Supporting Information for
“An EPIC Tikhonov regularization: Application to quasi-static fault slip inversion”

F. Ortega-Culaciati\textsuperscript{1}, M. Simons\textsuperscript{2}, J. Ruiz\textsuperscript{1}, L. Rivera\textsuperscript{3}, N. Diaz-Salazar\textsuperscript{1}

\textsuperscript{1}Departamento de Geofísica, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile.
\textsuperscript{2}Seismological Laboratory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA.
\textsuperscript{3}Université de Strasbourg, CNRS, IPGS UMR 7516, F-67000 Strasbourg.

Contents

1. Text S1 to S13
2. Figures S.1 to S.73

Introduction

The following provides a brief description, range of contained figures and starting page of each supplementary information in this document.

Text S1  Relates observation, model prediction and prior information covariance matrices to posterior model estimates perturbations. ................................................................. 2
Text S2  Description of illposedness of quasi-static slip inversion and effectiveness of higher order Tikhonov regularization. Figures S.1 - S.2 ........................................... 5
Text S3  Detailed description and implementation of the EPIC ..................................................... 8
Text S4  Definition of the subduction model for Japan Trench megathrust. Figure S.3 ............... 15
Text S5  Definitions for calculation of posterior spatial correlation lengths of slip estimates. Figures S.4 - S.5 .................................................................................. 18
Text S6  Inversion results for the recovery of synthetic checkerboard slip models as a function of the noise realization. Figures S.6 - S.12 ....................................................... 22
Text S7  Inversion results for the recovery of synthetic elliptical (M\textsubscript{s} 8.0) slip models as a function of the noise realization. Figures S.13 - S.19 ....................... 30
Text S8  Inversion results for the recovery of synthetic elliptical (M\textsubscript{s} 8.5) slip models as a function of the noise realization. Figures S.20 - S.26 ....................................... 38
Text S9  Inversion results for the recovery of synthetic elliptical (M\textsubscript{s} 9.0) slip models as a function of the noise realization. Figures S.27 - S.33 ......................... 46
Text S10 Inversion results for the recovery of all synthetic slip cases as a function of the regularization parameter. Figures S.34 - S.37 ........................................... 54
Text S11 Generalized Cross Validation cost functions and selected estimated models with different values of the regularization parameter for the recovery of synthetic slip distributions shown in Figures 2 and 3 of the main body of the article. Figures S.38 - S.51 ........ 59
Text S12 Details of the inversion results for the 2011 (M\textsubscript{s}9.0) Tohoku-Oki earthquake using ET2 regularization. Figures S.52 - S.61 ................................................... 74
Text S13 Details of the inversion results for the 2011 (M\textsubscript{s}9.0) Tohoku-Oki earthquake using ET2 regularization and positivity constraints. Figures S.62 - S.73 ............... 85
Text S14 EPIC feasibility and selection of range of possible values for target posterior standard deviation \( \sigma_t \). Figures S.74 - S.75 ......................................................... 98

Corresponding author: F. Ortega-Culaciati, ortega.francisco@uchile.cl
Text S1  On the perturbation of model parameter estimates due to observational, model prediction and prior information uncertainties

Here we discuss the stability of the model parameter estimates upon perturbations on the data space, i.e., due to uncertain observations or model predictions, and upon perturbations on the space describing the regularized quantity $h$.

For the general linear least squares problem, we solve

$$\min_m \phi(m)$$ \hspace{1cm} (S.1)

where $\phi(m)$ is the objective function,

$$\phi(m) = \frac{(Gm - d)^\top C_x^{-1} (Gm - d)}{\phi_d(m)} + \frac{(Hm - h^0)^\top C_h^{-1} (Hm - h^0)}{\phi_h(m)}$$

$$= \phi_d(m) + \phi_h(m)$$ \hspace{1cm} (S.2)

Let $\delta d$ be a perturbation on the data space, for example, due to a combination of observational and model prediction errors, we can write the objective function of the perturbed general least squares problem as,

$$\tilde{\phi}(m) = \phi_{d + \delta d}(m) + \phi_h(m)$$ \hspace{1cm} (S.3)

where,

$$\phi_{d + \delta d}(m) = \frac{(Gm - (d + \delta d))^\top C_x^{-1} (Gm - (d + \delta d))}{\phi_d(m)}$$

$$= \frac{(Gm - d)^\top C_x^{-1} (Gm - d) + \delta d^\top C_x^{-1} \delta d - (Gm - d)^\top C_x^{-1} \delta d - \delta d^\top C_x^{-1} (Gm - d)}{\phi_d(m)}$$

$$= \phi_d(m) + \delta d^\top C_x^{-1} \delta d - 2\delta d^\top C_x^{-1} (Gm - d)$$

$$= \phi_d(m) + \delta d^\top C_x^{-1} \delta d + 2\delta d^\top C_x^{-1} \delta d - 2\delta d^\top C_x^{-1} Gm$$

$$= \phi_d(m) - 2\delta d^\top C_x^{-1} Gm + \kappa(\delta d)$$ \hspace{1cm} (S.4)

where, $\kappa(\delta d)$ is a function that does not depend on model parameters $m$ (i.e. a constant for $m$), but only on the data $d$ and the data perturbation $\delta d$.

Now, from (S.2), (S.3) and (S.4), we obtain an expression of the cost function for the perturbed least squares problem that depends on the cost function of the unperturbed problem,

$$\tilde{\phi}(m) = \phi_{d + \delta d}(m) + \phi_h(m)$$

$$= \phi_d(m) + \phi_h(m) - 2\delta d^\top C_x^{-1} Gm + \kappa(\delta d)$$ \hspace{1cm} (S.5)

Let $m^*$ and $C_m$ be the posterior maximum likelihood model and covariance matrix of the unperturbed problem (S.2). The cost function of the unperturbed problem can be written in terms of $m^*$ and $C_m$ as,

$$\phi(m) = (m - m^*)^\top C_m^{-1} (m - m^*) + R$$ \hspace{1cm} (S.6)

where $R$ is a constant that does not depend on $m$ [e.g., Tarantola, 2005]. Thus, the objective function of the perturbed problem can be written as,

$$\tilde{\phi}(m) = (m - m^*)^\top C_m^{-1} (m - m^*) + R + \kappa(\delta d) - 2\delta d^\top C_x^{-1} Gm$$ \hspace{1cm} (S.7)
Now, in order to obtain the solution of the perturbed problem \( \min_m \tilde{\phi}(m) \), we impose the condition
\[
\nabla \tilde{\phi}(m) = 0,
\]

obtaining,
\[
\nabla \tilde{\phi}(m) = 2C_m^{-1}(m - m^*) - 2G^\top C_X^{-1}d = 0 \quad (\text{S.8})
\]

\[
\Leftrightarrow C_m^{-1}(m - m^*) = G^\top C_X^{-1}d \quad (\text{S.9})
\]

\[
\Leftrightarrow m - m^* = C_m G^\top C_X^{-1}d \quad (\text{S.10})
\]

Now denote \( \delta m^{\text{ad}} = m - m^* \) to the perturbation of the estimated model \( m \) from the maximum likelihood model \( m^* \) (estimated without perturbation) due to the data perturbation \( d \), and \( \delta d^m = G^\top C_X^{-1}d \) to the projection of the data perturbation into model parameter space. Then, equation (S.10) can be written as,
\[
\delta m^{\text{ad}} = C_m \delta d^m \quad (\text{S.11})
\]

In what follows we estimate the covariance matrix of \( \delta m^{\text{ad}} \). If we assume that the misfit covariance matrix \( C_X \) is well chosen, in the sense that it truly represents any data space perturbation \( d \), we can state that \( \delta d \) follows an unbiased normal distribution with covariance matrix \( C_X \). Thus, using the error propagation formula, we can calculate the covariance matrix for \( \delta d_X = W_X d \) as,
\[
C_{\delta d_X} = W_X C_X W_X^\top \quad \text{but, } C_X = \left(W_X^\top W_X\right)^{-1}
\]

\[
= W_X \left(W_X^\top W_X\right)^{-1} W_X^\top
\]

\[
= W_X W_X^{-1} W_X^\top W_X^\top
\]

\[
\Rightarrow C_{\delta d_X} = I \quad \text{← the identity matrix. (S.12)}
\]

Now, the data perturbation projected into model parameter space becomes,
\[
\delta d^m = G^\top C_X^{-1}d = (W_X G)^\top (W_X \delta d) = (W_X G)^\top \delta d_X \quad (\text{S.13})
\]

with covariance matrix obtained using the error propagation formula as follows:
\[
C_{\delta d^m} = (W_X G)^\top C_{\delta d_X} (W_X G)
\]

\[
\text{but } C_{\delta d_X} = I
\]

\[
= G^\top C_X^{-1} G
\]

\[
= P_u = C_m^{-1} \quad (\text{S.14})
\]

where \( P_u \) and \( C_m^{-1} \) are the precision matrix and posterior covariance matrix of the non-regularized inversion problem \( (W_X G m = W_X d) \), respectively. Note that we only need \( P_u \), and that here we abuse notation since for ill-posed or rank deficient problems the inverse of \( P_u \) (i.e., \( C_m^{-1} \)) might be challenging to compute or it simply can not be directly computed. \( P_u \) could be close to a singular matrix (or be one), one of the reasons for requiring a regularization term to solve the inverse problem.

Then, the covariance matrix of the perturbation of the inverse problem solution due to a data perturbation, \( \delta m^{\text{ad}} = C_m \delta d^m \), can be written as,
\[
C_{\delta m}^{\text{ad}} = C_m C_{\delta d^m} C_m^\top
\]

\[
= C_m G^\top C_X^{-1} G C_m
\]

\[
= C_m P_u C_m \quad (\text{S.15})
\]

Given the structure of \( C_{\delta m}^{\text{ad}} \) it is difficult to conclude about the variances for the perturbations of model parameters from the optimum value due to a data perturbation for the regularized inversion problem. However, as we show in the following, those variances are bounded.

Following a similar procedure as done to deduce the formula for \( \delta m^{\text{ad}} \), it can be shown that the perturbation of estimated model parameters from their optimum value due to a perturbation \( \delta h \)
of the chosen (uncertain/somewhat arbitrary) values \( h \), that we assumed has unbiased Gaussian errors with covariance matrix \( C_h \), can be written as:

\[
\delta m^{\delta h} = C_m H^T C_h^{-1} \delta h
\]  

(S.16)

The covariance matrix for \( \delta m^{\delta h} \) can be calculated using the error propagation formula as:

\[
C_{\delta m} = (C_m H^T C_h^{-1} C_h (C_m H^T C_h^{-1}))^T
\]

\[
= C_m H^T C_h^{-1} C_h H C_m
\]

\[
= C_m H^T C_h^{-1} H C_m
\]

(S.17)

Now, due to linearity of the inverse problem, it is direct to show that model perturbations \( \delta m \) due to errors in both data and regularized quantity spaces can be written as:

\[
\delta m = \delta m^{\delta d} + \delta m^{\delta h}
\]

\[
= C_m G^T C_h^{-1} \delta d + C_m H^T C_h^{-1} \delta h
\]

(S.18)

To find the covariance matrix of \( \delta m \), recall that data and prior information are assumed to be independent while formulating the inverse problem [e.g., Tarantola, 2005], thus we can write:

\[
C_{\delta m} = C_{\delta m}^{\delta d} + C_{\delta m}^{\delta h}
\]

\[
= C_m G^T C_h^{-1} G C_m + C_m H^T C_h^{-1} H C_m
\]

\[
= C_m \left[ G^T C_h^{-1} G + H^T C_h^{-1} H \right] C_m
\]

\[
\Rightarrow C_{\delta m} = C_m
\]

(S.19)

Then, model parameter perturbations from the a posteriori maximum likelihood model due to errors on observations, model prediction and regularization, follow an unbiased Gaussian distribution with covariance matrix equal to the a posteriori covariance matrix of the estimated model parameters \( C_m \). Although the previous statement is trivial, as \( C_m, C_{\delta m}^{\delta d}, C_{\delta m}^{\delta h} \) covariance matrices are (semi-) definite positive matrices, we can conclude that model perturbations from posterior maximum likelihood model due to observational or model prediction uncertainties, follow an unbiased Gaussian probability distribution whose variances are bounded by the posterior marginal variances of the estimated model parameters. Thus, posterior model parameters marginal variances will control and bound the amplitude of the model perturbations due to uncertain observations, model prediction and prior information. The direction of the perturbations will depend on posterior correlations between estimated model parameters.
Text S2  Precision of theoretical/physical model and Tikhonov regularization

For the ill-posed quasi-static slip inversion problem, if we consider a fine enough fault slip discretization, slip on adjacent fault patches will be highly anticorrelated. Therefore, while it will be straightforward to estimate the average slip of both fault patches, it will be challenging to distinguish differences of slip between them, leading to the phenomenon known as checkerboarding in the fault slip estimates. Figure S.1 illustrates a typical example of the aforementioned anticorrelation between slip in two adjacent fault patches for the unregularized problem.

