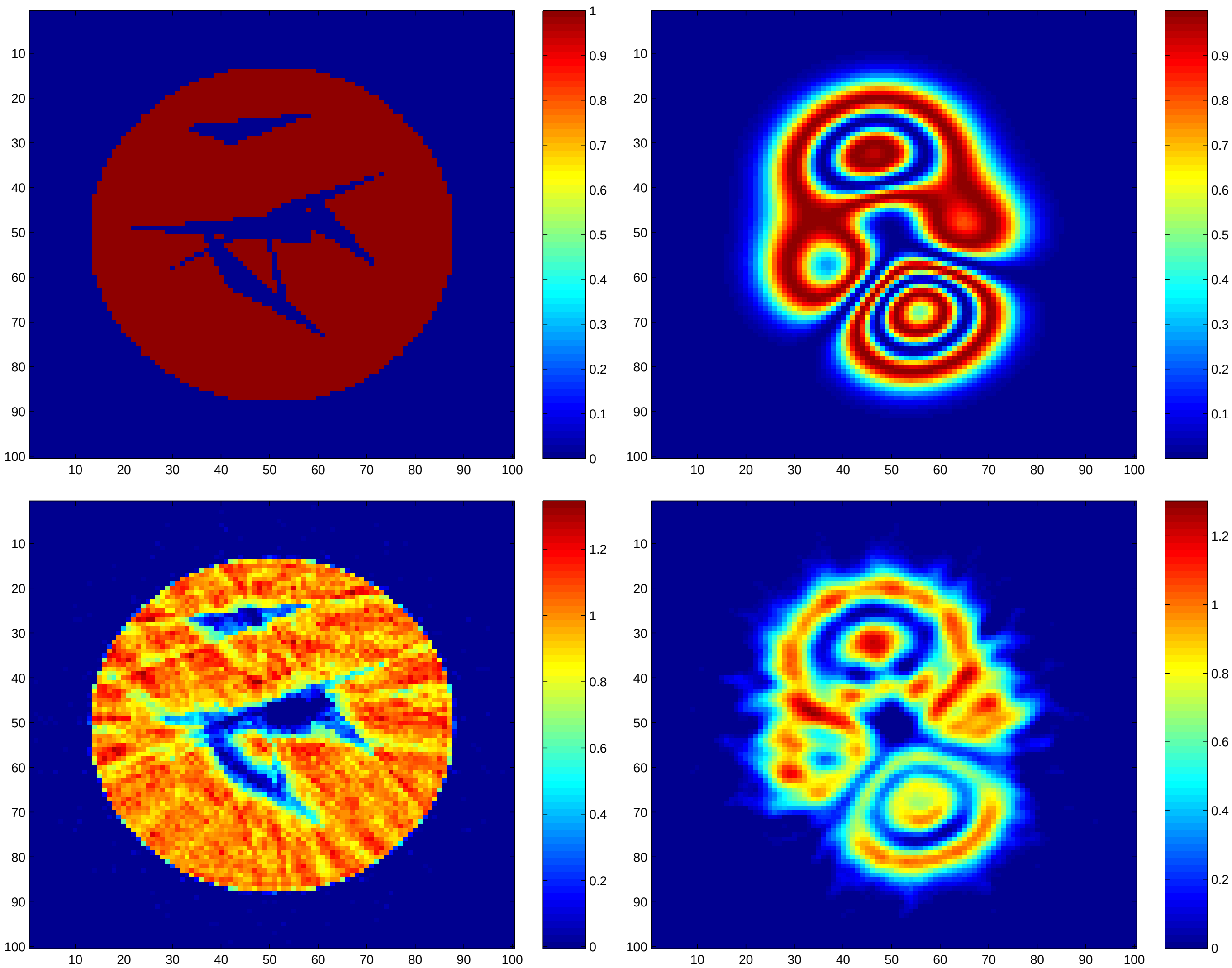




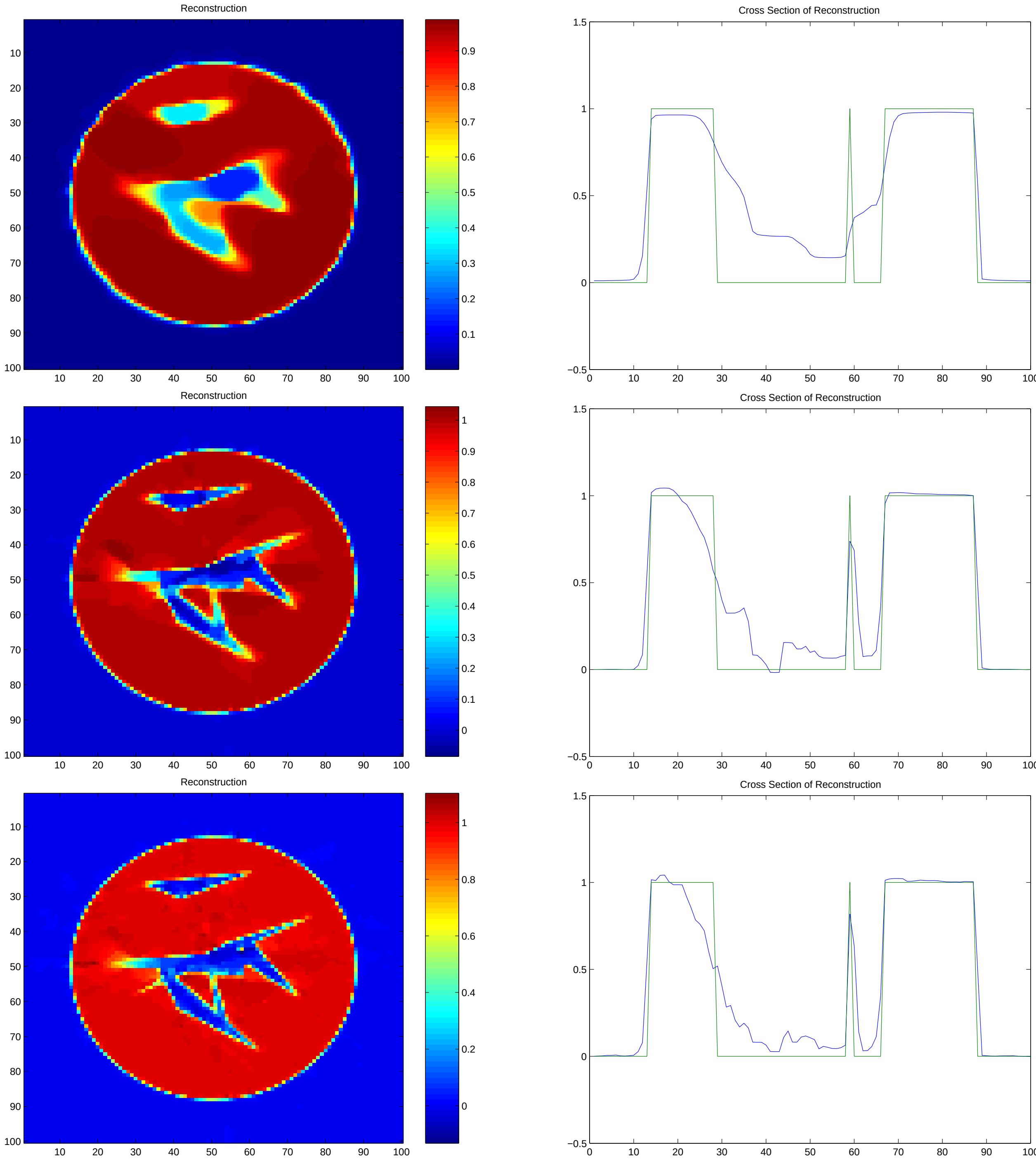
A Closer Look: Simulated Radiographs

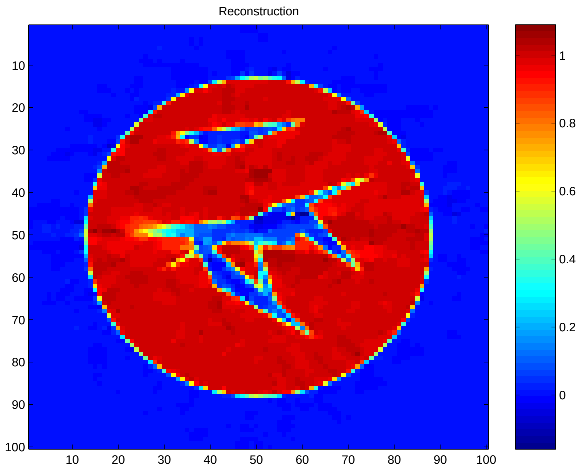
In these experiments we have been using a modified version of the fixed point code of Curt Vogel's to begin to explore:

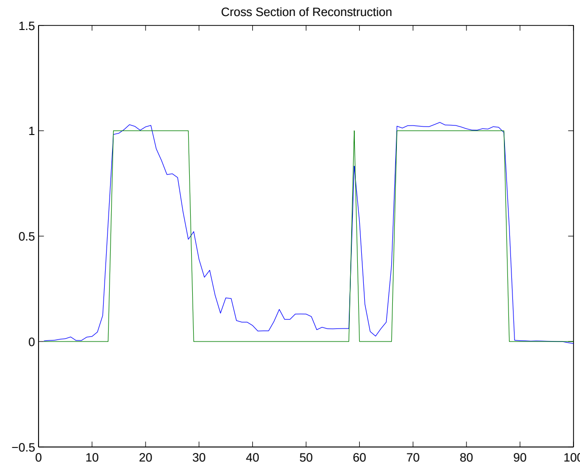
The object and truncated SVD reconstruction: The actual objects.



