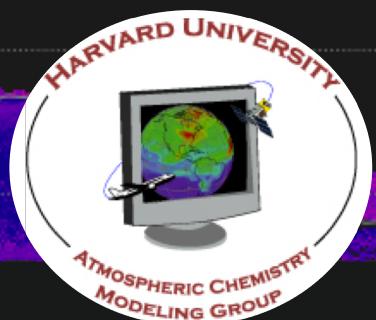


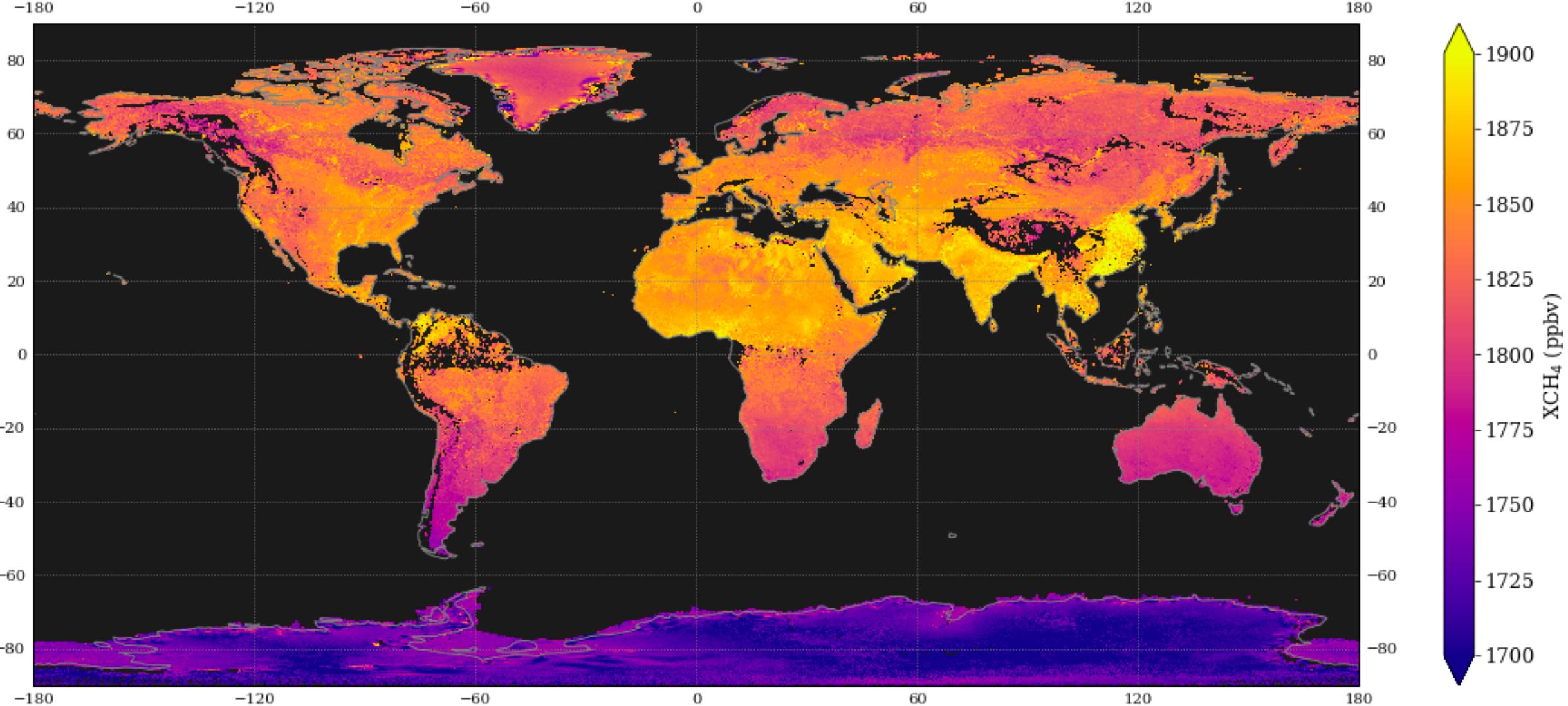
Decreasing the computational cost of analytic inversions of high-resolution satellite observations

Hannah Nesser (hnesser@g.harvard.edu), Daniel Jacob



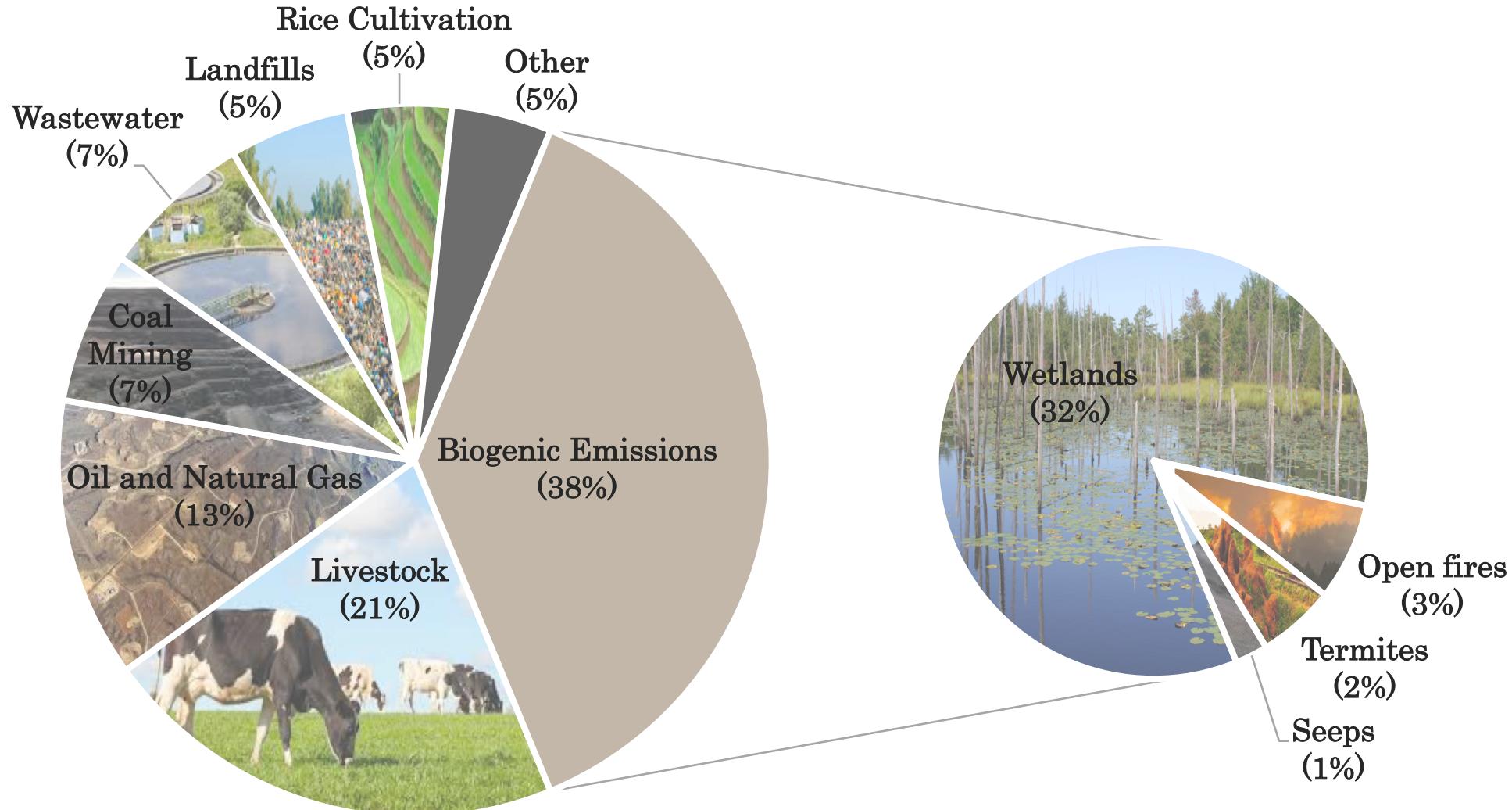
TROPOMI methane observations provide a unique opportunity to improve constraints on emission estimates

TROPOMI CH₄ Observations (May 2018 – March 2019)