Figure S.1. Top panels show triangulated fault mesh with a selection of 4 pairs of adjacent fault slip patches (left) and slip sensitivity to Japan GEONET GPS network normalized to its minimum value (right). Bottom panels show 4 posterior conditional pdfs for perturbations of slip on each of the 4 pairs of adjacent fault patches shown in the top left panel. The slip perturbations are from the maximum likelihood model, ranging from 0 to 12 m. Conditional pdfs are normalized by their maximum likelihood value.

In order to understand Figure S.1, we will focus on the unregularized slip inversion problem, where prior information can be written as the homogeneous state of information[see Tarantola, 2005]. Thus, if we assume cartesian model parameters, \( f_{prior}(m) = \text{const.} \), a constant function in model parameter space. Let \( m^* \) be the maximum likelihood model of the posterior pdf of model parameters \( f_{post}(m) \). It can be shown that the a posteriori pdf for model parameters can be written as

\[
  f_{post}(\Delta m) = \text{const.} \cdot \exp \left( -\frac{1}{2} \Delta m^T P \Delta m \right)
\]  

(S.20)

where \( \Delta m = m - m^* \) is a perturbation from the maximum likelihood model, and \( P = G^T C^{-1} G \) is the Precision Matrix of the problem. Note that \( P \) could be taken as the inverse of the posterior covariance matrix of slip estimates \( m \) for the unregularized problem if the inverse of \( P \) is well defined. Each box at the bottom in Figure S.1 corresponds to the conditional pdf (normalized by its maximum likelihood value) for deviations of slip from its maximum likelihood value on two adjacent fault slip patches, \( \Delta m_i, \Delta m_j \), given that slip at the rest of the fault surface corresponds to their maximum
likelihood values. Equation (S.20) can be written in terms of the deviated values of slip as

\[
f_{\text{post}}(\Delta m_i, \Delta m_j \mid \Delta m_k = 0) = \text{const.} \cdot \exp \left( \frac{1}{2} \begin{bmatrix} \Delta m_i & \Delta m_j \end{bmatrix} \begin{bmatrix} P_{ii} & P_{ij} \\ P_{ji} & P_{jj} \end{bmatrix} \begin{bmatrix} \Delta m_i \\ \Delta m_j \end{bmatrix} \right) \tag{S.21}
\]

where \(P_{ij}\) is the precision matrix for \(m_i\) and \(m_j\) (the inverse of the posterior conditional covariance matrix of \(m_i\) and \(m_j\), given that the rest of model parameters have their maximum likelihood value).

Note that under the assumption of a fine enough discretization of the fault surface, the predicted surface displacements due to unit slip at adjacent fault patches will be very similar, i.e., the \(i\)-th and \(j\)-th columns of \(G\) will be very close. Then, as \(P_{ij} = G_i^T C^{-1} G_j\), where \(G_k\) is the \(k\)-th column of \(G\) (the data prediction due to a unit slip on \(k\)-th fault patch), \(P_{ij}\) will be similar to the norm \(\|W G_k\|_2^2\).

If we make the approximation that \(P_{ii} \approx P_{ij} \approx P_{jj}\), for a fine enough discretization, the matrix \(P_{ij}\) can be approximated by

\[
P_{ij} \approx P_{ii} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}
\]

with eigenvalue decomposition

\[
P_{ij} \approx P_{ii} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2P_{ii} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}
\]

By looking at the eigenvalue decomposition of \(P_{ij}\), \(2P_{ii}\) corresponds to the reciprocal of the posterior conditional variance in the positively correlated direction for \([\Delta m_i \quad \Delta m_j]^T\). Also we obtain that, for our approximation, the anticorrelated direction has infinite conditional variance. Note that for the case in which we do not make the approximation that \(P_{ii} \approx P_{ij} \approx P_{jj}\), for a fine fault discretization, we still obtain a much larger value of the conditional variance for the anticorrelated direction (as shown in Figure S.1), which is responsible for the "checkerboarding" problem of slip estimates.

A first attempt to solve the checkerboarding problem would be to simply minimize the difference between slip in adjacent fault patches, as done in first order Tikhonov regularization. Effectively, 1\(^{st}\) order Tikhonov regularization imposes a Gaussian prior on the difference \(\Delta m_i - \Delta m_j\) with variance \(1/\epsilon^2\), i.e., the prior information for a pair of adjacent fault patches can be represented by the pdf

\[
f(m_i, m_j) = \text{const.} \cdot \exp \left( -\frac{1}{2} \frac{(m_i - m_j)^2}{\epsilon^{-2}} \right) \tag{S.24}
\]

where \(m_i\) and \(m_j\) is the slip on adjacent fault patches. Figure S.2 shows the fdp corresponding to equation (S.24) for different values of the damping or regularization parameter \(\epsilon^2\).

**Figure S.2.** Prior pdf for 1\(^{st}\) order Tikhonov regularization of slip on a pair of adjacent fault patches. Each box shows the pdf for a different value of \(\epsilon^2\). The pdf’s are normalized by maximum likelihood value.
In the same manner, 2\textsuperscript{nd} order Tikhonov regularization minimizes the difference between slip in a given subfault with slip at all neighboring subfaults through minimizing the application of a scaled second order finite difference operator to fault slip. In this case, if a simple finite difference approximation of the Laplacian operator is used, the pdf describing prior information for slip at the i-th fault patch can be written as

\[
f_i(m_i|m_k, k \in I_i) = \text{const.} \cdot \exp \left(- \frac{1}{2} \frac{\left(m_i - \frac{1}{N_{I_i}} \sum_{k \in I_i} m_k\right)^2}{\varepsilon^2} \right)
\]  

(S.25)

where \(I_i\) is the set of indices of the fault patches adjacent to the i-th fault patch, and \(N_{I_i}\) is the number of fault patches that are adjacent to the i-th fault patch.

Note that when increasing the value of the damping parameter \(\varepsilon\), one decreases prior variance on the difference of slip \((\sigma^2 = \frac{1}{\varepsilon^2})\), thus imposing smoother models, as we allow for smaller differences of slip in adjacent fault patches. However, Tikhonov regularization imposes the prior constraints with the same precision \(\varepsilon^2\) throughout the fault surface, not accounting for the spatial variability of posterior model precision in the unregularized problem, which in Figure S.1 is shown to vary in a great manner throughout the fault surface.
Text S3 Equal Posterior Information Condition (E.P.I.C.) to calculate prior information variances for a spatially variable Tikhonov regularization

The linear least squares problem with Tikhonov regularization can be written as,

$$\min_m (Gm - d)^\top C^{-1}_x (Gm - d) + \varepsilon^2 (Hm)^\top (Hm)$$  \hspace{1cm} (S.26)

where \( m \) is the vector of unknown slip, \( G \) is the Green’s functions or design matrix, \( d \) the data vector, \( C_x \) the misfit covariance matrix (representing observational and model prediction uncertainties), \( H \) the regularization operator, and \( \varepsilon^2 \) is the regularization parameter.

As shown in the main body of this article, (S.26) is a particular case of the general least squares problem

$$\min_m (Gm - d)^\top C^{-1}_x (Gm - d) + (Hm - h^o)^\top C^{-1}_h (Hm - h^o)$$  \hspace{1cm} (S.27)

where \( h^o = 0 \), \( C_h = \sigma^2_h I \), \( I \) is the identity matrix and \( \sigma^2_h = \frac{1}{\varepsilon^2} \). Thus, the prior covariance matrix for the regularized quantity \( h = Hm \) is a diagonal matrix, where all prior variances on \( h \) are the same, i.e., assuming independent and identically distributed (i.i.d.) uncertainties on the quantity \( h \) to which prior information is posed.

Here we extend Tikhonov regularization to a case in which prior information on \( h \) represents uncertain quantities that are independent, but not identically distributed, achieving a spatially variable strength on the smoothing regularization on slip implied by higher order Tikhonov regularization. We identify 2 issues with Tikhonov constraint, the information gained on slip on faults is not homogeneous throughout the fault, as posterior variances on estimated slip vary throughout the fault, and that \( \varepsilon^2 \) is a regularization parameter that is difficult to interpret in terms of direct prior information. For instance, it may have units of spatial derivatives of slip, and the values to search for during the inversion might be not easy or direct to find.

Instead of searching directly for the optimum value of the regularization parameter \( \varepsilon^2 \) (or the variances of prior information \( \sigma^2_h \)), we ask for prior information variances that produce posterior information gain on slip, with respect to a base case (homogeneous state of information), that is uniform throughout the fault. As Gaussian uncertainties are assumed for all quantities, we show below that the posterior marginal variance serves as a measure of information gained on slip on a single fault patch. Then, for a given target posterior variance \( \sigma^2_m = \sigma^2_t \), we estimate prior information variances, a diagonal \( C_h \) matrix. Since we use \( \sigma^2_t \) as regularization parameter instead of \( \varepsilon^2 \), we refer to it as target posterior variance on model parameters (slip in this case).

A measure of the information content on a continuous random variable is given by the differential entropy. If \( f(x) \) is the PDF of the continuous random variable \( X \), its differential entropy is defined as,

$$h(x) = - \int_X f(x) \log(f(x)) \, dx$$  \hspace{1cm} (S.28)

Also, Kullback-Leibler divergence gives us a measure of the “distance” from the PDF of the random variable \( x \) \( f(x) \) to another PDF associated to the same random variable \( (\mu(x)) \),

$$D_{KL}(f||\mu) = - \int_X f(x) \log \left( \frac{f(x)}{\mu(x)} \right) \, dx$$  \hspace{1cm} (S.29)

If the variable \( x \) is a Cartesian parameter and \( \mu(x) \) is a uniform PDF representing its state of homogeneous information [see Tarantola, 2005], \( \mu(x) \) is constant, and the Kullback-Leibler divergence of \( f(x) \) with respect to \( \mu(x) \) is proportional to the differential entropy of \( x \). From here on, we assume that model parameters are all cartesian in nature, thus the pdf representing the state of homogenous information can be approximated by a wide enough uniform distribution.
Each component of \( \mathbf{m} \) inferred by solving the problem (S.27), follows a normal pdf. Then, the differential entropy associated to the posterior marginal pdf of each model parameter becomes,

\[
h(m_i) = \frac{1}{2} \ln \left( 2\pi e \sigma_{m_i}^2 \right) \tag{S.30}
\]

and depends only on its posterior marginal variance \( \sigma_{m_i}^2 \). Then, setting the Equal Posterior Information Condition (EPIC) on \( \mathbf{m} \), is equivalent to imposing that posterior marginal variances are equal to \( \sigma_{m_i}^2 \) for all components of \( \mathbf{m} \). The application of the EPIC, for a given target posterior variance \( \sigma_t^2 \), allows the estimation of the prior information variances (or regularization weights) that modulate the strength of a slip smoothing regularization operator \( \mathbf{H} \) (e.g., higher order Tikhonov). As a consequence, posterior co-variances (off diagonal elements of the posterior covariance matrix \( \tilde{\mathbf{C}}_m \) of model parameters) or correlations between closer model parameters, will be in turn modified in order to produce a uniform posterior marginal variances structure. From an information theory perspective, we seek to equalize the information content gained on each fault patch by manipulating the amount of information shared with its neighbors, i.e., the mutual information between neighboring model parameters (as mutual information is related to Pearson correlation under normality assumption).

