Reconciling Bayesian Regularization and Total Variation Regularization

Andrew Stuart

Computing and Mathematical Sciences
California Institute of Technology

Matt Dunlop, Charlie Elliott, Viet Ha Hoang
AFOSR, DARPA, EPSRC, ONR
Imperial College

May 9th 2018

Overview

Underlying Concepts

Mesh Invariance

Bayesian Level Set Inversion

Phase Field Inversion

Conclusions and References

Underlying Concepts

Inverse Problem

Problem Statement

Find u from y where $\mathsf{G} : \mathcal{U} \mapsto \mathcal{Y}$, where $\mathcal{U}, \mathcal{Y}$ are Banach spaces, η is noise and

$$y = \mathsf{G}(u) + \eta, \quad \eta \sim \mathsf{N}(0, \gamma^2 I).$$

Issues

- η is not known;
- y may not be in the image space of G (no solution);
- Many u may satisfy $\|y - \mathsf{G}(u)\| \ll 1$ (multiple solutions).

[4] Engl, Hanke, Neubauer 1996

[6] Kaipio, Somersalo 2006

Regularization Via Optimization

Least Squares Regularization

$$\text{Least Squares Functional:} \quad \Phi(u; y) := \frac{1}{2\gamma^2} \|y - \mathsf{G}(u)\|^2,$$

$$\text{Regularized LSQ Functional:} \quad I(u; y) := \Phi(u; y) + R(u).$$

Regularization: Spaces Of Functions

$$\text{Unknown Function:} \quad u : D \subset \mathbb{R}^d \mapsto \mathbb{R},$$

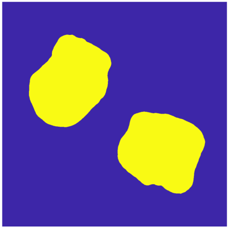
$$\text{TV Regularization:} \quad R(u) := \lambda \int_D |\nabla u(x)| dx,$$

$$\text{Tikhonov Regularization:} \quad R(u) := \lambda \|u\|^2_{H^s(D;\mathbb{R})}.$$

Tikhonov:[4] Engl, Hanke, Neubauer 1996

TV: [10] Rudin, Osher, Fatemi 1992, [12] Schoenleib 2015

TV Regularization for Binary Functions



Binary Functions

$$\begin{aligned}\text{Binary Function:} \quad & u(x) \in \{\pm 1\}, \quad \forall x \in D,\\ \text{Regularized LSQ Functional:} \quad & I(u; y) := \Phi(u; y) + R(u),\\ \text{TV Regularization:} \quad & R(u) := \lambda \times |\text{perimeter}|.\end{aligned}$$

Bayesian Regularization

Set-Up

Assume u, y, η are random, $u \perp \eta$ and $\eta \sim N\big(0, \gamma^2 I\big)$. Then:

$$\begin{aligned} \text{Prior:} \quad & u \sim \nu_0(du), \\ \text{Likelihood:} \quad & y|u := N\big(\mathsf{G}(u), \gamma^2 I\big), \\ \text{Posterior:} \quad & u|y \sim \nu(du). \end{aligned}$$

Bayes Theorem

$$\nu(du) \propto \exp\big(-\Phi(u;y)\big)\nu_0(du), \qquad 2\gamma^2\Phi(u;y)\|y - \mathsf{G}(u)\|^2.$$

$$\nu_0(du) = \exp\big(-R(u)\big)\,du \Longrightarrow \nu(du) = \exp\big(-I(u;y)\big)\,du.$$

[6] Kaipio, Somersalo 2006, [12] Stuart 2010, [2] Dashti, Law, Stuart, Voss 2010

Mesh Invariance

Tikhonov Setting

Key Question

- Let $R_N(u)$ be a finite dimensional approximation of $\|u\|^2_{H^s}$.
- Does $\exp\big(-R_N(u)\big)du$ approach an $N \to \infty$ limiting measure?
- e.g. $s = d = 1$ and $R_N(u) \propto \sum_{j=0}^{N-1} |u_{j+1} - u_j|^2$.