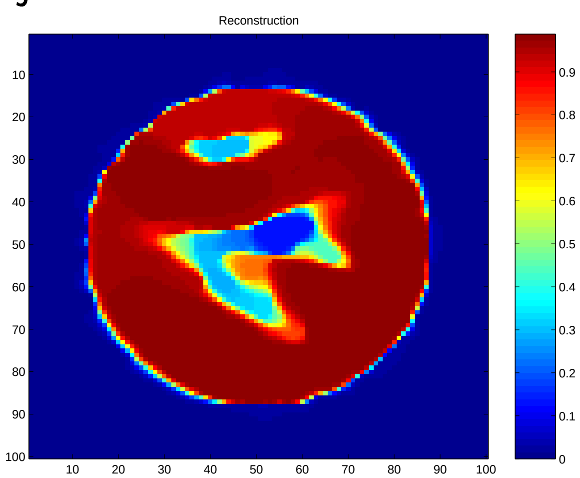
Object 2 results: 1% noise

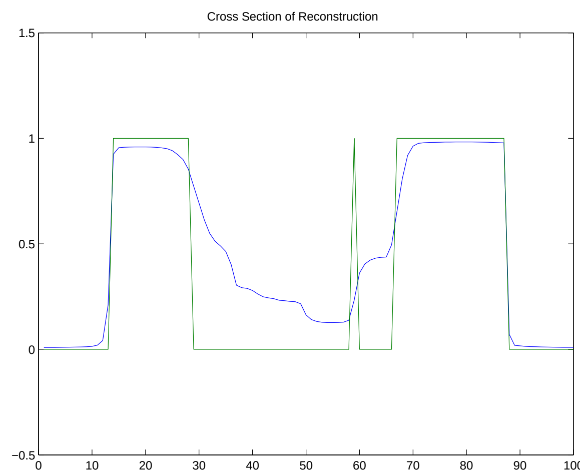