For the damped least squares problem in (S.27), posterior model parameter estimates \( \tilde{\mathbf{m}} \) follow a normal distribution with posterior covariance matrix

\[
\tilde{\mathbf{C}}_m = \left[ \mathbf{G}^\top \mathbf{C}_x^{-1} \mathbf{G} + \mathbf{H}^\top \mathbf{C}_h^{-1} \mathbf{H} \right]^{-1} \tag{S.31}
\]

Imposing the EPIC allows to decompose \( \tilde{\mathbf{C}}_m \) into a constant diagonal matrix and a symmetric Hollow matrix, \( \tilde{\mathbf{C}}_m = \sigma_t^2 \mathbf{I} + \tilde{\mathbf{C}}_m^\text{eff} \). Thus, imposing that all the posterior variances are equal to \( \sigma_t^2 \), while leaving the covariances of \( \tilde{\mathbf{C}}_m \) (i.e., the elements of \( \tilde{\mathbf{C}}_m^\text{eff} \)) to be determined by the structure of \( \mathbf{G}, \mathbf{H} \) and \( \mathbf{C}_x \).

For higher order Tikhonov regularization, \( \mathbf{H} \) is a linear smoothing operator that is proportional to the finite difference approximation of the first or second order derivatives (i.e, gradient or Laplacian) of slip, that effectively induces correlations in the posterior model parameter estimates. Then, we only need to have spatially variable variances of the independent uncertainties of prior information on \( \mathbf{h} \) in order to modulate the posterior variances on model parameters and their correlation. Therefore, for adopting the EPIC, we use a diagonal covariance matrix for prior information,

\[
\mathbf{C}_h = \delta_{ij} \sigma_{h_i}^2 \tag{S.32}
\]

in which each \( \sigma_{h_i}^2 \) will be determined using the EPIC for a given target posterior variance for the estimated model parameters.

Recall the solution of problem (S.27) is,

\[
\tilde{\mathbf{m}} = \left( \mathbf{G}^\top \mathbf{C}_x^{-1} \mathbf{G} + \mathbf{H}^\top \mathbf{C}_h^{-1} \mathbf{H} \right)^{-1} \left( \mathbf{G}^\top \mathbf{C}_x^{-1} \mathbf{d}^{\text{obs}} + \mathbf{H}^\top \mathbf{C}_h^{-1} \mathbf{h}^o \right) \tag{S.33}
\]

with posterior covariance matrix

\[
\tilde{\mathbf{C}}_m = \left[ \mathbf{G}^\top \mathbf{C}_x^{-1} \mathbf{G} + \mathbf{H}^\top \mathbf{C}_h^{-1} \mathbf{H} \right]^{-1} = \left( \mathbf{P} + \mathbf{H}^\top \mathbf{C}_h^{-1} \mathbf{H} \right)^{-1} \tag{S.34}
\]

where \( \mathbf{P} = \mathbf{G}^\top \mathbf{C}_x^{-1} \mathbf{G} \) is the precision matrix of the unregularized problem.

Now, given a target posterior variance \( \sigma_t^2 \), using the EPIC to obtain the prior information variances \( \sigma_{h_i}^2 \) in (S.32), corresponds to solve the following nonlinear equations,

\[
\left[ \tilde{\mathbf{C}}_m \right]_{ii} = \left[ \left( \mathbf{P} + \mathbf{H}^\top \mathbf{C}_h^{-1} \mathbf{H} \right)^{-1} \right]_{ii} = \sigma_t^2 \tag{S.35}
\]

where, \( \left[ \tilde{\mathbf{C}}_m \right]_{ii} \) is the i-th diagonal element of the posterior covariance matrix on slip, \( \mathbf{C}_{h_ii} = \delta_{ij} \sigma_{h_i}^2 \) the covariance matrix of prior information, and \( \sigma_t^2 \) the target posterior variance for slip. Note that \( \sigma_t^2 \)
is now used as a regularization parameter. Therefore, as part of the model class selection scheme, we first define a set of candidate values of \( \sigma^2_t \). For each candidate, one must determine \( \mathbf{C}_h \), to then obtain posterior model parameter estimates \( \mathbf{m} \). Subsequently, the optimum value of \( \sigma^2_t \) can be obtained using L-curve, cross-validation, or any other standard model class selection technique.

Note that in previous equations, a single target posterior variance \( \sigma^2_t \) is used, being adequate when inverting for an unknown scalar field discretized through a spatial or temporal domain (fault surface in this case) into the elements of \( \mathbf{m} \). An example would be the inversion of the values of quasi-static fault slip in which the direction of slip is set to be known and fixed (as for the inversions constrained with synthetic data in this work). Instead, if a vector field is the unknown of the inverse problem (for instance, searching for slip in dip-slip and strike-slip orthogonal directions), one could consider using different target posterior variances for each component of the vector field being inverted. For instance, if a vector field with 2 dimensions whose directions are defined by \( \mathbf{v}_1 \) and \( \mathbf{v}_2 \) is considered, one only need to replace \( \sigma^2_t \) on the right hand side of the EPIC nonlinear equations (S.35) by

\[
\sigma^2_t = \begin{cases} 
(\sigma^2_t^{v_1})^2 & \text{if } i \in I_{v_1} \\
(\sigma^2_t^{v_2})^2 & \text{if } i \in I_{v_2}
\end{cases}
\]  

(S.36)

where \( \sigma^2_t^{v_k} \) and \( I_{v_k} \) corresponds to the target posterior variance and the set of indices of the vector of model parameters \( \mathbf{m} \) associated to the \( k \)-th direction (\( \mathbf{v}_k \)) of the vector field being estimated. In this example, one could evaluate several pairs of regularization parameters \( (\sigma^2_t^{v_1}, \sigma^2_t^{v_2}) \) using L-curve, cross-validation, or any other model class selection technique in order to obtain a solution to the inverse problem. Other option, as done in this work for the example of the inversion of co-seismic slip for the Tohoku-Oki earthquake, is to search for the target posterior variance for rake parallel slip, while leaving fixed the one for the rake perpendicular slip, for example, at 1 meter.

On the Independence of Prior Information on Data and Model Parameters

Prior information must be obtained independently of the data used for the inversion. Then, the probability density functions (pdf’s) representing prior information on data and model parameters are not related and the joint prior pdf [eq 1.82 in Tarantola, 2005] can be written as the product of the prior pdf’s for observable data and model parameters.

In our method, we define prior information in a quantity \( \mathbf{h} \), assuming a normal distribution with mean \( \mathbf{h}^0 \) and covariance matrix \( \mathbf{C}_h \). While we state that for most use cases, the mean value will be chosen as \( \mathbf{h}^0 = \mathbf{0} \), we define \( \mathbf{C}_h \) as diagonal matrix when defining prior information. The consequence of such definition is that \( \mathbf{C}_h \) is defined in a parametric way as

\[
\mathbf{C}_h = \begin{bmatrix} 
\sigma^2_{h_1} & 0 & 0 & 0 & \ldots & 0 \\
0 & \sigma^2_{h_2} & 0 & 0 & \ldots & 0 \\
0 & 0 & \sigma^2_{h_3} & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 0 & \sigma^2_{h_{N_h}} 
\end{bmatrix}
\]

where, \( N_h \) is the number of regularization constraints, and the diagonal elements of \( \mathbf{C}_h \) are defined as unknown variances of the uncorrelated prior information on \( \mathbf{h} \), to be determined, as we generally lack the knowledge of the values for such variances.

Defining parametric variances of prior information is widely accepted and used when such quantities lack prior information. For instance, as thoroughly exposed in our manuscript, Tikhonov regularization defines \( \mathbf{C}_h = \mathbf{e}^{-2} \mathbf{I} \), where \( \mathbf{I} \) is an identity matrix and \( e \) is often called the regularization parameter or damping constant to be determined. Tarantola [2005] (Examples 5.5 and 5.13) also
has an example defining a prior model covariance matrix as

$$[C_h]_{ij} = [C_m(\lambda)]_{ij} = \sigma^2 e^{-\frac{d_{ij}}{\lambda}}$$

where $\sigma^2$ is a prior variance on $m$, typically fixed to some justifiable value, $d_{ij}$ is a measure of distance between $i$–th and $j$–th discretized model parameters and $\lambda$ is a prior, unknown, correlation length, defined as a regularization parameter to be determined. As an example, such methodology was successfully used by Radiguet et al. [2011] to determine the slip distribution of a slow slip event in Mexico.

As is the case for both aforementioned examples and for our work, the unknown parameters in $C_h$ define a family or class of solutions of the inverse problem that depend on such unknown parameters $(\{\sigma_h_i\}_{i=1}^{N_h})$ in our work, $\varepsilon$ for Tikhonov regularization and $\lambda$ for Tarantola [2005], Radiguet et al. [2011]). For Tikhonov regularization and Tarantola [2005], model class selection methods include L-curve (e.g., [Hansen and O’Leary, 1993], Cross Validation – CV [Wahba, 1990], Generalized Cross Validation - GCV [Craven and Wahba, 1979], Akaike Bayesian Information Criterion - ABIC [Akaike, 1980], and Bayesian Model Class Selection [e.g. Muto and Beck, 2008], and are used to determine the “best” value of the parameters of the prior covariance matrix $C_h$. Given the generally nonlinear dependence between the inverse problem solution and the elements of $C_h$, the model class selection procedure is typically carried out using a grid-search approach, i.e., performing inversions for several values of the regularization parameter ($\varepsilon$ or $\lambda$). Then, the prediction of each of the estimated posterior models is compared with the data and/or regularized quantity to select the best regularization parameter value using a criterion defined by the chosen model class selection method.

In our work, we could have pursued a similar procedure. However, as we define a large ($N_h$) number of unknown parameters in our diagonal $C_h$ matrix, searching for the best set $(\{\sigma_h_i\}_{i=1}^{N_h})$ using a grid-search scheme would have been impractical due to computational and time constraints.

In our work, instead of a blind and exhaustive search for plausible values for all $\sigma_h_i$, we only search for $\sigma_h_i$ values that comply with the EPIC. Here, we find it useful and practical to ask the question: “From the family of posterior models defined by an unknown diagonal prior covariance matrix $C_h$, what is the subset of posterior models that have spatially homogeneous variances for all (or ad-hoc subsets of) model parameters?” Then, in our paradigm, the EPIC is thought as a complement of the standard model class selection methods, by selecting posterior models that have a particular uncertainty structure. The EPIC defines a nonlinear relationship between our selection of a diagonal parametric $C_h$ and the structure (or shape) of posterior model parameters variances. Thus, defining the latter as a straightforward, simplified and reduced set of regularization parameters to be determined using a complementary choice among standard model class selection methods (e.g., L-curve, CV, GCV, ABIC). In this sense, the EPIC acts as filter of all possible posterior models implied by data and prior information with unknown diagonal covariance matrix $C_h$, as a complement of a choice of standard model class selection method, and not used to formally define the prior state of information on model parameters through the quantity $h$.

**Implementation of the EPIC for Tikhonov Regularization**

In order to implement the EPIC to calculate $\sigma^2_h$, we use a nonlinear least squares estimation scheme. Particularly we use Scipy non linear least squares optimization software, in which we use the Trust Region Reflective constrained optimization method [Coleman and Li, 1996]. In order to avoid numerical problems, we limit the ratio between minimum and maximum possible values of $\sigma^2_h$, so calculations have enough significative digits.