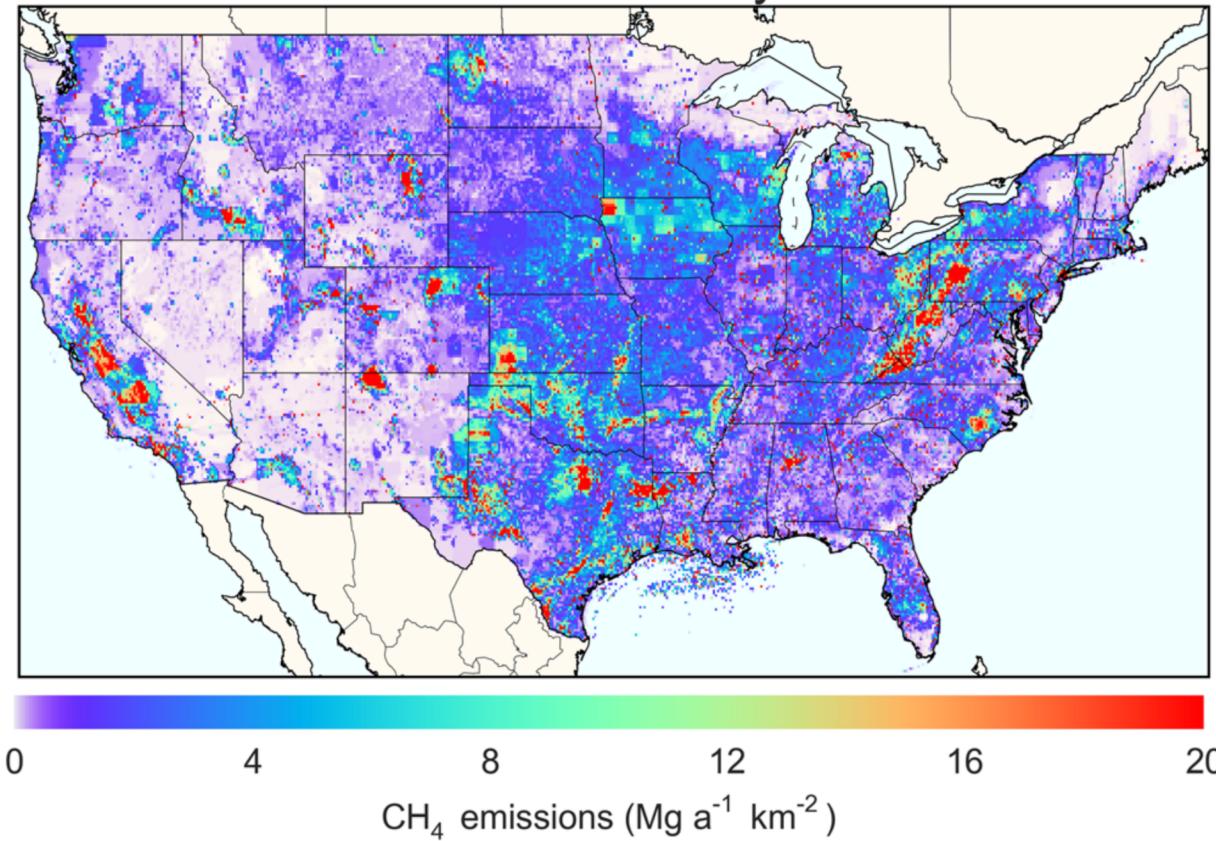
The spatial and temporal distribution of methane emissions is uncertain

Total Methane Emissions: 550 Tg/a

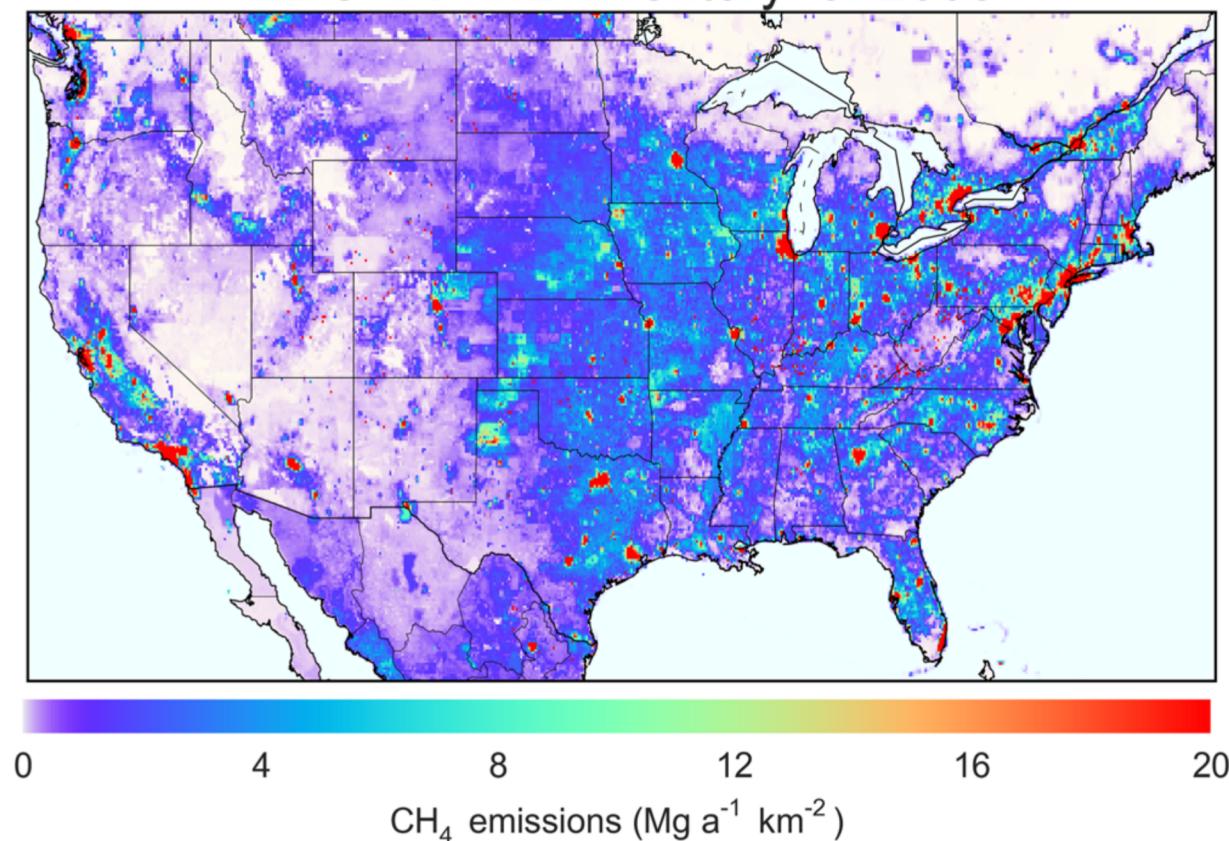


Significant errors exist in prior emission inventories

Gridded EPA Inventory for 2012



EDGAR v4.2 Inventory for 2008

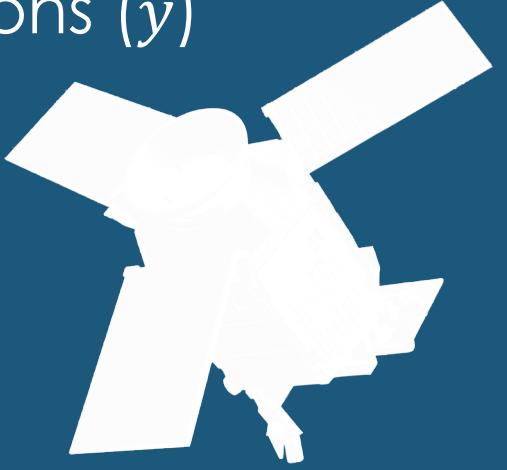


Observations of methane reflect emissions, atmospheric transport, and atmospheric chemistry