Answer: Yes for Tikhonov

$$\text{Limiting Measure:} \quad N\big(0, (-\triangle)^{-s}\big) \text{ on } L^2(D;\mathbb{R}),$$

$$\text{Limiting Regularizer:} \quad R(u) = \frac{1}{2}\langle u, (-\triangle)^s u\rangle.$$

[7] Lasanen 2002

[12] Stuart 2010

TV Setting

Key Question

- Let $R_N(u)$ be a finite dimensional approximation of $\|u\|_{\mathrm{TV}}$.
- Does $\exp\big(-R_N(u)\big)du$ approach an $N \to \infty$ limiting measure?
- e.g. $d = 1$ and $R_N(u) \propto \sum_{j=0}^{N-1} |u_{j+1} - u_j|$.

Answer: No for TV

$$
\begin{aligned}
\text{Limiting Measure:} \quad & \exp\big(-R_N(u)\big)du \text{ diverges/Gaussian}, \\
\text{Besov Alternative:} \quad & R(u) = \lambda \|u\|_{B^1_{1,1}(D;\mathbb{R})}, \\
\text{Note:} \quad & \|u\|_{\mathrm{TV}} \leq C\|u\|_{B^1_{1,1}(D;\mathbb{R})}.
\end{aligned}
$$

[8] Lassas, Siltanen 2004

[9] Lassas, Saksman, Siltanen 2009

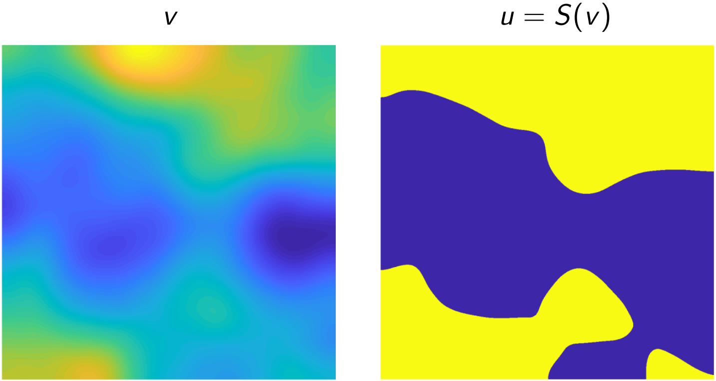
Bayesian Level Set Inversion

[3] Iglesias, Lu, Stuart 2016 [3] Dunlop, Elliott, Hoang, Stuart 2018

The Level Set Function

$$S(v) = 1, \;\; v \geq 0, \quad \text{and} \quad S(v) = -1, \;\; v < 0.$$

Lift $S : \mathbb{R} \mapsto \mathbb{R}$ to $S : C(D; \mathbb{R}) \mapsto L^{\infty}(D; \mathbb{R})$.



Gaussian Measures

Precision Operator Via The Laplacian

$A : H^4_{\mathrm{per}}(D;\mathbb{R}) \to L^2(D;\mathbb{R})$ given by the identity

$$Av = \delta\triangle^2 v - q\delta\triangle v + \tau^2\delta v.$$

Set $\mu_0 := N(0, A^{-\alpha/2})$, $\alpha > \frac{d}{2}$.

Regularity Theorem

Let $v \sim \mu_0$. Then μ_0 is supported on $C^s(D;\mathbb{R})$ if $s < \alpha - \frac{d}{2}$.

Level Set

Perimeter Theorem

$$BV_{\text{binary}}(D;\mathbb{R}) := \{\psi \in BV(D;\mathbb{R}) : \ \psi(D) \subset \{\pm 1\}\}.$$

Let $\nu_0 = S^\sharp \mu_0$, $\alpha > 1 + \frac{d}{2}$. Then ν_0, ν are supported on BV_{binary}

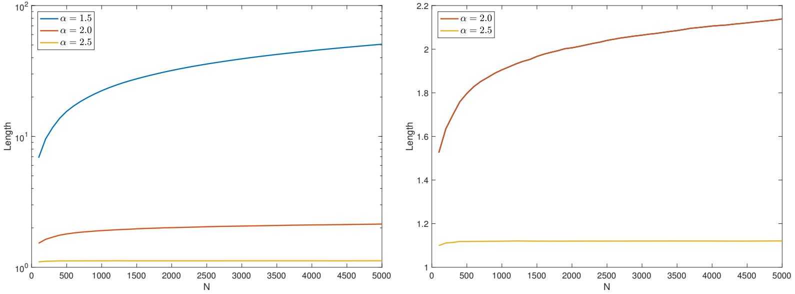


Figure: $d = 2$. Perimeter versus N for $\alpha = 1.5$ (blue), 2.0 (brown) and 2.5 (yellow). Left log scale; right natural scale. Nested prior w.r.t N.