Here we solve the optimization problem,

$$\min_k \varphi(\sigma^2_{h_k}) = \sum_i F_i^2$$  \hspace{1cm} (S.37)
where

\[ F_i = \left[ \hat{C}_m \right]_{ii} - \sigma_i^2 \]  
⇔ \[ F_i = \left[ (P + H^T C_h^{-1} H)^{-1} \right]_{ii} - \sigma_i^2 \]  

(S.38)

where \( \sigma_i^2 \) is the target posterior variance on \( m \).

Recall, from (S.32) we have that,

\[ C_{h_{ij}}^{-1} = \delta_{ij} \frac{1}{\sigma_{h_{ii}}} \]  

(S.40)

Now as we need to calculate both \( \sigma_{h_{ii}}^2 \) and \( \frac{1}{\sigma_{h_{ii}}} \), we model \( \sigma_{h_{ii}}^2 \) as a strictly positive quantity \( (\sigma_{h_{ii}}^2 > 0) \), i.e., as a Jeffreys quantity. Thus, in order to avoid numerical problems, such as dividing by a very small value of \( \sigma_{h_{ii}}^2 \), we perform a variable change: \( \beta_i = \ln \left( \frac{1}{\sigma_{h_{ii}}} \right) \)

\[ \Rightarrow \beta_i = \ln \left( \frac{1}{\sigma_{h_{ii}}} \right) \Leftrightarrow \sigma_{h_{ii}}^2 = e^{-\beta_i} \]  

(S.41)

then,

\[ C_{h_{ij}}^{-1} = \delta_{ij} e^{\beta_i} \]  

(S.42)

and

\[ F_k = F_k(\beta) \]  

(S.43)

Now in order to implement the nonlinear least squares algorithm, we need to calculate the Jacobian matrix of \( F(\beta) \),

\[ J_{Fkl} = \frac{\partial F_k}{\partial \beta_l} \]  

(S.44)

Let’s define

\[ A = A(\beta) = P + H^T C_h^{-1}(\beta) H \]  

(S.45)

then

\[ F_k(\beta) = [A^{-1}(\beta)]_{kk} - \sigma_m^2 \]  

(S.46)

and the Jacobian of \( F(\beta) \)

\[ \frac{\partial F_k}{\partial \beta_l} = \left[ \frac{\partial A^{-1}}{\partial \beta_l} \right]_{kk} = \left[ -A^{-1} \frac{\partial A}{\partial \beta_l} A^{-1} \right]_{kk} \]  

(S.47)

for the values of \( \beta \) in which \( A^{-1} \) is well defined, so that there exist a least squares solution to the inverse problem.

Now,

\[ \frac{\partial A}{\partial \beta_l} = \frac{\partial P^0}{\partial \beta_l} + H^T \frac{\partial C_h^{-1}(\beta)}{\partial \beta_l} H \]  

(S.48)

and,

\[ \frac{\partial C_{h_{ij}}}{\partial \beta_l} = \frac{\partial}{\partial \beta_l} (e^{\beta_j}) \]

\[ = \delta_{ij} \frac{\partial e^{\beta_j}}{\partial \beta_l} \]

\[ = \delta_{ij} e^{\beta_j} \frac{\partial \beta_j}{\partial \beta_l} \]

\[ \Rightarrow \frac{\partial C_{h_{ij}}}{\partial \beta_l} = \delta_{ij} \delta_{jl} e^{\beta_j} \]  

(S.49)
Now from (S.48) and (S.49) we have,

$$\frac{\partial A_{ij}}{\partial \beta_l} = H_{ri} \frac{\partial c^{-1}_{hi}}{\partial \beta_l} H_{sj}$$

$$= H_{ri} \delta_{rl} \delta_{ji} e^{\beta_l} H_{sj}$$

$$= H_{ri} \delta_{rl} e^{\beta_l} H_{ij}$$  \hspace{1cm} (S.50)

$$\Rightarrow \frac{\partial A_{ij}}{\partial \beta_l} = e^{\beta_l} H_{il} H_{lj}$$  \hspace{1cm} (S.51)

Now from (S.47),

$$\frac{\partial F_k}{\partial \beta_l} = \left[ \frac{\partial A^{-1}}{\partial \beta_l} \right]_{kk}$$

but,

$$\frac{\partial A^{-1}_{qi}}{\partial \beta_l} = -A^{-1}_{qi} \frac{\partial A_{ij}}{\partial \beta_l} A^{-1}_{ji}$$

$$= -A^{-1}_{qi} e^{\beta_l} H_{il} H_{lj} A^{-1}_{ji}$$

$$= -e^{\beta_l} A^{-1}_{qi} H_{il} H_{lj} A^{-1}_{ji}$$  \hspace{1cm} (S.52)

lets define

$$B = H \cdot A^{-1}$$  \hspace{1cm} (S.53)

then,

$$B_{lt} = H_{il} A^{-1}_{jt}$$

also,

$$B_{iq} = H_{ij} A^{-1}_{iq} \left\langle \text{ but } A^{-1}_{jq} = A^{-1}_{ij} \right\rangle$$

$$= A^{-1}_{iq} H_{li}$$  \hspace{1cm} (S.54)

then,

$$\frac{\partial A^{-1}_{qi}}{\partial \beta_l} = -e^{\beta_l} B_{lt} B_{it} \left\langle \text{ No summation over } q, t, l \right\rangle$$  \hspace{1cm} (S.55)

Now recall that we only need the diagonal terms for the Jacobian of \(F\), then \(q = t = k\)

$$\frac{\partial F_k}{\partial \beta_l} = \frac{\partial A^{-1}_{kk}}{\partial \beta_l} = -e^{\beta_l} B_{lk}^2$$

then the Jacobian of \(F\) becomes:

$$\frac{\partial F_k}{\partial \beta_l} = -e^{\beta_l} B_{lk}^2$$  \hspace{1cm} (S.56)

For coding purposes, if we define a diagonal matrix

$$D_{ij} = e^{\beta_i} \delta_{ij}$$  \hspace{1cm} (S.57)

and the matrix

$$Q = D \left[B \otimes B\right]$$  \hspace{1cm} (S.58)

where \(\otimes\) indicate element-wise multiplication, then

$$Q_{lk} = D_{li} B_{rk}^2$$

$$= e^{\beta_i} \delta_{ri} B_{rk}^2$$

$$= e^{\beta_l} B_{lk}^2$$
An the closed expression for the Jacobian of $F$ becomes,

$$
\Rightarrow J_F = -Q^T \quad \text{(S.59)}
$$

$$
\Leftrightarrow J_{F_{ks}} = -Q_{ks}^T \quad \text{(S.60)}
$$

$$
= -Q_{sk} \quad \text{(S.61)}
$$

$$
= -e^{\beta_s} B^2_{sk} \quad \text{(S.62)}
$$

Nonlinear inversion algorithms often require an initial solution that is close to the optimum being search, one could search for $\sigma_{h_i}^2$ in 1 or 2 steps. In both cases one can set the initial solution to constant value $\sigma_{\alpha}^2$ (such that initial $\sigma_{h_i}^2 = \sigma_{\alpha}^2 \ \forall i$). In the 1 step approach, the final solution is searched directly from the initial one using the previously described methodology. The 2 step approach consist in first iterating to find the best $\sigma_{h_i}^2$ for the case of $\sigma_{h_i}^2 = \sigma_{\alpha}^2 \ \forall i$ (i.e., assuming that prior information variances are the same throughout the fault, like as in standard Tikhonov regularization), to then use the previously described algorithm to find $\sigma_{h_i}^2 \ \forall i$, using $\sigma_{h_i}^2 = \sigma_{\alpha}^2 \ \forall i$ as initial model for the nonlinear optimization algorithm. During our testing, for both approaches we obtain the same solutions, but the convergence time of the algorithm is significantly reduced when the 2 stage approach is used.

Regarding the computation of prior information variances ($C_h$), the EPIC is well achieved for ET2 regularization, with RMS residuals in the order of machine precision for the whole range (0.01 to 25 meters) of tested target posterior standard deviation ($\sigma_t$). Also, as the estimation of $C_h$ involves solving the nonlinear problem described in equation (14) of the main body, we tested the procedure for several initial guesses obtaining stable results of $C_h$. For the synthetic tests discussed here, the number of model parameters is 1936, and the calculation of $C_h$ takes at most 15 seconds on a 20 core workstation, requiring only 10 to 20 iterations of the nonlinear solver. Estimation of $C_h$ when using ET1 regularization is more challenging. Here, the EPIC is well achieved for most of the usable cases of $\sigma_t$ with RMS misfits of the order of machine precision, but for $\sigma_t > 2$ meters, RMS residuals start to degrade, reaching values that, while still low, are of the order of $1e-4$ meters for $\sigma_t = 25$ meters. As an averaged view for all the tested values of $\sigma_t$, when solving the EPIC for ET1 the nonlinear solver performs hundreds of iterations, resulting in a scenario in which the time needed to compute the solution of the EPIC for ET1 is about 1 to 2 orders of magnitude longer than the time needed for ET2. See Text S14 to see the final values of the cost function of the EPIC for the examples developed in this work.

**Selection of Range of Possible Values for the Target Posterior Standard Deviation $\sigma_t$**

Please see Supplementary Text S14 on page 98.
Text S4 Japan Trench Subduction Model Definition, Fault Slip Parameterization and Regularization Operators

Fault Geometry of the Area Affected by the 2011 (Mw 9.0) Tohoku-Oki Earthquake

We construct a 3D triangulated fault surface based on the fault geometry published in Simons et al. [2011]. The geometry is built based on interpretations of seismic reflection and refraction profiles, providing constraints for the first 30 km depth. Published seismic tomographic and earthquake hypocenters. See Simons et al. [2011] for further detail. For modeling purposes, we performed a triangulation scheme on the fault geometry in order to construct a 3D triangulated surface to model the complex 3-D geometry of the megathrust. We use a total of 1936 triangular patches, with quasi-homogeneous characteristic dimension ranging between 7 and 15 km.

Elastic Structure

Figure S.3 shows the elastic structure obtained from Simons et al. [2011]. The media is represented by a 1D layered elastic structure with properties computed by averaging the 3D tomography from NIED (Japan National Research Institute of Earth Science and Disaster Prevention) at the epicentral region of the Tohoku-Oki earthquake for the first 16 km depth. The elastic structure below 16 km depth is taken from Takahashi et al. [2004]. The elastic Green’s Functions relating an unit dislocation on a given fault patch with the prediction of displacements at the GPS sites are computed by approximating a finite triangular elastic dislocation in the layered elastic half space. Each triangular dislocation is represented by summation of a set of point sources evenly distributed over the triangle’s surface in order to simulate a finite dislocation over the triangle.

![Elastic Structure](image)

**Figure S.3.** 1D elastic structure for Japan Trench Megathrust from Simons et al. [2011]

Fault Slip Parameterization

For the extensive suite of fault slip inversions constrained by synthetic data used to compare the effect of the different regularization operators exposed in this work, we model slip to occur in a direction that is parallel to a reference direction, with azimuth of 104°, given by the orientation of the average slip vector defined by the focal mechanism of the Tohoku-Oki (Mw 9.0) event from Global CMT project. As all triangular fault patches of the 3D triangulated fault surface have different
For the slip inversion constrained by GPS observations of co-seismic offsets produced by the Tohoku-Oki (MW 9.0) mainshock event, we model slip to occur in two directions, a rake parallel direction and a rake perpendicular direction. The rake parallel direction is defined as described in the previous paragraph, defining a rake angle ($\lambda_{\text{par}}$) that is different for each triangular fault patch. The rake perpendicular direction for each triangular fault patch is defined as the slip vector associated to the rake angle $\lambda_{\text{perp}} = \lambda_{\text{par}} + 90^\circ$.

Slip throughout the fault surface is represented using piecewise constant basis functions, in which slip is constant within each fault patch discretization and discontinuous between adjacent fault patches.