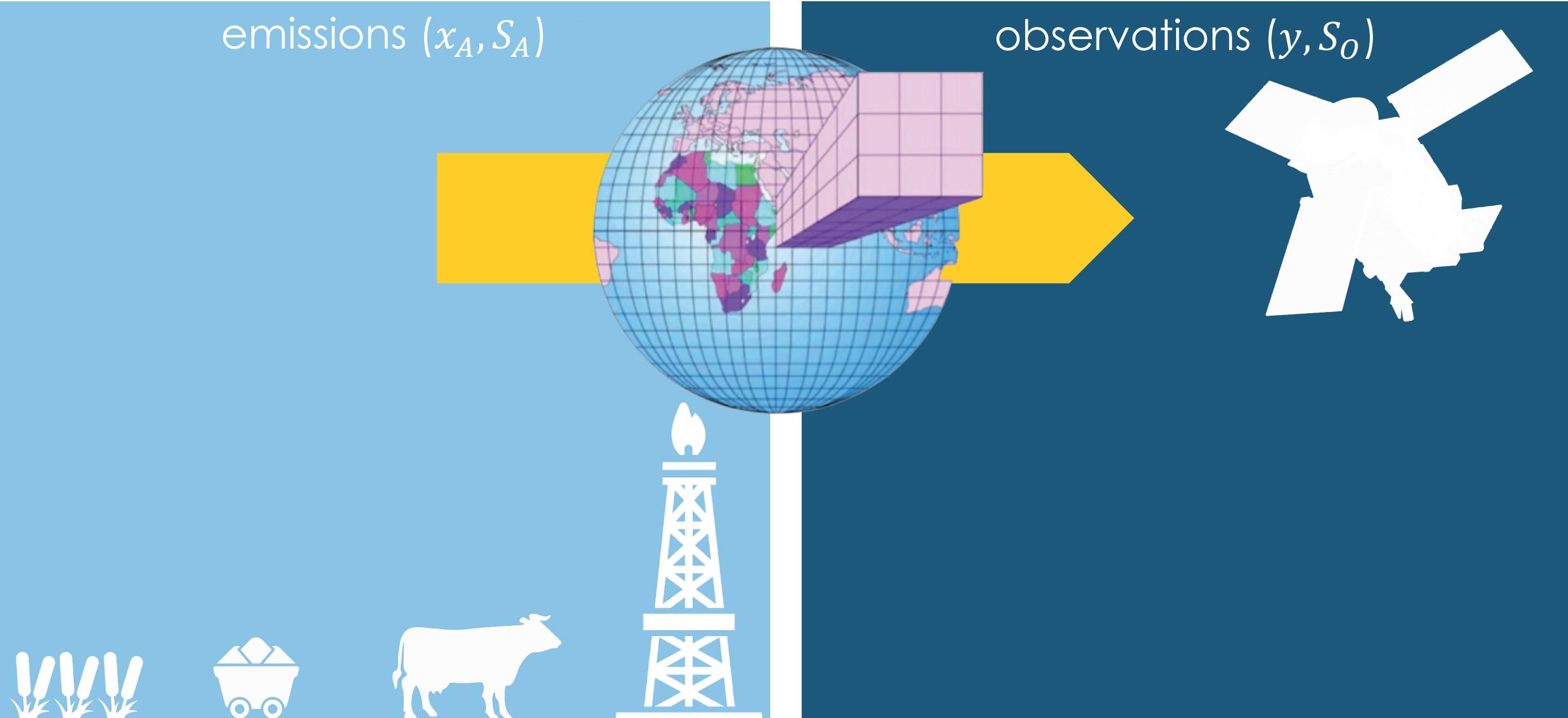
emissions (x)



observations (y)



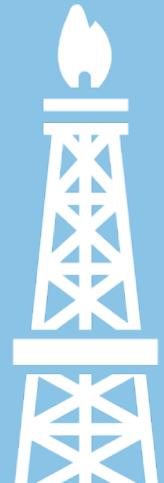
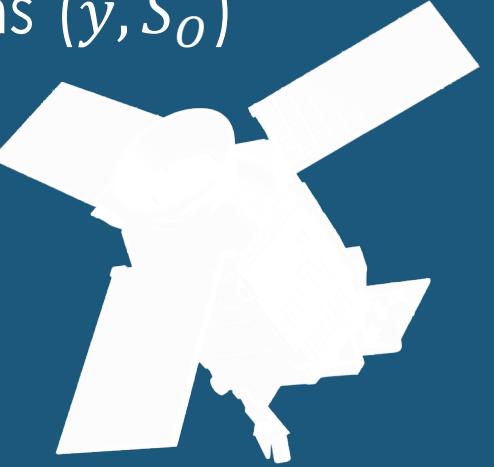
Forward models replicate transport and chemistry to describe the dependence of observations on estimated emissions



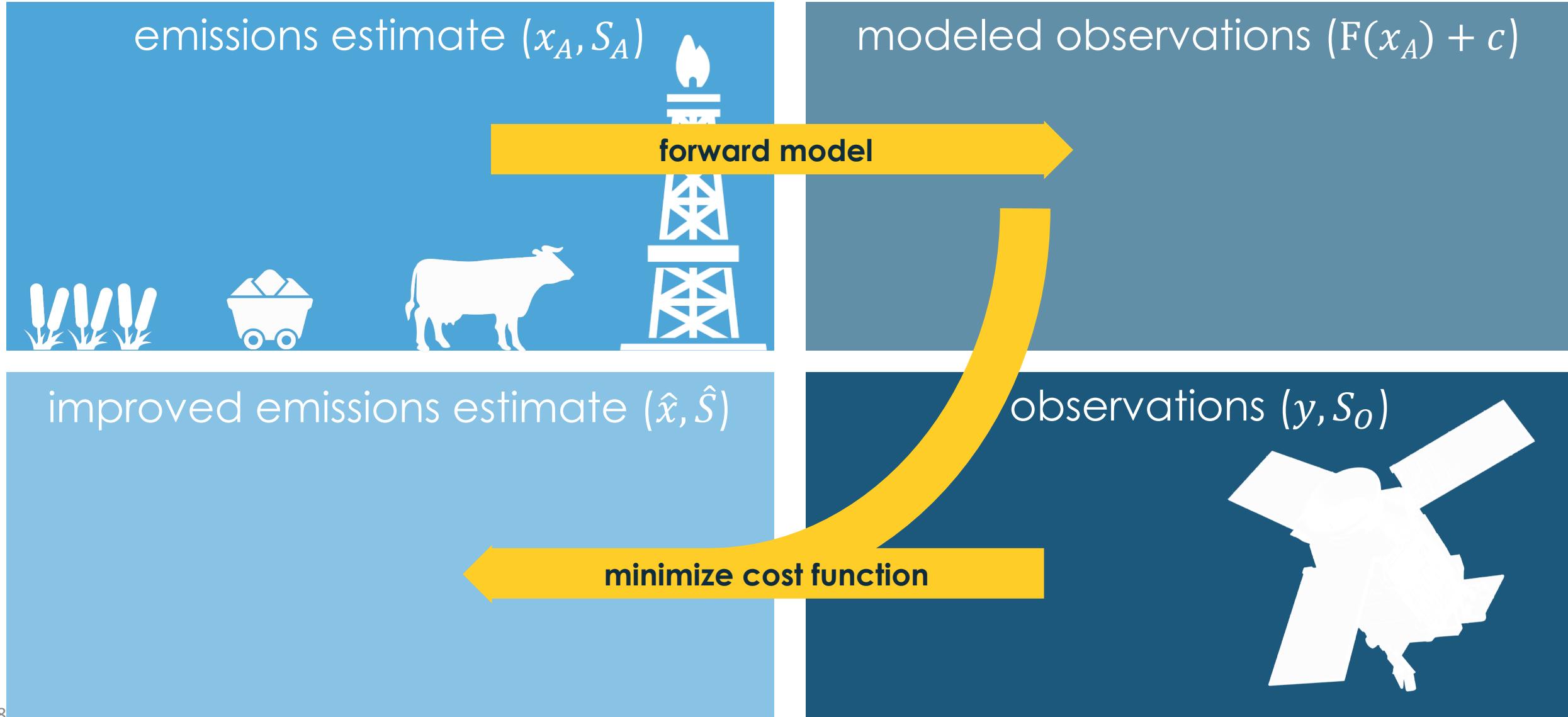
Inversions describe the dependence of emissions on observations

emissions (x_A, S_A)

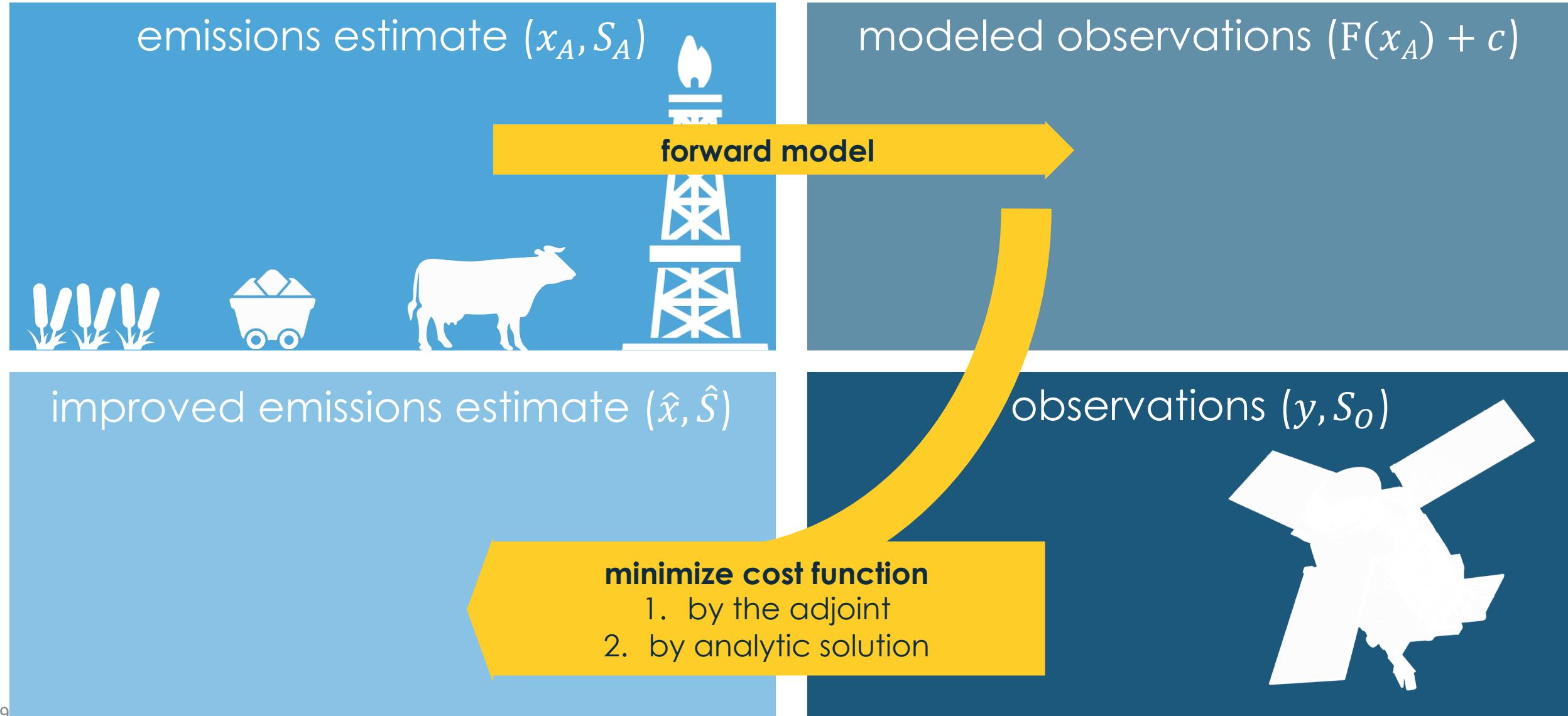
observations (y, S_O)