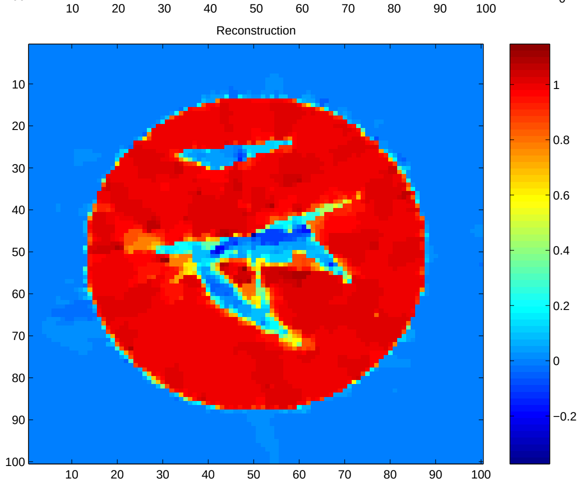


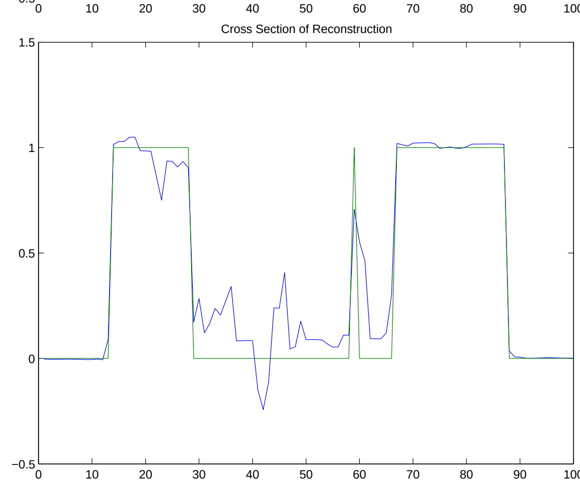


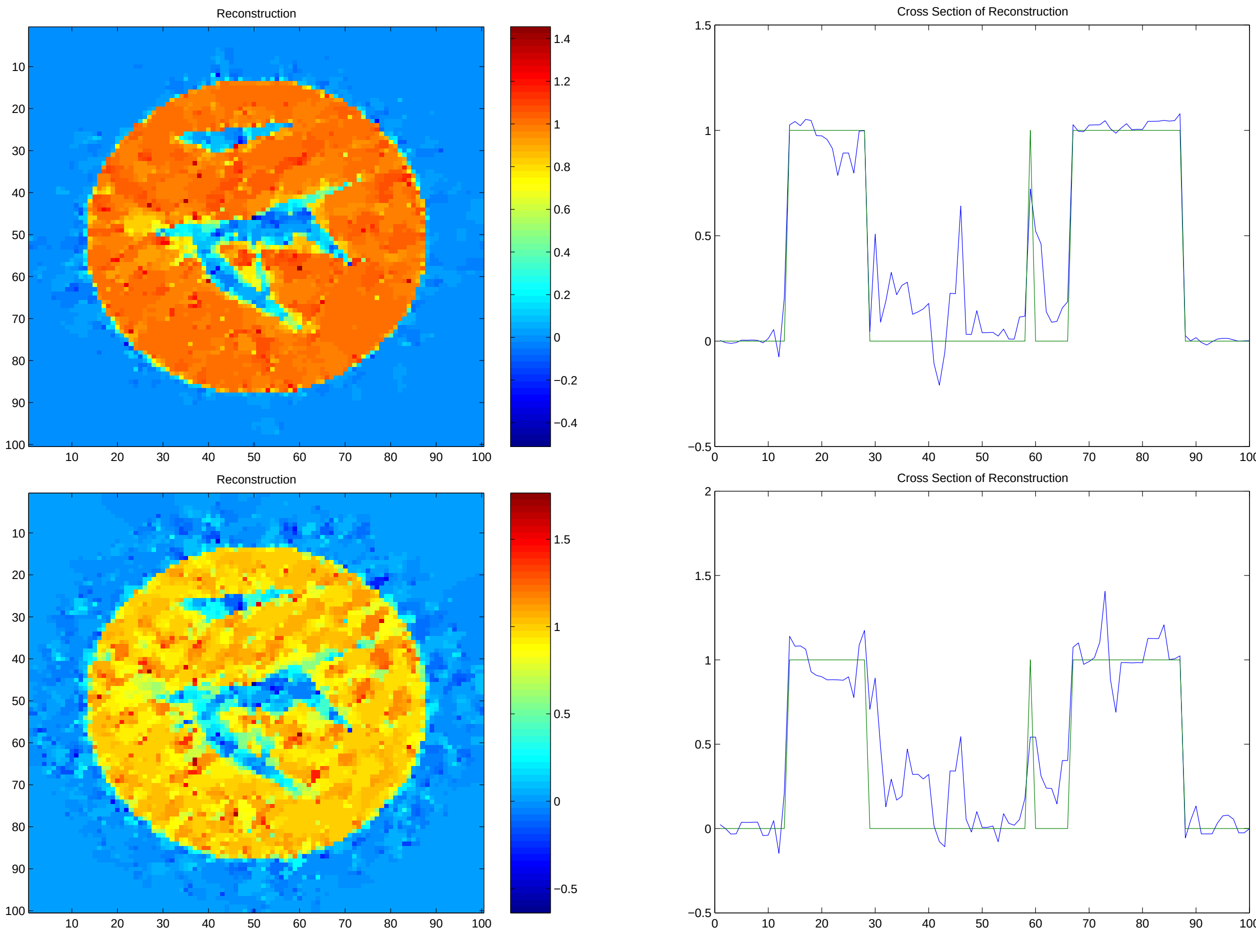