Example: Linear Observations

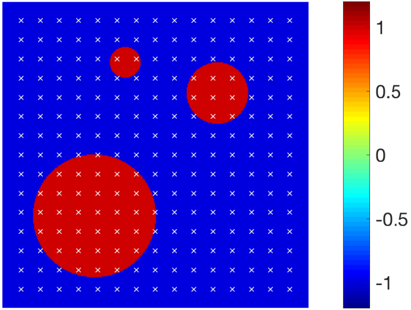
Linear Inverse Problem

Find $u|y$ from y, where $u \sim \nu_0$ and $y|u$ is defined by

$$y_j = u(x_j) + \eta_j.$$

Concatenate

$$y = G(u) + \eta.$$



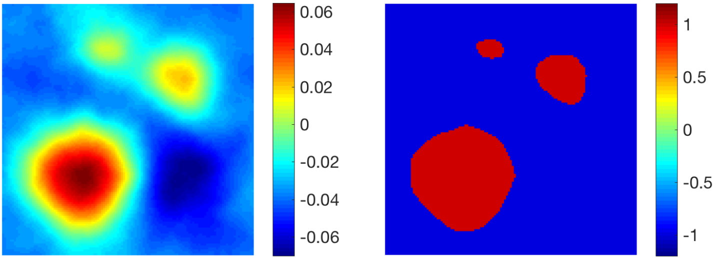
Example: Linear Observations

Posterior Means: Level Set

Inverse problem: find v from y where

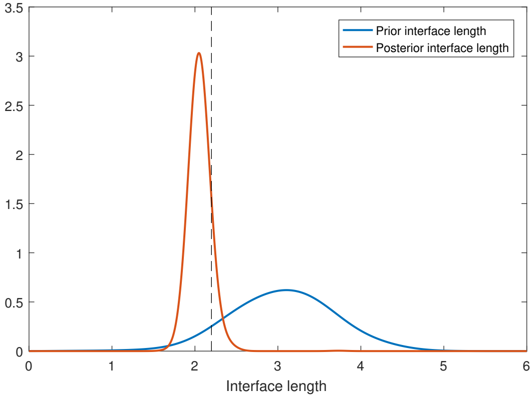
$$y = (G \circ S)(v) + \eta.$$

Let $v \sim \mu_0$ and compute posterior on $v|y$. Use MCMC to estimate $\mathbb{E}(v)$ (left) and $S(\mathbb{E}(v))$ (right) under this posterior.



Perimeter Learning

Distribution of the perimeter of the zero level set under the prior (blue) and posterior (brown) distributions. Perimeter is learnt.



Phase Field Inversion

[3] Dunlop, Elliott, Hoang, Stuart 2018

Example: Linear Observations With Small Noise

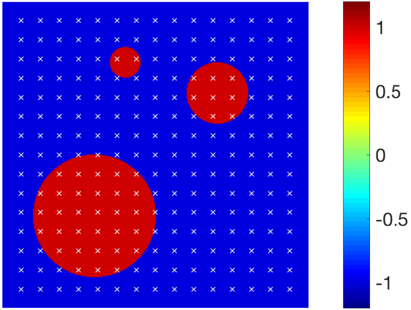
Linear Inverse Problem

Find $u|y$ from y, where $u \sim \nu_0$ and $y|u$ is defined by

$$y_j = u(x_j) + \varepsilon^c \eta_j, \quad \eta_j \sim \mathrm{N}(0,1).$$

Concatenate

$$y = G(u) + \varepsilon^c \eta, \quad \eta_j \sim \mathrm{N}(0,I).$$



Phase Field Priors

Precision Operator Via The Laplacian

$A : H^4_{\mathrm{per}}(D;\mathbb{R}) \to L^2(D;\mathbb{R})$ given by the identity

$$Au = \delta\varepsilon^{-2a_1}\triangle^2 u - q\delta\varepsilon^{-2a_2}\triangle u + \tau^2\delta\varepsilon^{-2a_3}u.$$

Let $\mu_0 := N(0, A^{-1})$.