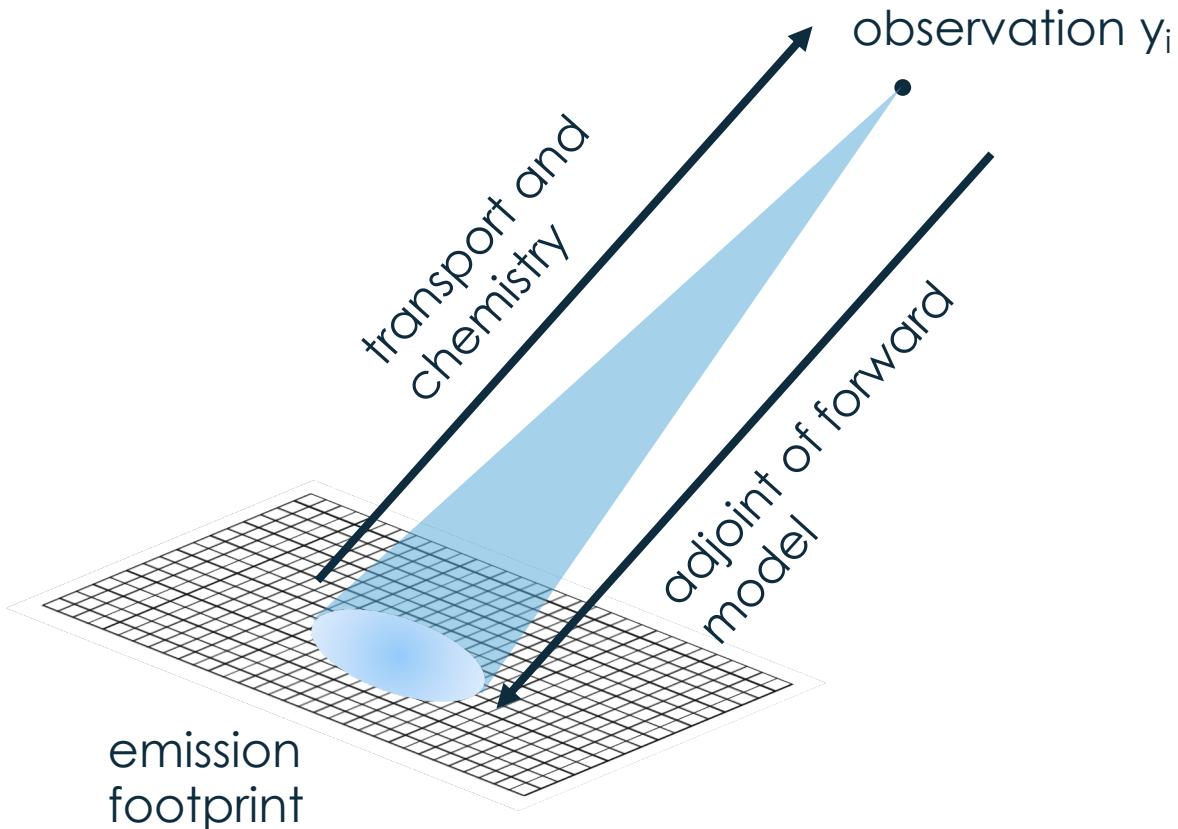
A Bayesian inversion accounts for these errors by maximizing the probability of the emissions given the observations



A Bayesian inversion accounts for these errors by maximizing the probability of the emissions given the observations



While the adjoint efficiently computes the derivative, it does not support analytic solution of errors or extensive sensitivity testing



adjoint solutions do not:

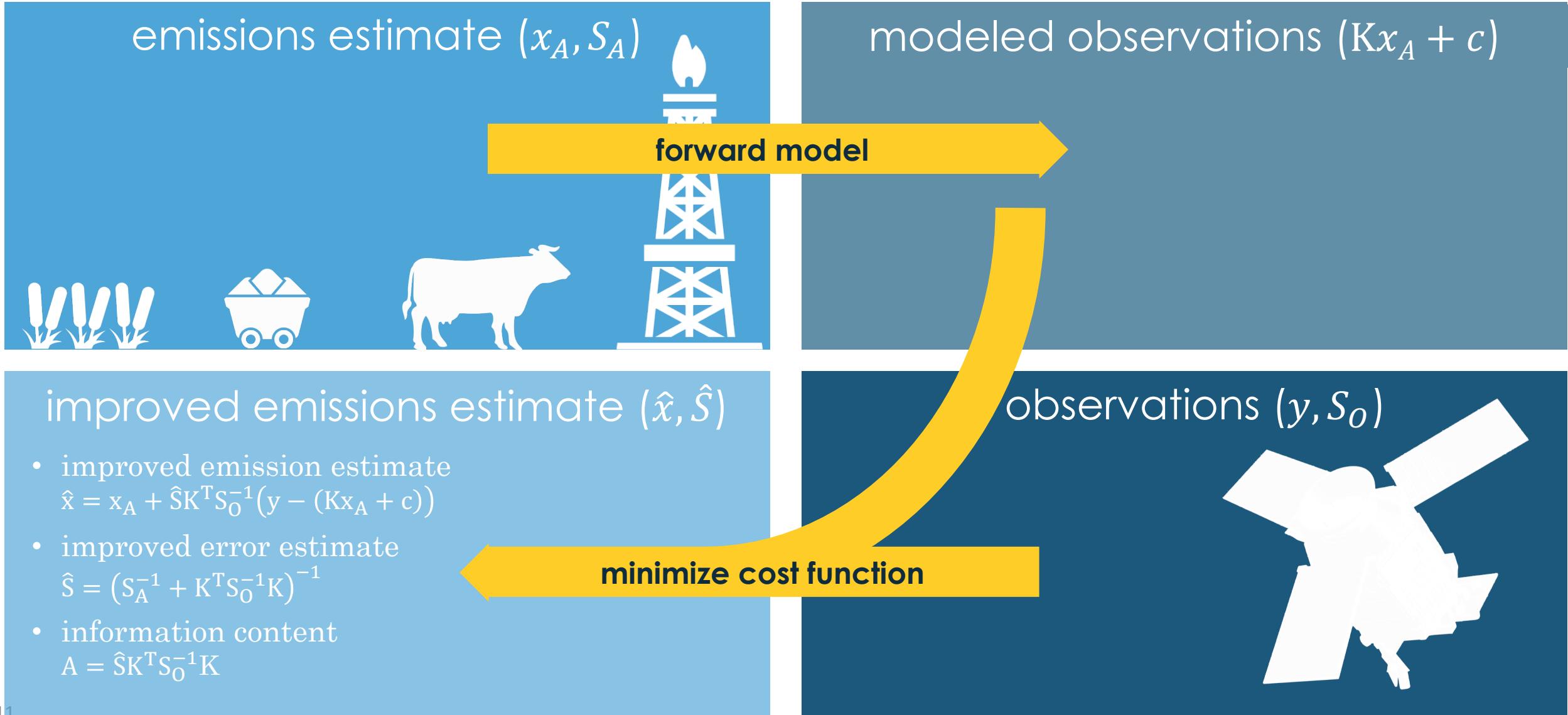
find the true minimum
of the cost function