**Model Prediction Uncertainties**

We acknowledge that the defined forward model is an approximation of the true physics underlying the slip inversion problem, for instance, due to uncertain fault geometry and elastic structure. However, we do not include epistemic uncertainties in our comparative analysis of regularization schemes, as we constrain slip models with synthetic datasets, thus considering an exact forward model in this case.

When using real Tohoku-Oki earthquake data, we include epistemic uncertainties using the approach of Minson et al. [2014]. Here, model prediction uncertainties are assumed to follow an unbiased normal distribution with covariance matrix

$$C_p = \alpha^2 \text{diag}(d^2)$$

where $\text{diag}(d^2)$ is a diagonal matrix formed by the squared values of the data vector $d$, and $\alpha$ is a proportionality constant. We set $\alpha = 0.1$ following Minson et al. [2014] as we compare our results with those of Minson et al. [2014].

**Definition of Laplacian and Gradient Operators for the Triangulated Fault Mesh**

The prior information imposed in the regularized linear least squares problem (equation (S.27)), prescribed some restrictions in a quantity $h = Hm$. When prior information is expressed using 1st and 2nd order Tikhonov regularization, the operator $H$ is proportional to a finite difference approximation of the gradient ($\nabla$) or the Laplacian ($\nabla^2$) differential operators defined at the triangulated mesh representing the fault surface.

As slip is assumed to be constant within each fault patch discretization, we use an element wise approximation of the gradient and Laplacian differential operators.

For 1st order Tikhonov, we use an approximation of the gradient operator ($\nabla$), based on the differences of slip between two adjacent fault patches, where two triangular fault patches are defined to be adjacent if they are different and share a unique edge. Thus, the regularization operator, applied to the unknown model parameters, can be written as:

$$[Hm]_i = m_{k(i)} - m_{q(i)}$$  \hspace{1cm} (S.63)

where $i$ indicates the $i$-th gradient operator between $k(i)$-th and $q(i)$-th fault patches. Here $k(i)$ and $q(i)$ define mappings relating the index of the gradient operator and a pair of indices $(k, q)$ corresponding to two adjacent triangles. The mappings are defined so that a given pair of triangles is defined without repetition, thus, a given pair and its permutation are not present at the same time. Note that, while the number of columns in $H$ is equal to the number of model parameters ($\dim(m)$),...
the number of rows in $H$ is equal to the number of interior edges of the triangulated mesh representing the fault surface, which is generally larger than the number of triangular fault patches.

For $2^{nd}$ order Tikhonov, we use an approximation of the Laplacian operator ($\nabla^2$), based on the differences between slip on a given fault patch and slip on its adjacent fault patches. Thus, the regularization operator, applied to the unknown model parameters, can be written as:

$$[Hm]_i = N^i_A m_i - \sum_{k \in I_i} m_k$$  \hspace{1cm} (S.64)

where $i$ is the index of the fault patch where the slip Laplacian operator is applied, and $I_i$ is a set of $N^i_A$ indices of the fault patches that are adjacent to the $i$-th triangular fault patch ($N^i_A \leq 3$).

In both cases, $1^{st}$ and $2^{nd}$ order Tikhonov regularization, the differential operators, as shown, imply boundary conditions that are consistent with the smoothing condition of the regularization operator without adding additional constraints to the problem. Thus, the used boundary condition corresponds to a Neumann boundary condition, implying the assumption that the field that is being search during the inversion (slip in this case) in a vicinity just outside the parameterized fault surface is equal to the field modeled at the nearest border triangular element.

**Generalized Cross Validation (GCV)**

As explained in the main body of the manuscript, for the synthetic tests we use Generalized Cross Validation (GCV) to select the best model from the family of solutions defined by the regularization parameters. It can be shown [e.g., Aster et al., 2013] that the best model selected using GCV, corresponds to the one that minimizes the loss function $g(\alpha)$:

$$g(\alpha) = \frac{1}{N} \sum_{k=1}^N \left( \frac{[G\tilde{m}_\alpha]_k - d_k}{1 - [GG^\dagger]_k} \right)^2$$

where, $\alpha$ represents a specific value of one of the damping parameters discussed in the main body of the article ($\Gamma, \lambda, \alpha_t$), $\tilde{m}_\alpha$ is the vector of estimated model parameters given a regularization parameter $\alpha$ and $G^\dagger$ is the generalized inverse of $G$. For the general least squares problem described in the main body, considering $h^0 = 0$ as in all the examples developed in our work, the generalized inverse becomes:

$$G^\dagger = \left(G^\dagger C_x^{-1}G + H^\dagger C_h^{-1}H\right)^{-1} G^\dagger C_x^{-1}$$
**Text S5 Posterior variances and correlations**

In Text S1 we show that perturbations on estimated model parameters due to uncertainties on observations, model prediction and prior information follow a normal distribution with a covariance matrix equal to the posterior covariance matrix of the estimated model parameters ($C_{\text{m}}$). Uncertainties on posterior model estimates can be interpreted in two ways that are tied to each other, the posterior marginal variances and correlations (or covariances) between model parameters.

The square root of the i-th component of the diagonal of the posterior model covariance matrix $\sqrt{C_{\text{m}_{ii}}}$, gives the marginal variance of the i-th model parameter ($m_i$), which is a measure of the amount of the perturbation the model might suffer due to uncertain observations, model prediction and prior information. Thus, $\sqrt{C_{\text{m}_{ii}}}$ can be used to define confidence intervals for each estimated model parameter.

On the other hand, off diagonal elements of $\tilde{C}_{m}$ provide information on linear dependency of estimated model parameters. An easier to interpret quantification of linear dependence between model parameter estimates $m_i$ and $m_j$, is given by the Pearson correlation coefficient,

$$\rho_{m_{ij}} = \frac{\tilde{C}_{m_{ij}}}{\sqrt{\tilde{C}_{m_{ii}} C_{m_{jj}}}}$$

which has extreme values of -1, 1 for negatively and positively fully correlated model parameters, respectively, and a null value when two model parameters are uncorrelated. In that sense, a set of model parameters that are all positively or negatively correlated, will tend to be perturbed from the estimated (mean) model, in the same direction due to uncertainties in observations, model prediction and prior information, with perturbation amplitudes proportional to the posterior marginal standard deviations.

When model parameters are defined from a spatial discretization of a continuous scalar field (a component of fault slip, for instance), a measure of correlation between model parameters can be used to estimate a spatial correlation length, defining the location and size of regions in which model parameters are highly correlated, regions being potentially perturbed from the estimated (mean) model in the same direction upon the presence of uncertainties. One form of spatial correlation can be given by a distance exponentially decaying correlation function:

$$\rho_{ij} = e^{-\frac{d_{ij}}{\lambda_i}}$$

where $d_{ij}$ is a measure of the distance between i-th and j-th discretized model elements, $\lambda_i$ is the spatial correlation length for the i-th discretized element, and $p$ is the exponent of the $L_p$ norm used to define the distances $d_{ij}$ [e.g., Tarantola, 2005; Radiguet et al., 2011]. $\lambda_i$ measures how fast the correlation between model parameter $m_i$ and any other model parameter $m_j$ decays as a function of their distance $d_{ij}$. In this work, based on a trial and error approach, we decided to use $p = 2$, and for each slip model parameter associated to a discretized element of the fault mesh, we determine the correlation length using a grid-search method by looking for 500 possible values of each $\lambda_i$ between 2 km and 250 km. An example of estimated a posteriori correlation length can be seen in Figure S.5, corresponding to the one associated to the best solution of an inversion test with synthetic data generated using a checkerboard slip distribution.

As when assessing uncertainty of the results of an inverse problem, we need to care not only on posterior marginal variances but also about the covariances or correlations, we compare those quantities in Figure S.4 for inversion using all the regularization schemes analyzed in this work. Figure S.4 shows 8 panels ((a) to (h)), (a) shows the True model of slip and the square root of model precision (diagonal terms), (b) to (h) panels have 4 boxes showing (from left to right) prior information variances ($\sqrt{(C_m)}$), posterior model parameter standard deviations ($\sqrt{(C_{\text{m}})}$, labeled as $\sqrt{(C_{\text{m}}^{\text{post}}})$ for explicitness), posterior correlation lengths, and estimated model parameters. The
Label at the left side of panels (b) to (h) indicates the regularization scheme used to estimate model parameters, and hence, posterior covariance matrix, correlation length and setup of prior information.

Colorbars for $\sqrt{C_h}$ and $\sqrt{C_{\text{Post}}^m}$ show the base 10 logarithm of the quantities normalized to their minimum value. Colorbars for slip distributions are saturated at 0 and 100 cm values, so values higher or lower will show with the colors corresponding to 100 and 0 cm, respectively.

Each field plotted has an inset at its bottom right corner showing the minimum and maximum value of the plotted field. If an inset is present at the top-left corner, it indicates the ratio between maximum and minimum value of the plotted field (the latter without taking the logarithm if such quantity if plotted).

A simple inspection of Figure S.4, reveals that standard regularization approaches induce spatially variable posterior marginal variances and correlation lengths of model parameters, making it difficult to separate both effects for interpretation of model estimates uncertainties. On the contrary, when using the EPIC for higher order Tikhonov regularization (EPIC Tikhonov regularization), posterior variances are the same for all model parameters, being only the correlation length a spatially variable quantity. By fixing posterior marginal variances, one can directly analyze the spatial distribution of the uncertainties of model estimates in terms of their covariances or correlations. Here, the correlation length calculated for a particular estimated model parameter defines a region (of the discretization used to define the vector $m$) in which model perturbations upon realizations of misfit and regularization uncertainties are highly correlated. Thus, defining a proxy for a measure of spatial resolution of slip estimates, as observational and model prediction uncertainties induce perturbations on estimated model parameters that tend to occur in similar direction and amplitude at the region defined by the calculated correlation length on a given fault element. Also, one could easily separate model parameters into sets classified with respect to their significance relative to the marginal variances or correlation lengths of model estimates.
Figure S.4. Comparison of posterior variances and correlation lengths for slip inversion using Noise 2 synthetic data generated with a checkerboard slip models, and using all the regularization operators compared in this study. Bottom right inset in figure boxes show minimum and maximum value of the plotted field. When present, top-left inset shows the ratio between maximum and minimum value of the plotted field. See Text S5 for further figure explanation.
Figure S.5. Slip inversion for synthetic data generated with a checkerboard slip mode. From top to bottom, the first column of models show the true model used to generate the synthetic data, and the inversion solution constrained with Noise2 data, considering the different regularization approaches analyzed in this study. Second column shows estimated posterior correlation length for slip. 3rd to 7th columns show posterior Pearson correlation coefficient between slip in a selected fault patch (indicated at first row) and slip at the rest of the fault. Each red circle has a radius equal to the estimated slip spatial correlation length at the analyzed fault discretization.
Text S6  Synthetic inversions: Checkerboard slip models

We first generate synthetic data, for all GPS stations of the GEONET network [Sagiya et al., 2000] in north east Japan, using checkerboard style slip distributions with values ranging between 0 and 1 m, considering patterns with scales that vary from 60 to 140 km peak to peak characteristic length. Subsequently, we generate three unbiased noise cases, named Noise1, Noise2 and Noise3, using realizations of a multivariate normal distribution with zero mean and covariance matrix representing uncertainties on crustal displacement synthetic data (0.5 cm and 1 cm for horizontal and vertical components respectively) at the location of the GPS instruments used in this work (same as in Simons et al. [2011]).

Figures S.6 to S.12 show estimates of rake parallel slip (see Text S4) for all cases of synthetic data perturbed with all noise cases, each figure showing results for a different regularization operator. The first row of models show the slip distribution used to compute synthetic data (labeled as True model), and subsequent rows show slip estimates for synthetic data in which Noise1, Noise2 and Noise3 cases of random noise have been added.