Prior

For $r, b > 0$ define ν_0 via

$$\frac{d\nu_0}{d\mu_0} = \frac{1}{Z_0}\exp\left(-\frac{r}{\varepsilon^b}\int_D \frac{1}{4}\left(1-u(x)^2\right)^2 dx\right).$$

Then ν_0, ν are supported on $C^s(D;\mathbb{R})$ for $s < 2 - \frac{d}{2}$. They have no support on $BV(D;\mathbb{R})$.

Probability Maximizers

Parameter Choice

$$a_2 - a_1 = 1, \quad 3 + 2a_1 - b = -1,$$
$$3 + 2a_1 - 2c = 0, \quad 3 + 2(a_1 - a_3) = a > 0.$$

Γ−Limit Theorem

Small ball of vanishing radius are of maximizal probability when centred on minimizers of $I^{\varepsilon}(u)$. If

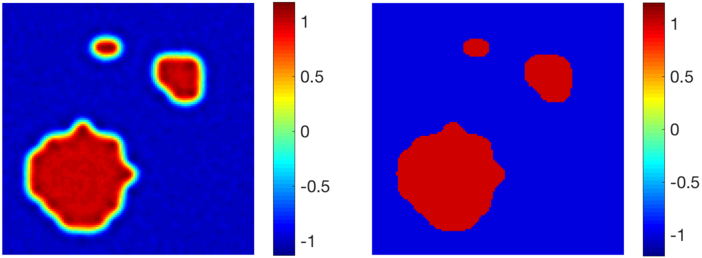
$$I_0^{\delta} = \frac{\lambda^{\delta}}{2} \int_D |\nabla u| \, dx + \frac{1}{2\gamma^2} |(y - Ku)|^2, \quad \textit{if} \;\; u \in BV_{\mathrm{binary}}(D)$$

then, for appropriately defined λ^{δ}, $I_0^{\delta} = \lim_{\varepsilon \to 0} I^{\varepsilon}$ in the sense of $\Gamma-$convergence in the strong $L^1(D; \mathbb{R})$ topology.

Example: Linear Observations

Posterior Means: Phase Field

We use MCMC samples to estimate $\mathbb{E}(v)$ (left) and $S(\mathbb{E}(v))$ (right) under the posterior.



Conclusions and References

Conclusions

- Inverse problems require regularization.
- Optimization versus statistical regularization.
- TV good for inversion with interfaces (optimization).
- TV hard to use explicitly in Bayesian mode (statistical).
- Bayesian level set method implicitly penalizes TV norm (statistical).
- Phase field inversion implicitly penalizes TV norm (optimization).

References

[1] S. L. Cotter, G. O. Roberts, A. M. Stuart, D. White, *MCMC methods for functions: modifying old algorithms to make them faster*, Statistical Science, 28 (2013), pp. 424–446.

[2] M. Dashti, K. J. Law, A. M. Stuart, and J. Voss, *MAP estimators and their consistency in Bayesian nonparametric inverse problems*, Inverse Problems, 29 (2013), p. 095017.

[3] M. M. Dunlop, C. M. Elliott, V. Ha Hoang and A. M. Stuart, *Reconciling Bayesian and Total Variation Methods for Binary Inversion*, arXiv preprint arXiv:1706.01960.

[4] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, vol. 375, Springer Science & Business Media, 1996.

[5] M. A. Iglesias, Y. Lu, and A. M. Stuart, *A Bayesian level set method for geometric inverse problems*, Interfaces and Free Boundary Problems, (2016).

[6] J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, vol. 160, Springer Science & Business Media, 2006.

References

[7] S. Lasanen, *Discretizations of generalized random variables with applications to inverse problems.*, PhD, University of Oulu, 2002.

[8] M. Lassas and S. Siltanen, *Can one use total variation prior for edge-preserving Bayesian inversion?*, Inverse Problems, 20 (2004), p. 1537.

[9] M. Lassas, E. Saksman, and S. Siltanen, *Discretization-invariant Bayesian inversion and Besov space priors*, Inverse Problems and Imaging, 3 (2009), pp. 87–122.

[10] L.I. Rudin, S Osher and E Fatemi, *invariant Nonlinear total variation based noise removal algorithms*, PhysicaD, 60 (1992), pp. 259–268.

[11] C.-B. Schœnlieb, *Partial Differential Equation Methods for Image Inpainting*, CUP, 2015.

[12] A. M. Stuart, *Inverse problems: a Bayesian perspective*, Acta Numerica, 19 (2010), pp. 451–559.