provide analytic posterior error

support extensive sensitivity testing

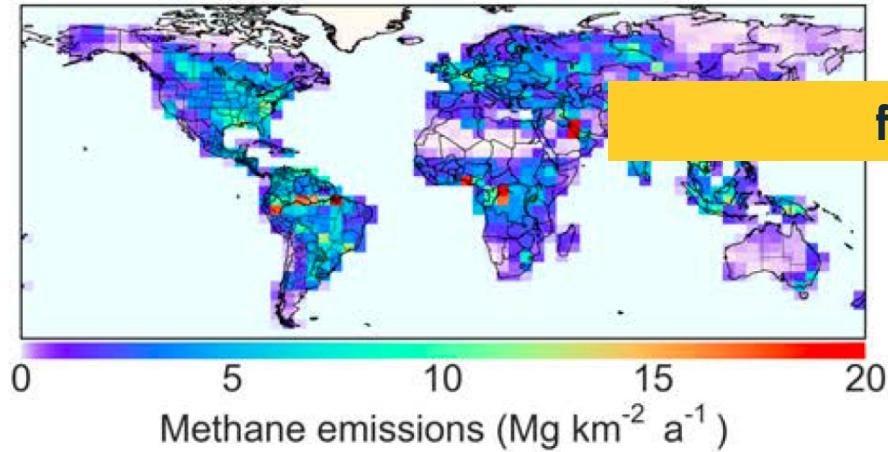
reduce adjoint development costs

An analytic solution to the cost function minimum exists when the forward model is linear



Past inversions optimized emissions on a coarse grid, but denser satellite observations support higher resolution inversions

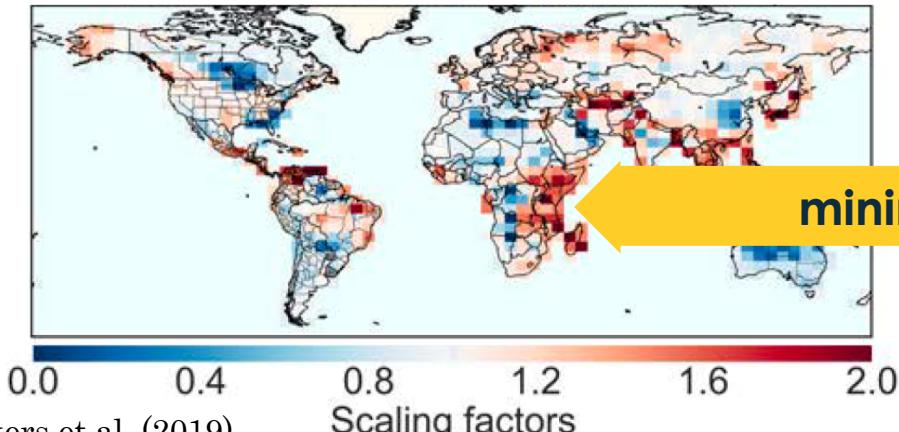
emissions estimate (x_A, S_A)



modeled observations ($Kx_A + c$)

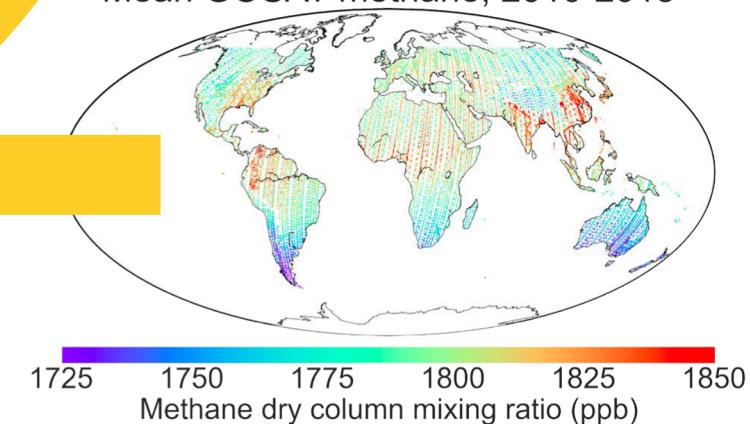
forward model

improved emissions estimate (\hat{x}, \hat{S})



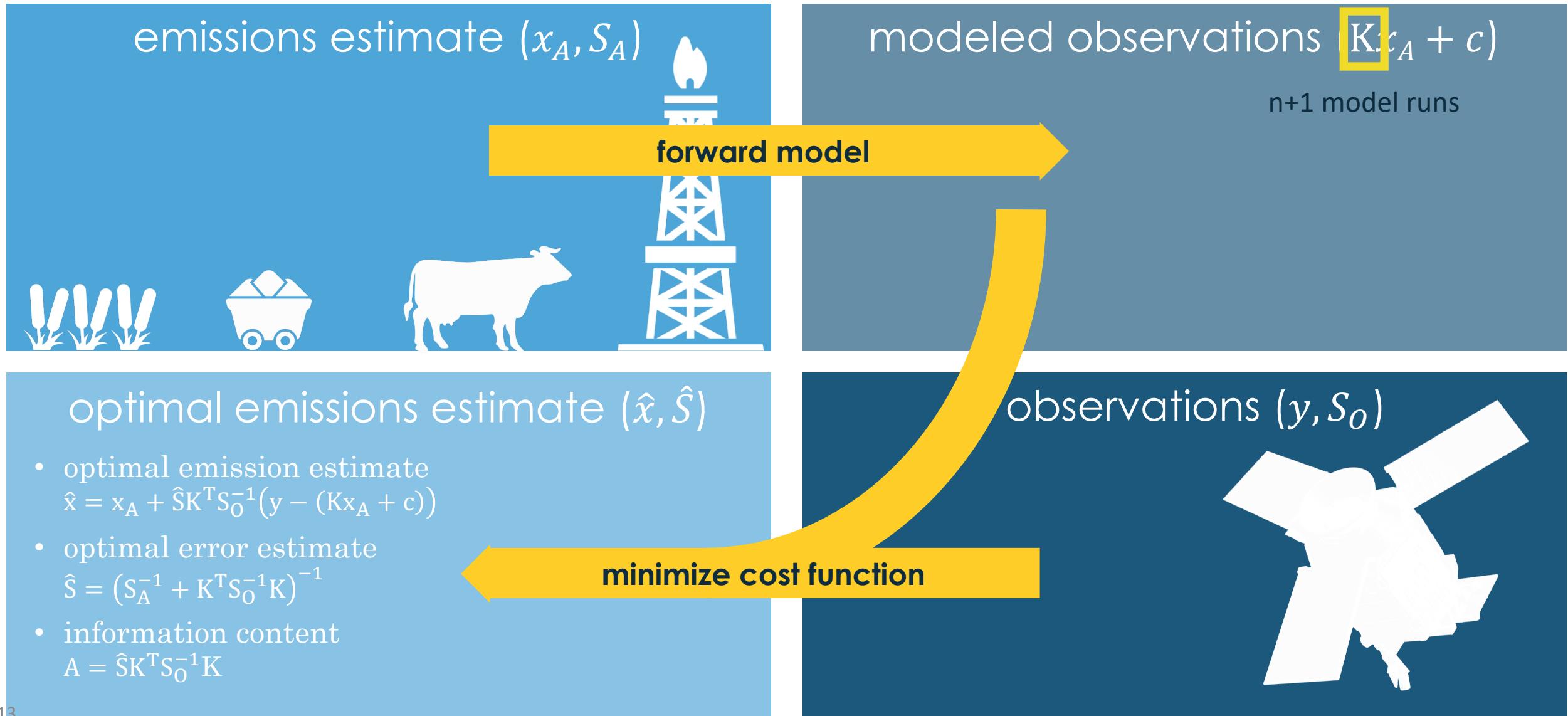
observations (y, S_O)

Mean GOSAT methane, 2010-2015



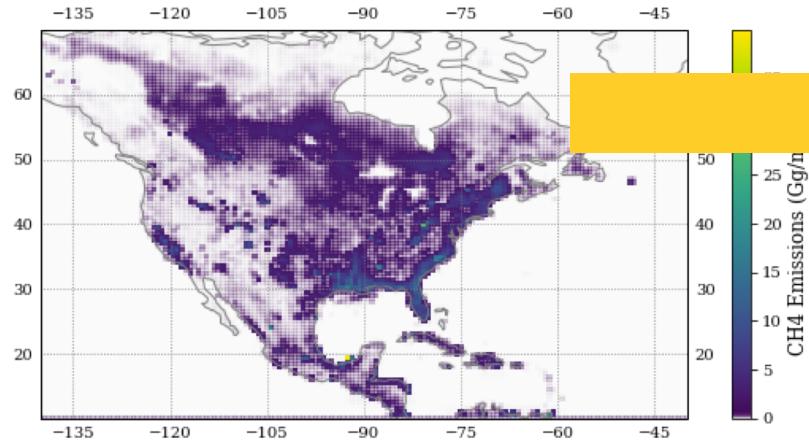
minimize cost function

Increasing inversion resolution increases computational cost, which is limited by the number of grid boxes optimized