Colorbars show slip in meters, up to 10% above maximum slip (1.1 m, gray tones) and up to 10% of back-slip (-0.1 m, pink tones). Each slip model is shown in a box with two insets. The top-left inset shows the moment magnitude of such slip distribution and the bottom-right inset shows minimum and maximum value of slip (being positive and negative values corresponding to reverse and normal slip, respectively). The models shown in the figures, are obtained by performing least squares inversion for a wide range of the regularization parameters and then selecting the best one using Generalized Cross Validation model selection technique [Craven and Wahba, 1979].
Figure S.6. Slip inversion test using $C_m$ regularization method. See Text S6 for further explanation.
Figure S.7. Slip inversion test using order 0 Tikhonov regularization. See Text S6 for further explanation.
Figure S.8. Slip inversion test using 1st order Tikhonov regularization. See Text S6 for further explanation.
Figure S.9. Slip inversion test using 2\textsuperscript{nd} order Tikhonov regularization. See Text S6 for further explanation.
Figure S.10. Slip inversion test using Sensitivity modulated 2\textsuperscript{nd} order Tikhonov regularization. See Text S6 for further explanation.
Figure S.11. Slip inversion test using the EPIC for 1st order Tikhonov regularization. See Text S6 for further explanation.
Figure S.12. Slip inversion test using the EPIC for 2nd order Tikhonov regularization. See Text S6 for further explanation.
Text S7 Synthetic inversions: Elliptic $M_w$ 8.0 earthquake.

We first generate synthetic data, for all GPS stations of the GEONET network [Sagiya et al., 2000] in north east Japan, using a slip distribution resembling an $M_w$8.0 earthquake with elliptical shape, considering several cases of centroid location at the parameterized fault surface. Subsequently, we generate three unbiased noise cases, named Noise1, Noise2 and Noise3, using realizations of a multivariate normal distribution with zero mean and covariance matrix representing uncertainties on crustal displacement synthetic data (0.5 cm and 1 cm for horizontal and vertical components respectively) at the location of the GPS instruments used in this work (same as in Simons et al. [2011]).

Figures S.13 to S.19 show estimates of rake parallel slip (see Text S4) for all cases of synthetic data perturbed with all noise cases, each figure showing results for a different regularization operator. The first row of models show the slip distribution used to compute synthetic data (labeled as True model), and subsequent rows show slip estimates for synthetic data in which Noise1, Noise2 and Noise3 cases of random noise have been added.

Colorbars show slip normalized by its maximum value, in which the colorbar saturates at 10% of back-slip (-10% of peak value, pink tones). Each slip model is shown in a box with two insets. The top-left inset shows the moment magnitude of such slip distribution and the bottom-right inset shows minimum and maximum value of slip (being positive and negative values corresponding to reverse and normal slip, respectively). The models shown in the figures, are obtained by performing least squares inversion for a wide range of the regularization parameters and then selecting the best one using Generalized Cross Validation model selection technique [Craven and Wahba, 1979].
Figure S.13. Slip inversion test using $C_{\text{m}}^4$ regularization method. See Text S7 for further explanation.
Figure S.14. Slip inversion test using order 0 Tikhonov regularization. See Text S7 for further explanation.
Figure S.15. Slip inversion test using 1st order Tikhonov regularization. See Text S7 for further explanation.
Figure S.16. Slip inversion test using 2nd order Tikhonov regularization. See Text S7 for further explanation.
Figure S.17. Slip inversion test using Sensitivity modulated 2\textsuperscript{nd} order Tikhonov regularization. See Text S7 for further explanation.
Figure S.18. Slip inversion test using the EPIC for 1st order Tikhonov regularization. See Text S7 for further explanation.
Figure S.19. Slip inversion test using the EPIC for 2nd order Tikhonov regularization. See Text S7 for further explanation.
Text S8 Synthetic inversions: Elliptic Mw 8.5 earthquake.

We first generate synthetic data, for all GPS stations of the GEONET network [Sagiya et al., 2000] in north east Japan, using a slip distribution resembling an Mw 8.5 earthquake with elliptical shape, considering several cases of centroid location at the parameterized fault surface. Subsequently, we generate three unbiased noise cases, named Noise1, Noise2 and Noise3, using realizations of a multivariate normal distribution with zero mean and covariance matrix representing uncertainties on crustal displacement synthetic data (0.5 cm and 1 cm for horizontal and vertical components respectively) at the location of the GPS instruments used in this work (same as in Simons et al. [2011]).

Figures S.20 to S.26 show estimates of rake parallel slip (see Text S4) for all cases of synthetic data perturbed with all noise cases, each figure showing results for a different regularization operator. The first row of models show the slip distribution used to compute synthetic data (labeled as True model), and subsequent rows show slip estimates for synthetic data in which Noise1, Noise2 and Noise3 cases of random noise have been added.

Colorbars show slip normalized by its maximum value, in which the colorbar saturates at 10% of back-slip ( -10% of peak value, pink tones). Each slip model is shown in a box with two insets. The top-left inset shows the moment magnitude of such slip distribution and the bottom-right inset shows minimum and maximum value of slip (being positive and negative values corresponding to reverse and normal slip, respectively). The models shown in the figures, are obtained by performing least squares inversion for a wide range of the regularization parameters and then selecting the best one using Generalized Cross Validation model selection technique [Craven and Wahba, 1979].
Figure S.20. Slip inversion test using $C_m^4$ regularization method. See Text S8 for further explanation.
Figure S.21. Slip inversion test using order 0 Tikhonov regularization. See Text S8 for further explanation.
**Figure S.22.** Slip inversion test using 1st order Tikhonov regularization. See Text S8 for further explanation.
Figure S.23. Slip inversion test using 2nd order Tikhonov regularization. See Text S8 for further explanation.
Figure S.24. Slip inversion test using Sensitivity modulated 2nd order Tikhonov regularization. See Text S8 for further explanation.
Figure S.25. Slip inversion test using the EPIC for 1st order Tikhonov regularization. See Text S7 for further explanation.
Figure S.26. Slip inversion test using the EPIC for 2\textsuperscript{nd} order Tikhonov regularization. See Text S8 for further explanation.
Text S9  Synthetic inversions: Elliptic $M_w$ 9.0 earthquake.

We first generate synthetic data, for all GPS stations of the GEONET network [Sagiya et al., 2000] in north east Japan, using a slip distribution resembling an $M_w$9.0 earthquake with elliptical shape, considering several cases of centroid location at the parameterized fault surface. Subsequently, we generate three unbiased noise cases, named Noise1, Noise2 and Noise3, using realizations of a multivariate normal distribution with zero mean and covariance matrix representing uncertainties on crustal displacement synthetic data (0.5 cm and 1 cm for horizontal and vertical components respectively) for synthetic data at the location of the GPS instruments used in this work (same as in Simons et al. [2011]).

Figures S.27 to S.33 show estimates of rake parallel slip (see Text S4) for all cases of synthetic data perturbed with all noise cases, each figure showing results for a different regularization operator. The first row of models show the slip distribution used to compute synthetic data (labeled as True model), and subsequent rows show slip estimates for synthetic data in which Noise1, Noise2 and Noise3 cases of random noise have been added.

Colorbars show slip normalized by its maximum value, in which the colorbar saturates at 10% of back-slip (-10% of peak value, pink tones). Each slip model is shown in a box with two insets. The top-left inset shows the moment magnitude of such slip distribution and the bottom-right inset shows minimum and maximum value of slip (being positive and negative values corresponding to reverse and normal slip, respectively). The models shown in the figures, are obtained by performing least squares inversion for a wide range of the regularization parameters and then selecting the best one using Generalized Cross Validation model selection technique [Craven and Wahba, 1979].
Figure S.27. Slip inversion test using $C_{\lambda m}$ regularization method. See Text S9 for further explanation.
Figure S.28. Slip inversion test using order 0 Tikhonov regularization. See Text S9 for further explanation.
Figure S.29. Slip inversion test using 1st order Tikhonov regularization. See Text S9 for further explanation.
Figure S.30. Slip inversion test using 2nd order Tikhonov regularization. See Text S9 for further explanation.
Figure S.31. Slip inversion test using Sensitivity modulated 2\textsuperscript{nd} order Tikhonov regularization. See Text S9 for further explanation.
Figure S.32. Slip inversion test using the EPIC for 1st order Tikhonov regularization. See Text S9 for further explanation.
Figure S.33. Slip inversion test using the EPIC for 2nd order Tikhonov regularization. See Text S9 for further explanation.
Comparing Inversion results with synthetic data: Best model selected using Generalized Cross Validation (GCV)

Here we compare directly the solutions of the inversion tests with synthetic datasets as a function of the regularization scheme used in their determination. For each synthetic dataset generated with checkerboard and elliptical earthquakes slip patterns (see Text S6 to Text S9), we generate synthetic data with added random noise and perform inversions with several test values of the respective regularization parameter. In figures S.34 to S.37, we only consider the Noise 2 case, and we show the best inverted models, corresponding to the ones in which the selected regularization parameter minimizes the Generalized Cross Validation [Craven and Wahba, 1979] cost function, computed during each least squares inversion.

Figures S.34 to S.37 show estimates of rake parallel slip (see Text S4), so within each Figure we compare the behavior of the different regularization schemes for different cases of a given synthetic slip pattern being inverted. The first row of models show the slip distribution used to compute synthetic data (labeled as True model), and subsequent rows show slip estimates using different regularization schemes (each labeled on the left side of the row), constrained by the synthetic data in which Noise 2 has been added.

Colorbar shows slip in meters when checkerboard slip patterns are shown, and slip normalized by its maximum value for the synthetic earthquake slip patterns. For checkerboard slip patterns, slip colorbar saturates at 10% above maximum slip (1.1 m, gray tones) and at 10% of back-slip (-0.1 m, pink tones). For elliptical earthquake slip patterns, the colorbar saturates at 10% of back-slip (-10% of peak value, pink tones). Each slip model is shown in a box with two insets. The top-left inset shows the moment magnitude of such slip distribution and the bottom-right inset shows minimum and maximum value of slip (being positive and negative values corresponding to reverse and normal slip, respectively).
Figure S.34. Comparison of regularization schemes effect on slip estimates when using synthetic data generated with checkerboard slip models at different scales. See Text S10 for further figure explanation.
Figure S.35. Comparison of regularization schemes effect on slip estimates when using synthetic data generated with synthetic Mw 8.0 earthquake slip models at different locations. See Text S10 for further figure explanation.
Figure S.36. Comparison of regularization schemes effect on slip estimates when using synthetic data generated with synthetic Mw 8.5 earthquake slip models at different locations. See Text S10 for further figure explanation.
Figure S.37. Comparison of regularization schemes effect on slip estimates when using synthetic data generated with synthetic Mw 9.0 earthquake slip models at different locations. See Text S10 for further figure explanation.
Text S11 Synthetic inversions: Generalized Cross Validation cost function and families of models

Here, for the inversion cases shown in Figures 2 and 3 of the main body of the article in which Noise 2 is added to the synthetic observations, we show a subset of the family of slip models, that are solutions of the inverse problem, defined by 7 selected values of the regularization parameter for each respective inversion. We compare 3 under-regularized models, the optimum or best model minimizing the Generalized Cross Validation (GCV) cost function [Craven and Wahba, 1979], and 3 over-regularized models, with the aim of observing the effect of the chosen value of the regularization parameter in our model estimates, as well as to analyze the stability of the inverse problem solutions around the optimum value of the regularization parameter determined using GCV model class selection method.

Figures S.38 to S.44 show estimates of rake parallel slip (see Text S4) when a checkerboard slip True model is used to generate synthetic data perturbed with Noise 2, each figure showing results for a different regularization operator. Colorbars show slip in meters, up to 10% above maximum slip of the True model (1.1 m, gray tones) and up to 10% of back-slip (-0.1 m, pink tones) as True model has no back-slip.