Object 2 results: 10% noise



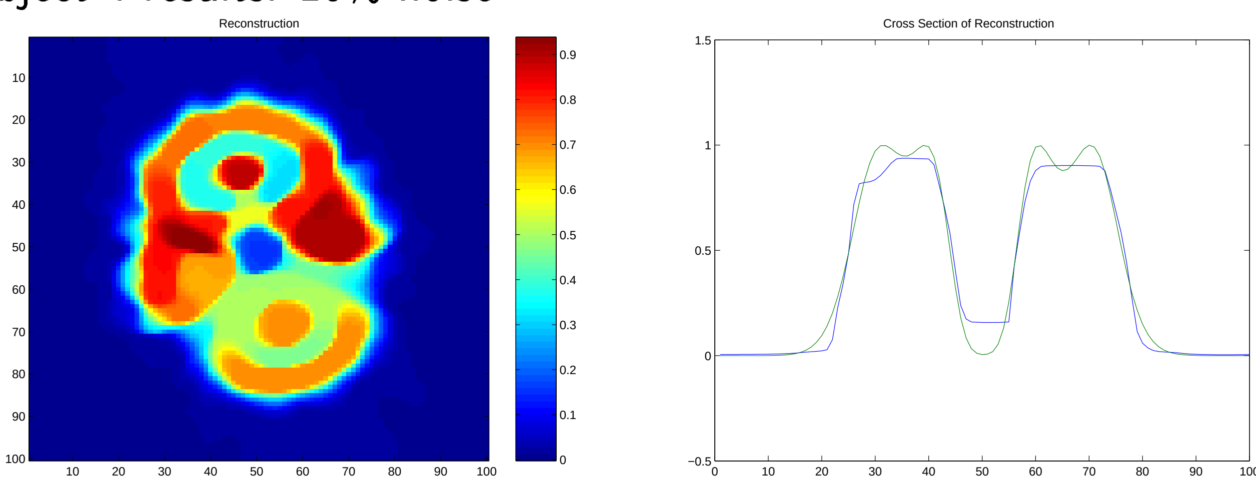


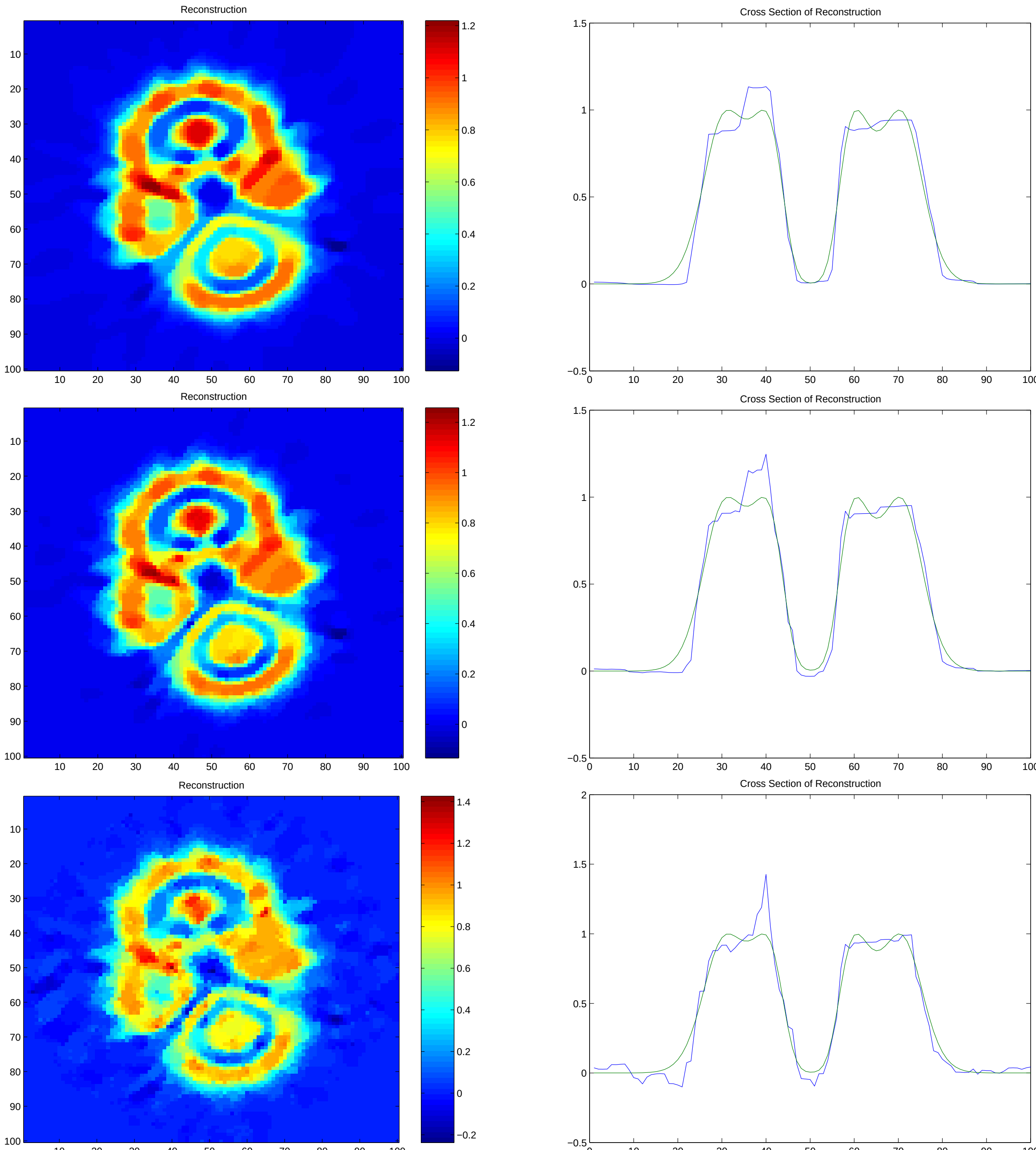