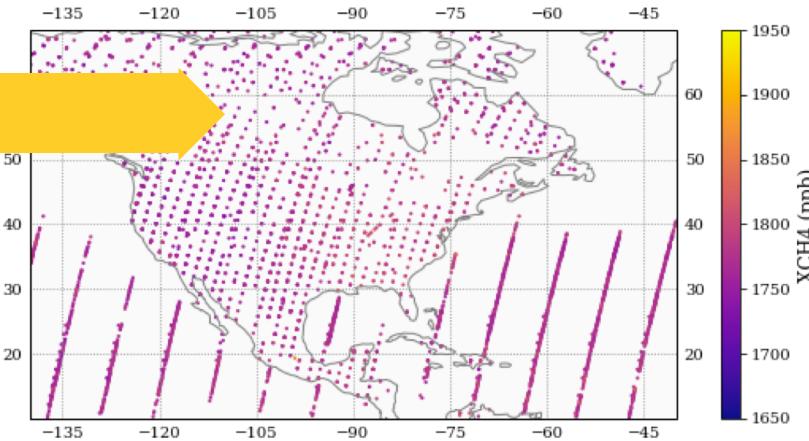
The Jacobian \mathbf{K} represents the sensitivity of observations to emissions: $\mathbf{y} = \mathbf{Kx} + \mathbf{c}$ with $\mathbf{K} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}}$

emissions estimate

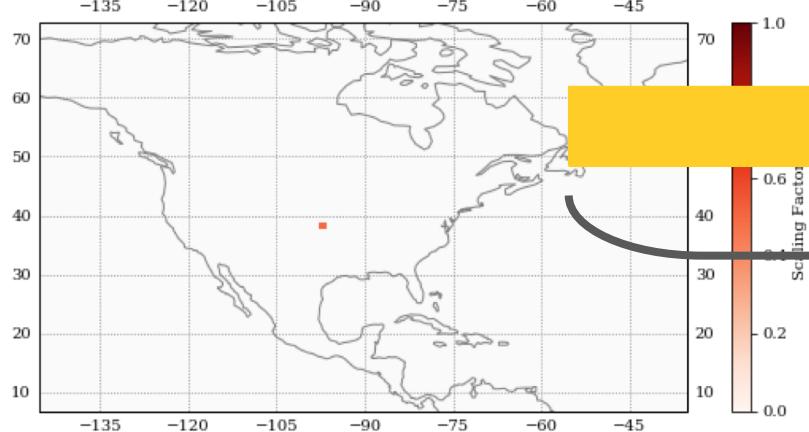


forward model

modeled observations

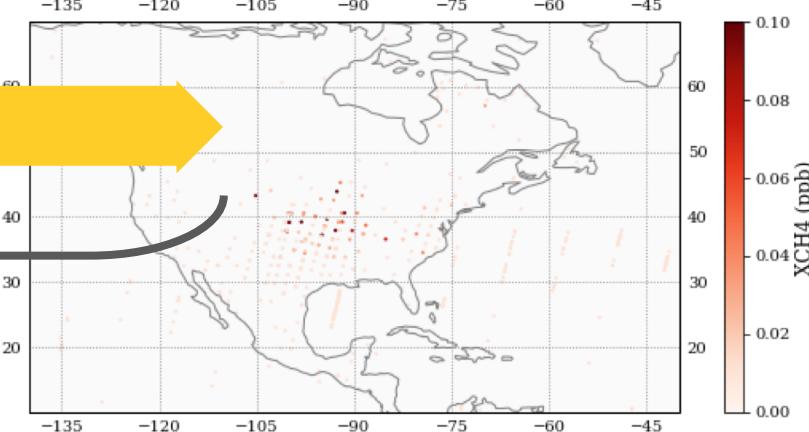


Δx



forward model

Δy

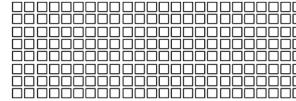


n perturbations,
 n model runs

The computational cost of an analytic Bayesian inversion is limited by the resolution

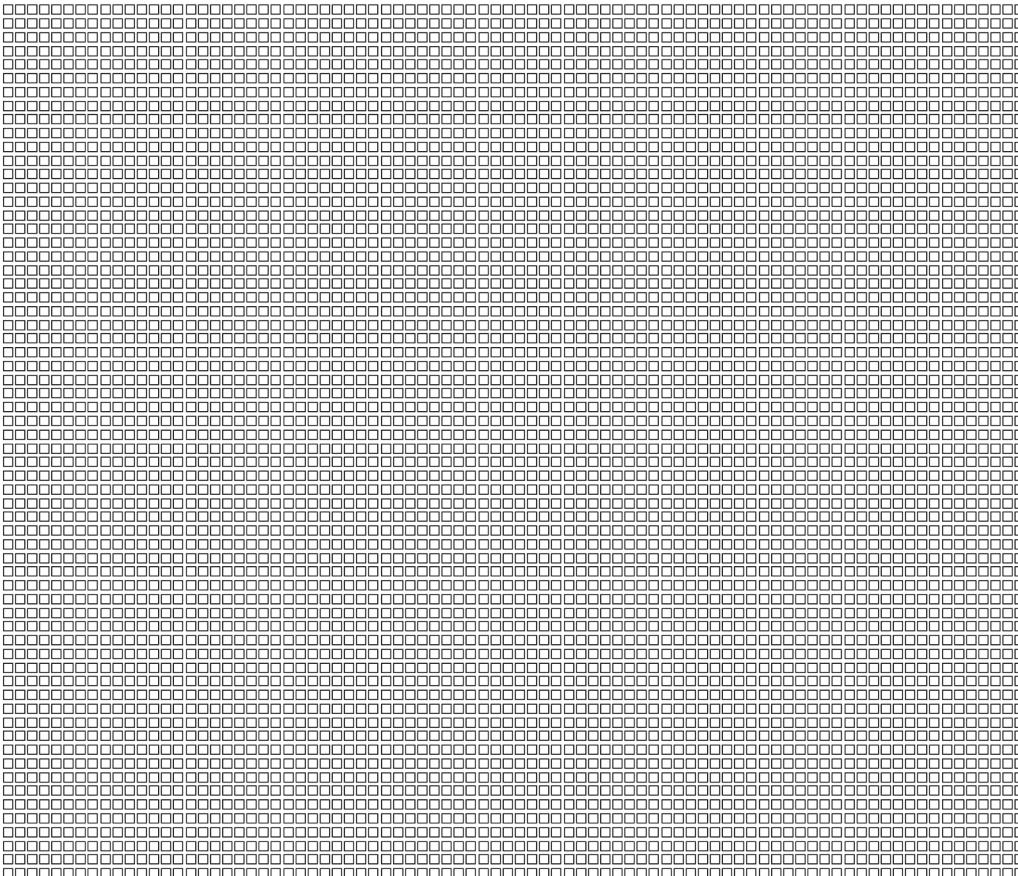
- $4^\circ \times 5^\circ : \sim 16,000 \text{ core-hours}$

1,000 grid cells, 8 cores, 2 hours per simulation-year



- $2^\circ \times 2.5^\circ : \sim 3,000,000 \text{ core-hours}$

4,000 grid cells, 32 cores, 24 hours per simulation-year



- $1^\circ \times 1.25^\circ : \sim 86,000,000 \text{ core-hours}$

16,000 grid cells, 32 cores, 7 days per simulation-year

How can analytic inversions:

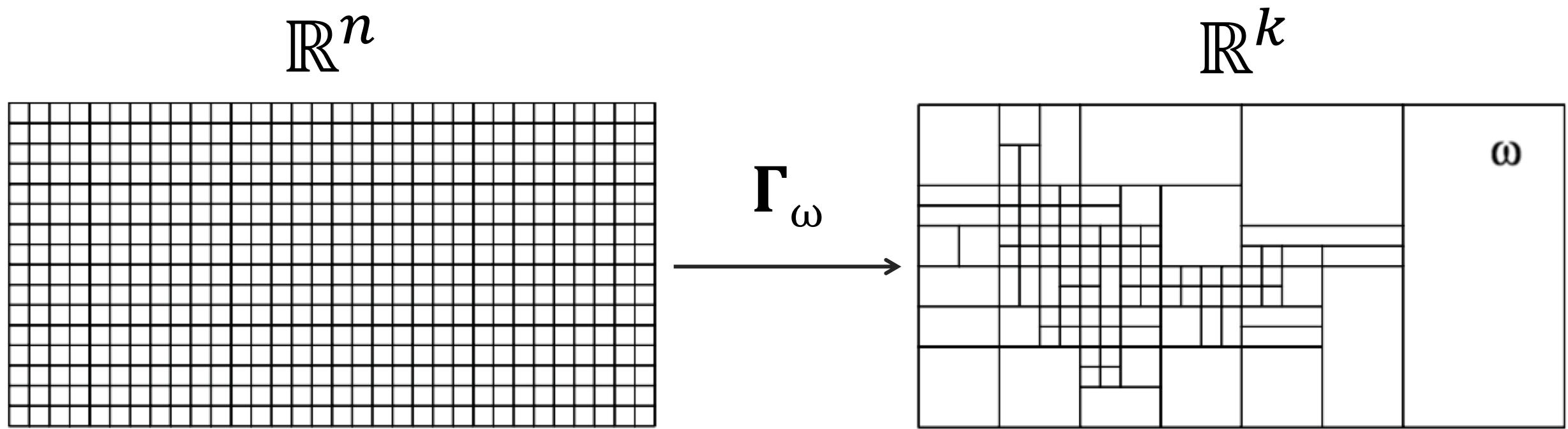
- i. increase resolution;
- ii. minimize computational cost;
- iii. and maintain information content?