Figures S.45 to S.51 show estimates of rake parallel slip (see Text S4) when an elliptical slip True model is used to generate synthetic data perturbed with Noise 2, each figure showing results for a different regularization operator. Colorbars show slip normalized by its maximum amplitude, and saturates at 10% of back-slip (-0.1 * maximum slip amplitude, pink tones) as True model has no back-slip.
Figure S.38. Slip inversion test using $C^4_m$ regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.39. Slip inversion test using order 0 Tikhonov (T0) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.40. Slip inversion test using 1st order Tikhonov (T1) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.41. Slip inversion test using 2nd order Tikhonov (T2) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.42. Slip inversion test using Sensitivity modulated 2nd order Tikhonov (ST2) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.43. Slip inversion test using the EPIC for 1st order Tikhonov (ET1) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.44. Slip inversion test using the EPIC for 2nd order Tikhonov (ET2) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
**Figure S.45.** Slip inversion test using $C_m^\lambda$ regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.46. Slip inversion test using order 0 Tikhonov (T0) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.47. Slip inversion test using 1\textsuperscript{st} order Tikhonov (T1) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.48. Slip inversion test using 2nd order Tikhonov (T2) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
**Figure S.49.** Slip inversion test using Sensitivity modulated 2nd order Tikhonov (ST2) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.50. Slip inversion test using the EPIC for 1st order Tikhonov (ET1) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude $M_w$. See Text S11 for further explanation.
Figure S.51. Slip inversion test using the EPIC for 2nd order Tikhonov (ET2) regularization scheme. Upper panels show Generalized Cross Validation (GCV) cost for all regularization parameters probed for the inversion (left) and color circles indicating a selection of 3 under-regularized (red), the optimum or best model minimizing GCV cost (green) and 3 over-regularized models (blue) at the right panel indicated as dashed circles. Bottom panels show True slip model used to generate synthetic observations in which Noise 2 case is added, and the solutions estimated for the 7 selected regularization parameter values indicated by the respective color number. The inset shows minimum and maximum slip and estimated moment magnitude Mw. See Text S11 for further explanation.
Text S12 Inversion of Tohoku-Oki Mw 9.0 co-seismic slip

We estimate co-seismic slip of the 2011 Tohoku-Oki Mw 9.0 earthquake constrained by onland co-seismic offsets measured at GEONET GPS sites [Sagiya et al., 2000] obtained from Minson et al. [2014]. Using a triangulated fault mesh, we estimate slip in 2 directions, rake parallel and rake perpendicular slip, where the rake direction at each fault patch is defined to be parallel to the average direction of movement obtained from the Global CMT focal mechanism of the event. See Text S4 for further information.

Prior information is defined for the second order differential of rake parallel slip, and for the amplitude of rake perpendicular slip. For the first one, the mean value is set to a null value and its independent uncertainties are defined as metaparameters to be determined during the inversion. For the second one, we prior information on rake perpendicular slip with a null mean and uncertainties equal to 1 meter or as a metaparameter to be determined during the inversion process depending on the chosen regularization scheme used (see below).

For comparison purposes, we perform the inversion using T2, ET2 and ST2 regularization schemes. For T2 and ST2 regularization, we search for $\epsilon^2$ as the regularization parameter, or reciprocal of variances of the prior information on rake parallel slip, and we leave fixed the prior variance of rake perpendicular slip at 1 meter as we expect little slip in the latter direction. For ET2 regularization, we search for target posterior variances as the regularization parameter, using the EPIC to find the variances of prior information on rake parallel and rake perpendicular slip as a function of the selected target posterior variance, searching for a range of target posterior standard deviations for rake parallel slip between 0.1 to 25 meters, and leaving fixed the target posterior standard deviation of rake perpendicular slip to 1 meter as we expect little slip in the latter direction.

Figure S.52 shows Generalized Cross Validation (GCV) cost function [Craven and Wahba, 1979] as a function of the regularization parameter index, ordered in decreasing regularization strength (from larger to smaller $\epsilon^2$ values, from smaller to larger target posterior standard deviations). Also, the figure shows estimated moment magnitudes for all solutions. From all the tested models, we selected 9 models in order to compare results of the regularization schemes that produce the same data misfit level. Such models are indicated with numbers 1,2,3,4 for over-regularized models, 6,7,8,9 for under-regularized models, and 5 for the models that minimize their respective GCV cost function. Slip estimates for all selected models are shown in Figures S.53 to S.61, along with GPS observations and model predicted data.
Figure S.52. GCV cost function (left) and moment magnitude (right) as a function of the regularization parameter index, for the inversions of the 2011 Tohoku-Oki ($M_w$ 9.0) earthquake using standard $2^{nd}$ order Tikhonov (T2), EPIC $2^{nd}$ order Tikhonov (ET2) and Sensitivity $2^{nd}$ order Tikhonov (ST2) regularization schemes. The order of the regularization parameter indices is from over-regularized inversions to under-regularized inversions. See Text S12 for further figure explanation.
Figure S.53: Case 1 choice of regularization parameter (see Figure S.52): Tohoku-Oki (M_w 9.0) 2011 earthquake slip inversion results. Left 3 columns indicate estimated slip amplitude and their rake parallel/perpendicular components, when using T2, EPIC T2 (ET2) and ST2 regularization schemes, respectively. Slip is normalized to its maximum absolute value. Inset shows minimum and maximum value of slip and estimated moment magnitude (M_w) or target a posteriori standard deviation (σ_t) when stated. Last 2 columns show horizontal and vertical GPS data and model prediction for each corresponding regularization case. See Supplementary Text S12 for further details.
Figure S.54: Case 2 choice of regularization parameter (see Figure S.52): Tohoku-Oki ($M_w$ 9.0) 2011 earthquake slip inversion results. Left 3 columns indicate estimated slip amplitude and their rake parallel/perpendicular components, when using T2, EPIC T2 (ET2) and ST2 regularization schemes, respectively. Slip is normalized to its maximum absolute value. Inset shows minimum and maximum value of slip and estimated moment magnitude ($M_w$) or target $a$ posteriori standard deviation ($\sigma_T$) when stated. Last 2 columns show horizontal and vertical GPS data and model prediction for each corresponding regularization case. See Supplementary Text S12 for further details.
Figure S.55: Case 3 choice of regularization parameter (see Figure S.52): Tohoku-Oki (Mw 9.0) 2011 earthquake slip inversion results. Left 3 columns indicate estimated slip amplitude and their rake parallel/perpendicular components, when using T2, EPIC T2 (ET2) and ST2 regularization schemes, respectively. Slip is normalized to its maximum absolute value. Inset shows minimum and maximum value of slip and estimated moment magnitude (Mw) or target a posteriori standard deviation (σT) when stated. Last 2 columns show horizontal and vertical GPS data and model prediction for each corresponding regularization case. See Supplementary Text S12 for further details.
Figure S.56: Case 4 choice of regularization parameter (see Figure S.52): Tohoku-Oki ($M_w$ 9.0) 2011 earthquake slip inversion results. Left 3 columns indicate estimated slip amplitude and their rake parallel/perpendicular components, when using T2, EPIC T2 (ET2) and ST2 regularization schemes, respectively. Slip is normalized to its maximum absolute value. Inset shows minimum and maximum value of slip and estimated moment magnitude ($M_w$) or target $a$ posteriori standard deviation ($\sigma_T$) when stated. Last 2 columns show horizontal and vertical GPS data and model prediction for each corresponding regularization case. See Supplementary Text S12 for further details.
Figure S.57: Case 5 choice of regularization parameter (see Figure S.52): Tohoku-Oki (Mw 9.0) 2011 earthquake slip inversion results. Left 3 columns indicate estimated slip amplitude and their rake parallel/perpendicular components, when using T2, EPIC T2 (ET2) and ST2 regularization schemes, respectively. Slip is normalized to its maximum absolute value. Inset shows minimum and maximum value of slip and estimated moment magnitude (Mw) or target a posteriori standard deviation (σt) when stated. Last 2 columns show horizontal and vertical GPS data and model prediction for each corresponding regularization case. See Supplementary Text S12 for further details.
Figure S.58: Case 6 choice of regularization parameter (see Figure S.52): Tohoku-Oki ($M_w 9.0$) 2011 earthquake slip inversion results. Left 3 columns indicate estimated slip amplitude and their rake parallel/perpendicular components, when using T2, EPIC T2 (ET2) and ST2 regularization schemes, respectively. Slip is normalized to its maximum absolute value. Inset shows minimum and maximum value of slip and estimated moment magnitude ($M_w$) or target $a$ posteriori standard deviation ($\sigma_t$) when stated. Last 2 columns show horizontal and vertical GPS data and model prediction for each corresponding regularization case. See Supplementary Text S12 for further details.
Figure S.59: Case 7 choice of regularization parameter (see Figure S.52): Tohoku-Oki ($\text{M}_\text{w} 9.0$) 2011 earthquake slip inversion results. Left 3 columns indicate estimated slip amplitude and their rake parallel/perpendicular components, when using T2, EPIC T2 (ET2) and ST2 regularization schemes, respectively. Slip is normalized to its maximum absolute value. Inset shows minimum and maximum value of slip and estimated moment magnitude ($\text{M}_\text{w}$) or target $a$ posteriori standard deviation ($\sigma_t$) when stated. Last 2 columns show horizontal and vertical GPS data and model prediction for each corresponding regularization case. See Supplementary Text S12 for further details.
Figure S.60: Case 8 choice of regularization parameter (see Figure S.52): Tohoku-Oki ($M_w$ 9.0) 2011 earthquake slip inversion results. Left 3 columns indicate estimated slip amplitude and their rake parallel/perpendicular components, when using T2, EPIC T2 (ET2) and ST2 regularization schemes, respectively. Slip is normalized to its maximum absolute value. Inset shows minimum and maximum value of slip and estimated moment magnitude ($M_w$) or target $a$ posteriori standard deviation ($\sigma_T$) when stated. Last 2 columns show horizontal and vertical GPS data and model prediction for each corresponding regularization case. See Supplementary Text S12 for further details.
Figure S.61: Case 9 choice of regularization parameter (see Figure S.52): Tohoku-Oki ($M_w$ 9.0) 2011 earthquake slip inversion results. Left 3 columns indicate estimated slip amplitude and their rake parallel/perpendicular components, when using T2, EPIC_T2 (ET2) and ST2 regularization schemes, respectively. Slip is normalized to its maximum absolute value. Inset shows minimum and maximum value of slip and estimated moment magnitude ($M_w$) or target a posteriori standard deviation ($\sigma_T$) when stated. Last 2 columns show horizontal and vertical GPS data and model prediction for each corresponding regularization case. See Supplementary Text S12 for further details.
Text S13 Testing EPIC T2 regularization for Inversion of Tohoku-Oki Mw 9.0 Co-seismic Slip with Positivity Constraints on Rake Parallel Slip

We estimate co-seismic slip of the 2011 Tohoku-Oki Mw 9.0 earthquake constrained by onland co-seismic offsets measured at Japanese GEONET GPS network sites obtained from Minson et al. [2014]. We use the same model setup and type of regularization as for the inversion shown in section 3.2 of the main text (see also Supplementary Text S6) but with the addition of positivity constraints. Here, we enforce rake parallel slip to be positive, allowing slip estimates with only a reverse component of slip, consistent with the relative movement of the megathrust fault during the Tohoku-Oki earthquake.

For comparison purposes, we perform the inversion using EPIC T2 (ET2), T2 and ST2 regularization schemes for rake parallel slip, while keeping the zeroth order Tikhonov regularization for rake perpendicular slip. Figures S.62, S.66 and S.70 relate data misfit, model roughness, maximum slip amplitude and slip model L2-norm for all considered values of regularization parameters when using EPIC T2 (ET2), T2 and ST2 regularization schemes, respectively. In the aforementioned figures, 6 numbered blue dots indicate models obtained for a selection of regularization parameters. Figures S.63, S.67 and S.71 show the slip distributions when using EPIC T2 (ET2), T2 and ST2 regularization, respectively. Each figure shows the 6 selected slip models labeled from 1 to 6. When comparing slip models using EPIC T2 (ET2), T2 and ST2 regularization, models with the same numerical label have similar data misfits. In the same manner, Figures S.64-S.65, S.68-S.69 and S.72-S.73 show mean model and standard deviation calculated using Montecarlo error propagation techniques, when using EPIC T2 (ET2), T2 and ST2 regularization, respectively. For the Montecarlo error propagation, by a process of trial and error, we consider that 20000 sampled a posteriori models represent well the statistics of the estimated slip.