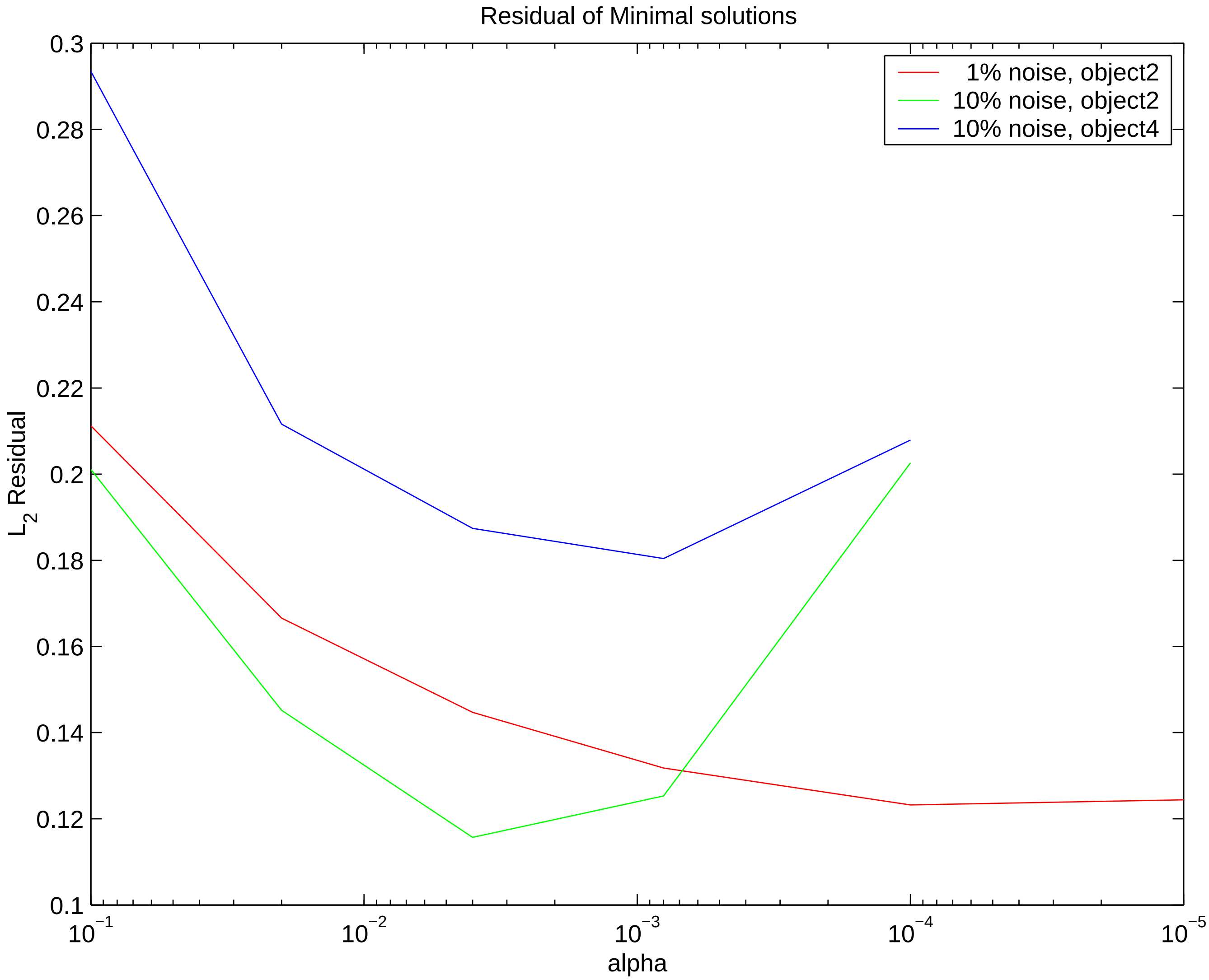


Object 4 results: 10% noise




Reconstruction
Cross Section of Reconstruction
Reconstruction
Cross Section of Reconstruction
Reconstruction
Cross Section of Reconstruction

Residual for the simulated radiograph results.