Reduced-Rank Jacobian Construction

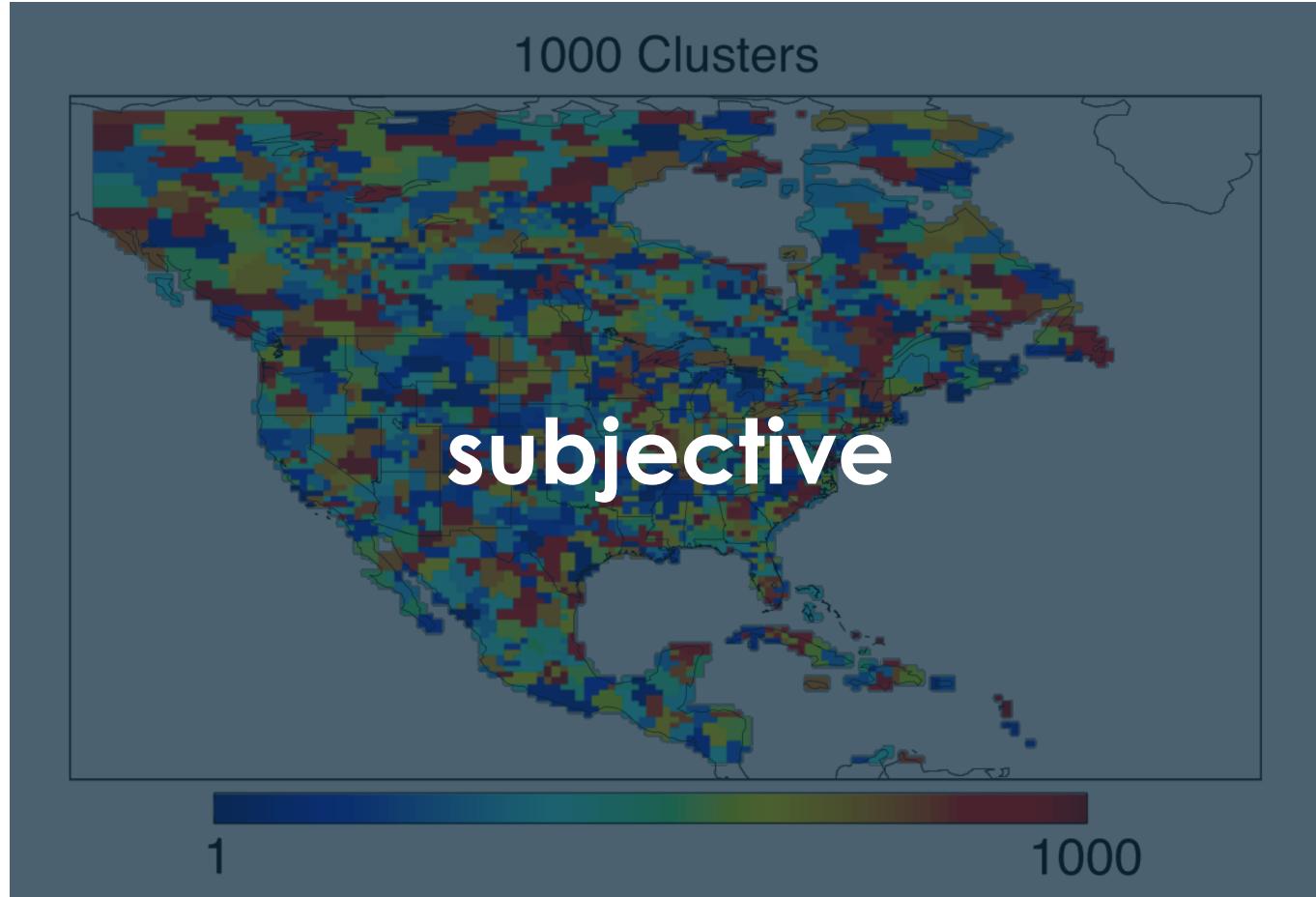
I. REDUCED-RANK INVERSIONS (BOUSSEREZ ET AL. 2018)

II. REDUCED-RANK JACOBIANS

Reducing the dimension of the state space from n to $p \ll n$ can reduce the computational cost of an analytic inversion

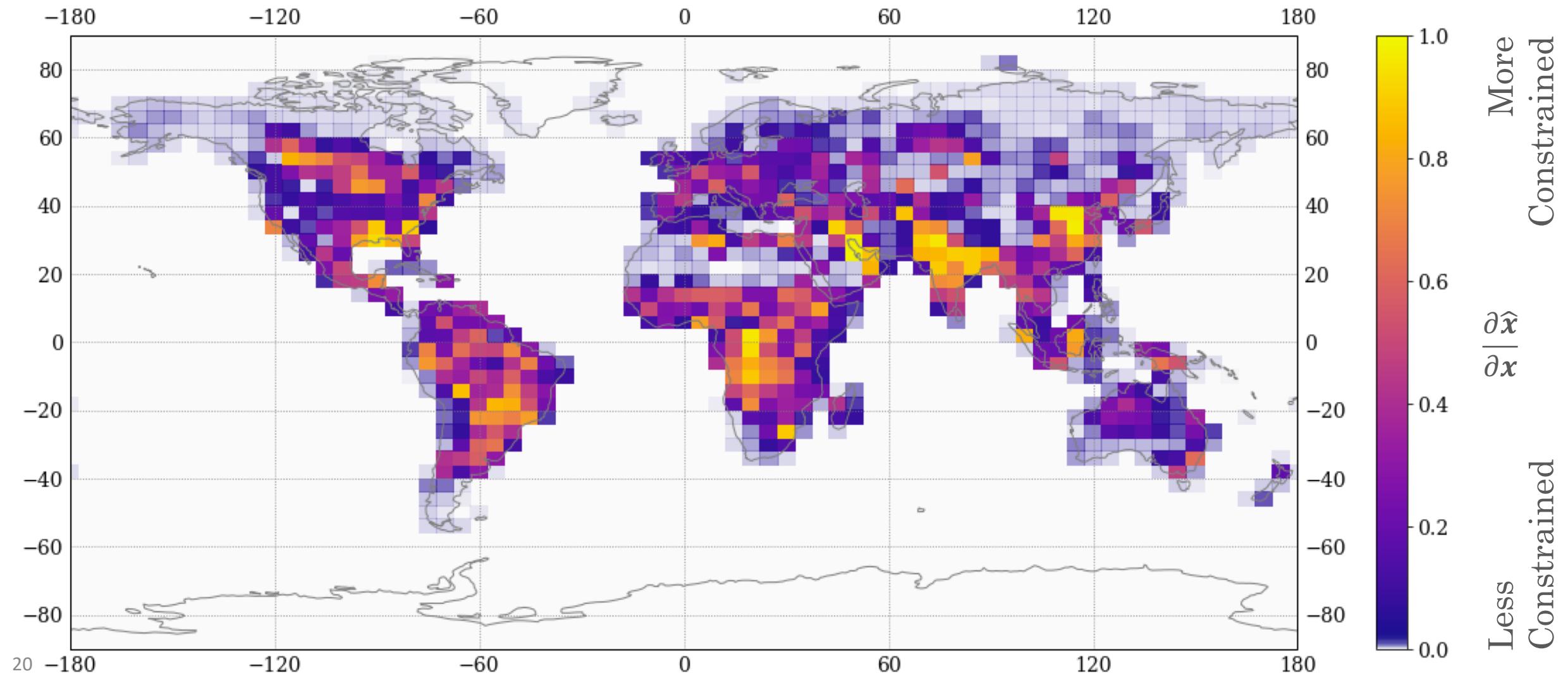


There are multiple methods of discretely clustering grid cells to reduce the state space dimension