For ET2 regularization we expect that the EPIC holds at least for the model parameters that present larger slip values (different from zero). Figures S.64 and S.65 show mean slip and standard deviations calculated using Monte Carlo error propagation techniques for the 6 selected slip models shown in Figure S.63. The numeric label of the models (1 to 6) are ordered in decreasing amount of regularization, i.e., increasing value of the regularization parameter σt. Models 1 to 5 present posterior slip standard deviations that vary little within the fault region that presents larger slip values. For model 6, which has a larger σt value, we observe much larger variations of posterior slip standard deviations. Thus, for large enough σt values, the EPIC does not hold when including positivity constraints. A similar behavior is observed when comparing estimated slip models (Figure S.63) and the mean of the models obtained using montecarlo error propagation techniques (Figure S.64). Here, the spatial distribution of the slip models 1 to 5 are similar, although smaller slip amplitude for the mean of the models, but the spatial distribution differs substantially for model 6. Such observation suggests that such spatial similarity might serve as a proxy to quantify the range of σt values such that the EPIC is useful when including positivity constraints. Nevertheless, model 6 has a value of σt that is much larger than the one that would be selected using the L-curve criterion (models 3 or 4), indicating that the amount of regularization chosen in model 6 is not enough.
Figure S.62. Relation between data misfit and model roughness ($\nabla^2 m$), estimated maximum slip amplitude, estimated model $L_2$ norm (panels (a), (b) and (c) respectively. Panel (d) shows a quasi-linear relationship between the logarithm of the regularization parameter and the maximum estimated slip. See Text S13 for further detail.
Figure S.63. Tohoku-Oki ($M_w 9.0$) 2011 earthquake estimated slip distributions for 6 selected values of the regularization parameter. Inset shows minimum and maximum value of slip amplitude, and estimated moment magnitude of the mainshock event. White lines indicate 10 m contours from Minson et al. [2014]. See Text S13 for further detail.
Figure S.64. Mean model from Montecarlo error propagation for 6 selected Tohoku-Oki ($M_w$ 9.0) 2011 earthquake estimated slip distributions. Inset shows minimum and maximum value of slip amplitude, and estimated moment magnitude of the mainshock event. White lines indicate 10 m contours from Minson et al. [2014]. See Text S13 for further detail.
Figure S.65. Standard deviation of models from Montecarlo error propagation for 6 selected Tohoku-Oki (Mw 9.0) 2011 earthquake estimated slip distributions. Inset shows minimum and maximum value of slip amplitude, and estimated moment magnitude of the mainshock event. White lines indicate 10 m contours from Minson et al. [2014]. See Text S13 for further detail.
T2 Regularization

Figure S.66. Relation between data misfit and model roughness ($\nabla^2 m$), estimated maximum slip amplitude, estimated model $L_2$ norm (panels (a), (b) and (c) respectively. Panel (d) shows a quasi-linear relationship between the logarithm of the regularization parameter and the maximum estimated slip. See Text S13 for further detail.
Figure S.67.  Tohoku-Oki (Mw 9.0) 2011 earthquake estimated slip distributions for 6 selected values of the regularization parameter. Inset shows minimum and maximum value of slip amplitude, and estimated moment magnitude of the mainshock event. White lines indicate 10 m contours from Minson et al. [2014]. See Text S13 for further detail.
Figure S.68. Mean model from Montecarlo error propagation for 6 selected Tohoku-Oki ($M_w$9.0) 2011 earthquake estimated slip distributions. Inset shows minimum and maximum value of slip amplitude, and estimated moment magnitude of the mainshock event. White lines indicate 10 m contours from Minson et al. [2014]. See Text S13 for further detail.
Figure S.69. Standard deviation of models from Montecarlo error propagation for 6 selected Tohoku-Oki (Mw 9.0) 2011 earthquake estimated slip distributions. Inset shows minimum and maximum value of slip amplitude, and estimated moment magnitude of the mainshock event. White lines indicate 10 m contours from Minson et al. [2014]. See Text S13 for further detail.
ST2 Regularization

Figure S.70. Relation between data misfit and model roughness ($\nabla^2 m$), estimated maximum slip amplitude, estimated model $L_2$ norm (panels (a), (b) and (c) respectively. Panel (d) shows a quasi-linear relationship between the logarithm of the regularization parameter and the maximum estimated slip. See Text S13 for further detail.
Figure S.71. Tohoku-Oki (Mw 9.0) 2011 earthquake estimated slip distributions for 6 selected values of the regularization parameter. Inset shows minimum and maximum value of slip amplitude, and estimated moment magnitude of the mainshock event. White lines indicate 10 m contours from Minson et al. [2014]. See Text S13 for further detail.
Figure S.72. Mean model from Montecarlo error propagation for 6 selected Tohoku-Oki (Mw 9.0) 2011 earthquake estimated slip distributions. Inset shows minimum and maximum value of slip amplitude, and estimated moment magnitude of the mainshock event. White lines indicate 10 m contours from Minson et al. [2014]. See Text S13 for further detail.
Figure S.73. Standard deviation of models from Monte carlo error propagation for 6 selected Tohoku-Oki (Mw 9.0) 2011 earthquake estimated slip distributions. Inset shows minimum and maximum value of slip amplitude, and estimated moment magnitude of the mainshock event. White lines indicate 10 m contours from Minson et al. [2014]. See Text S13 for further detail.
Text S14 EPIC Feasibility and Selection of Range of Possible Values for Target Posterior Standard Deviation $\sigma_t$

When performing a regularized inversion, typically one must define a range of allowable values of the regularization parameter, to then determine the best one using some model class selection criteria. For instance, when using second order Tikhonov regularization (T2) one must choose an initial set of values of the damping constant $\epsilon$. The range of values for $\epsilon$ is hard to guess, as the regularization parameter corresponds to the weight for the regularization equations and the units of $\epsilon$ are the ones of the reciprocal of a second order spatial derivative of slip. When using EPIC Tikhonov (e.g., ET2) we do have to select an initial range of values for the target posterior standard deviation ($\sigma_t$). Nevertheless, as one of the strengths of the proposed EPIC Tikhonov regularization, the selection of the range for our regularization parameter $\sigma_t$ remains physically interpretable, as it has the same units of the parameters being searched.

For the slip inversion problem, our regularization parameter(s) $\sigma_t$ represent the target standard deviation of inferred slip. In order to determine the range of $\sigma_t$ values, one can start by defining its higher limit. On one hand, if inverting an earthquake slip distribution, the largest $\sigma_t$ can be set to be a fraction of the maximum expected slip obtained from empirical scaling relationships between Mw and maximum slip. On the other hand, if one is inferring back-slip models to compute geodetic coupling at a fault, one can choose the highest possible value of $\sigma_t$ as a fraction of the rate of the rigid relative motion between the two sides of the fault (e.g., plate convergence rate in subduction zones).

As for the lower limit of the range of $\sigma_t$, in the context of the aforementioned examples, one can set it as a fraction of the average standard deviation of the surface displacements (or velocities) used to constraint the slip model, as we expect from experience that standard deviations of inferred slip will be larger.

For the synthetic exercises in this paper, we search for values of the target posterior standard deviation $\sigma_t$ between 1 centimeter and 15 meters. Also, for the actual Tohoku-Oki slip inversion we converged on search values of $\sigma_t$ for rake parallel slip between 10 centimeters and 15 meters, while leaving fixed the one for rake perpendicular slip at 1 meter. Then, having determined a range of possible $\sigma_t$ values, one determines the validity of each one by checking that allows a feasible solution for the EPIC. For the examples in this paper, we show in Figure S.74 the square root of the Root Mean Squared of the residual of the EPIC ($\sqrt{\text{RMS}_{\text{EPIC}}}$), where the residual is defined as the difference of both sides of equation 14 in main body of the paper. From Figure S.74, we can see that the EPIC is always well achieved, being consistently much better achieved for EPIC T2 than for EPIC T1.

Regarding the ability to solve the EPIC, for a fine enough fault mesh and in our experience, we have always been able to find a suitable range of $\sigma_t$ values (e.g., Figure S.74 for the examples developed in the paper). In order to further probe the ability to solve the EPIC for a given value of $\sigma_t$, we consider 3 different discretizations of the same fault in which we vary the coarseness of the discretization (see Figure S.75a,b,c). For the synthetic scenarios shown in our work and for ET2 regularization (our recommended regularization scheme), we compute the EPIC for an absurdly wide range of $\sigma_t$ values, between 5 millimeters and 500 meters. Figure S.75d shows the square root of the Root Mean Squared of the residual of the EPIC ($\sqrt{\text{RMS}_{\text{EPIC}}}$) for each considered $\sigma_t$ value and discretized fault mesh.

Figure S.75 shows that the cost function that is minimized when solving the EPIC (summation of squared differences between estimated posterior model variances and target posterior variance $\sigma_t^2$) reaches near zero minimum values for wide enough range of $\sigma_t$ values. Thus, showing the feasibility of the proposed methodology. The flat region on the left occurs because the model is extremely smoothed ($\sigma_t$ too small) and posterior variances are dominated by prior information. The larger values, $\sqrt{\text{RMS}_{\text{EPIC}}} > 10^{-1}$, indicate that for larger $\sigma_t$ values the EPIC becomes unfeasible. The unfeasibility of the EPIC for such large $\sigma_t$ occurs because the discretization size is not small enough for the EPIC to be able to reduce the posterior standard deviation of model parameters in
Figure S.74. Square root of the RMS of residual for the EPIC as a function of the target standard deviation of slip $\sigma_t$. Left panel refers to synthetic experiments in this paper. Right panel refers to the example of the inversion of the Tohoku-oki earthquake (EPIC T2), where the target standard deviation for rake parallel slip $\sigma_{t,\parallel}$ is being searched, and the one for rake perpendicular slip $\sigma_{t,\perp}$ is fixed at 1 meter.

Order to reach the desired $\sigma_t$. We note that the maximum value of feasible $\sigma_t$ depends on the size of the discretized elements of the mesh, being larger for finer fault discretizations (here we assume that everything else in the setting of the inverse problem - e.g., data and their uncertainties, remains unchanged). As a possible explanation, decreasing the size of a fault element, while maintaining the amount of slip, also decreases the data prediction signal to noise ratio, which in turn increases the variance of posterior slip on such fault element for the unregularized problem. The value of such posterior variance pose an upper limit on the values of the target posterior variance that have a feasible solution for the EPIC. As the EPIC defines weights for a smoothing operator, that in turn induces correlations between model parameters, the variances of posterior slip for the regularized problem will tend to be smaller than their value for the unregularized problem. Therefore, if the maximum feasible value of $\sigma_t$ is not large enough, one could avoid such problem by refining the discretization of the fault mesh. For the examples in this paper, all three fault meshes shown in Figure S.75 allow for a wide enough range of $\sigma_t$ values. We used the finer one (Figure S.75c) as we believe that, in an ideal world, we should always over-parametrize and only control the solution through explicit regularization (or prior information), mixing regularization and parameterization leads to confusion and issues with interpretation.
Figure S.75. (a), (b) and (c) show the same fault discretized at three levels of coarseness, with 1009, 1336 and 1936 triangular elements, respectively. The mesh used in the examples of this paper is the one shown in panel (c). For the synthetic scenarios in this work and considering ET2 regularization, panel (d) shows the square root of the RMS of residual for the EPIC as a function of the target standard deviation of slip $\sigma_t$. The color of each curve in (d) corresponds with the color of each mesh in (a), (b), (c).

References