Current work

Primal-Dual As mentioned above this method can potentially speed up the convergence to the minimum of $F(u)$. There may be a problem in the implementation related to the size of $P^T P$. We are currently working on this with David Strong.

... from the paper by Chan, Golub, and Mulet (next page) ...

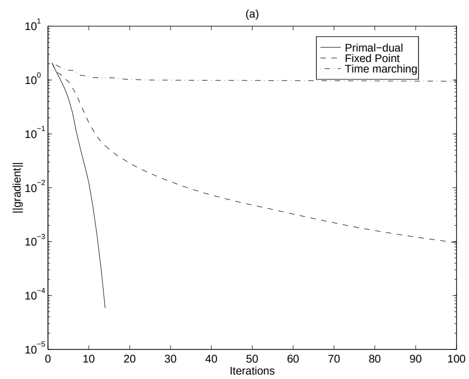


FIG. 7. *Plot of the L_2-norm of the gradient $g(u)$ of the objective function versus iterations for the different methods.*

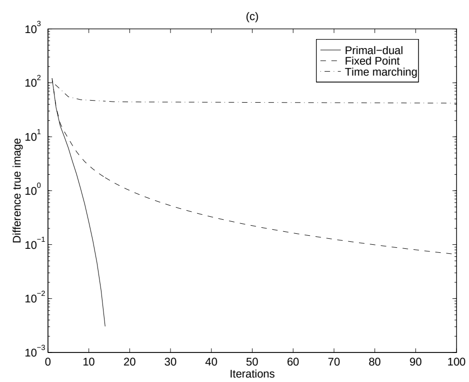


FIG. 8. *Plot of the L_2-norm of the difference between the current iterate and the solution for the problem computed by Newton's method with high accuracy versus iterations for the different methods.*

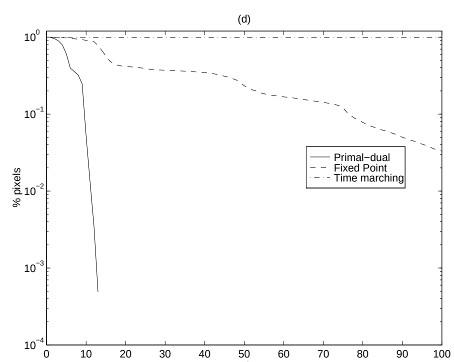


FIG. 9. *Plot of # pixels which differ more than .001 (relatively) from the solution for the problem computed by Newton's method with high accuracy versus iterations for the different methods.*

Bundle Methods In these methods, one is trying to use information obtained from a sequence of measurements of the non-smooth cost function to form an estimate of the sub-differential of the cost function.

Feature Corrections Because of the feature scale dependent manner in which the denoising effects the reconstructions, it makes sense to investigate recovered feature based corrections.

Restricted Searches In this formulation one simply looks only in the span of N (the null space of P) for increments, i.e. one solves

$$\arg\min_{\psi} TV(N^T\psi + u_{svd})$$

where ψ is the vector of coefficients for our null space perturbation.

Rigorous Results Convergence results and rates for modified methods suggested by the applications to radiography.

Other Priors Incorporation of other prior constraints into the variational framework.

Geometry and Dynamics Theoretical study of the parameterized dynamics induced by the iterative solution to the variational problem.

References

There are of course a large number of pertinent papers and monographs.

In particular I recommend:

[1] Papers to be found as CAM reports at UCLA: Chan, Osher, Mulet, Strong, Vese, ... url
http://www.math.ucla.edu/ imagers/htmls/reports.html

[2] Evans and Gariepy, “Measure Theory and Fine Properties of Functions”.

[3] Krantz and Parks, “The Geometry of Domains in Space”

[3] Ekeland and Temam, “Convex Analysis and Variational Problems”

[4] Bertsekas, “Nonlinear Programming”

[5] Zeidler, “Nonlinear Functional Analysis and it’s Applications III: Variational Methods and optimization”

[6] Vogel, “Computational Methods for Inverse Problems”