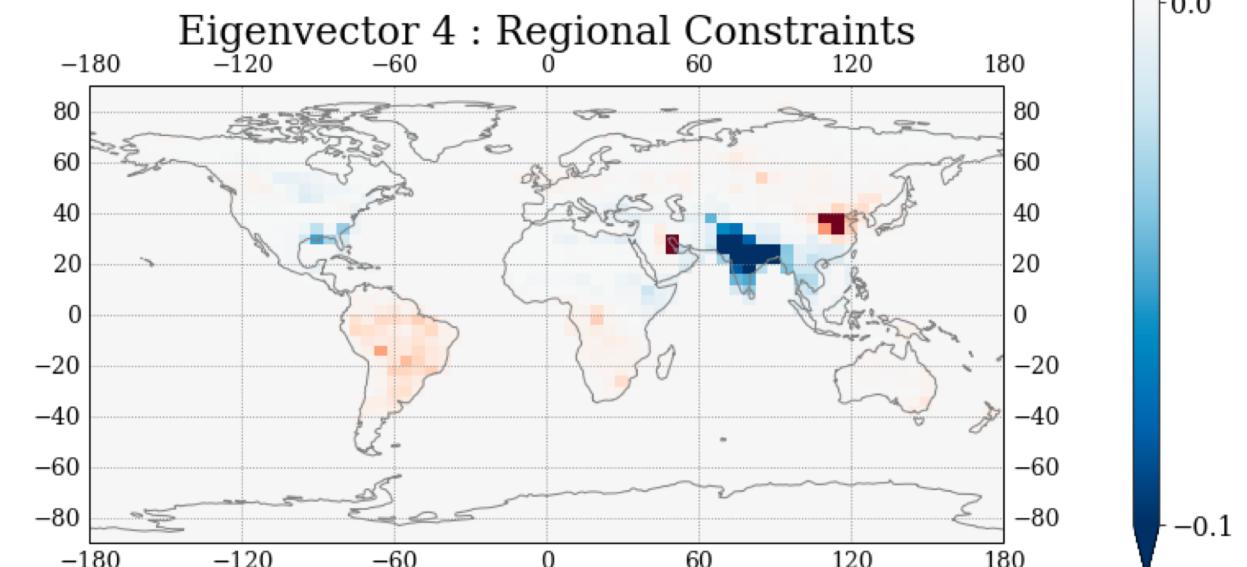
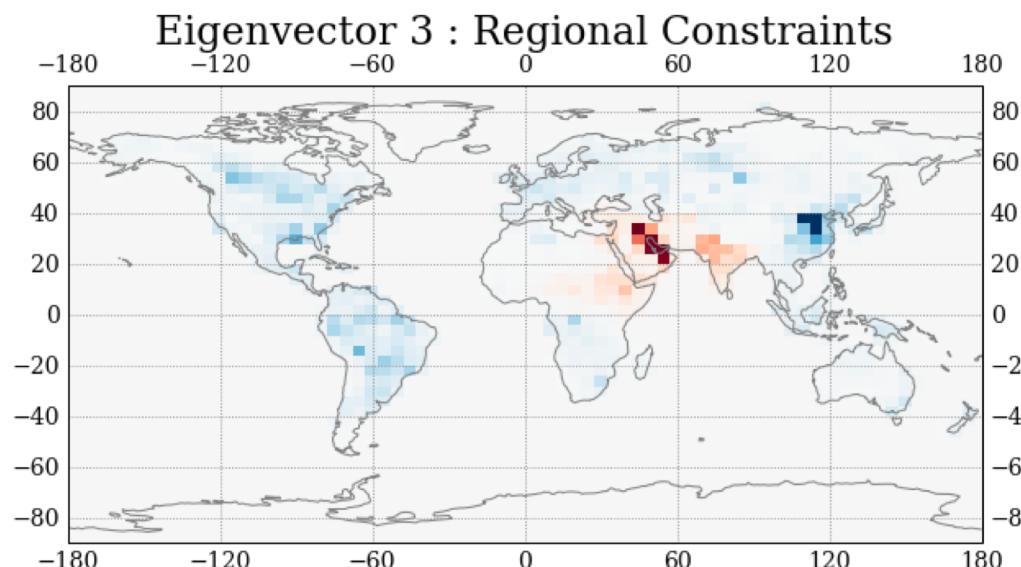
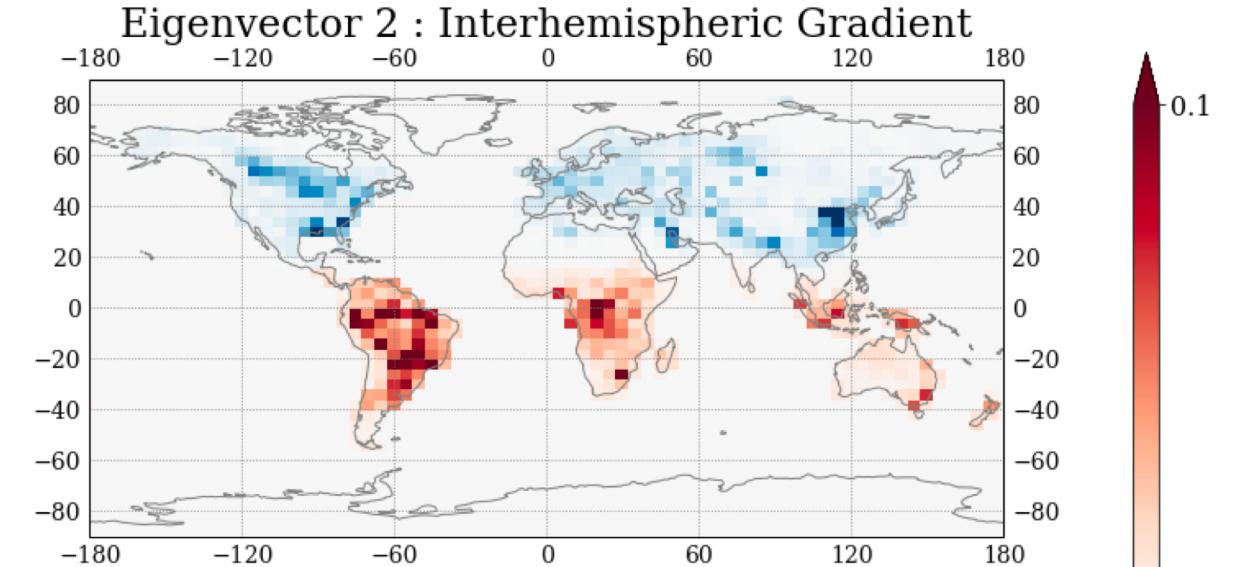
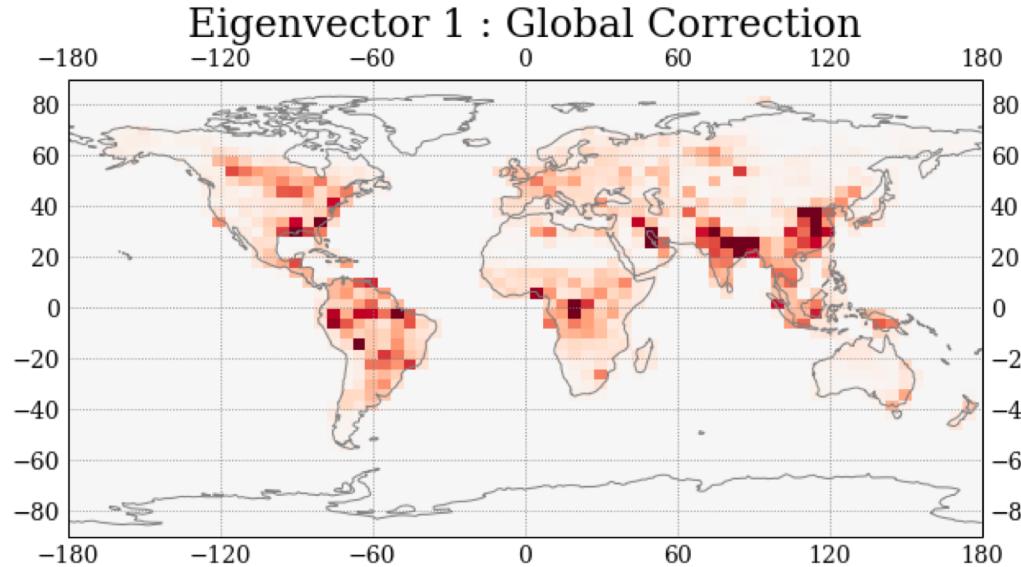


Known information about the emission-observation system can be used to find an optimal clustering analytically

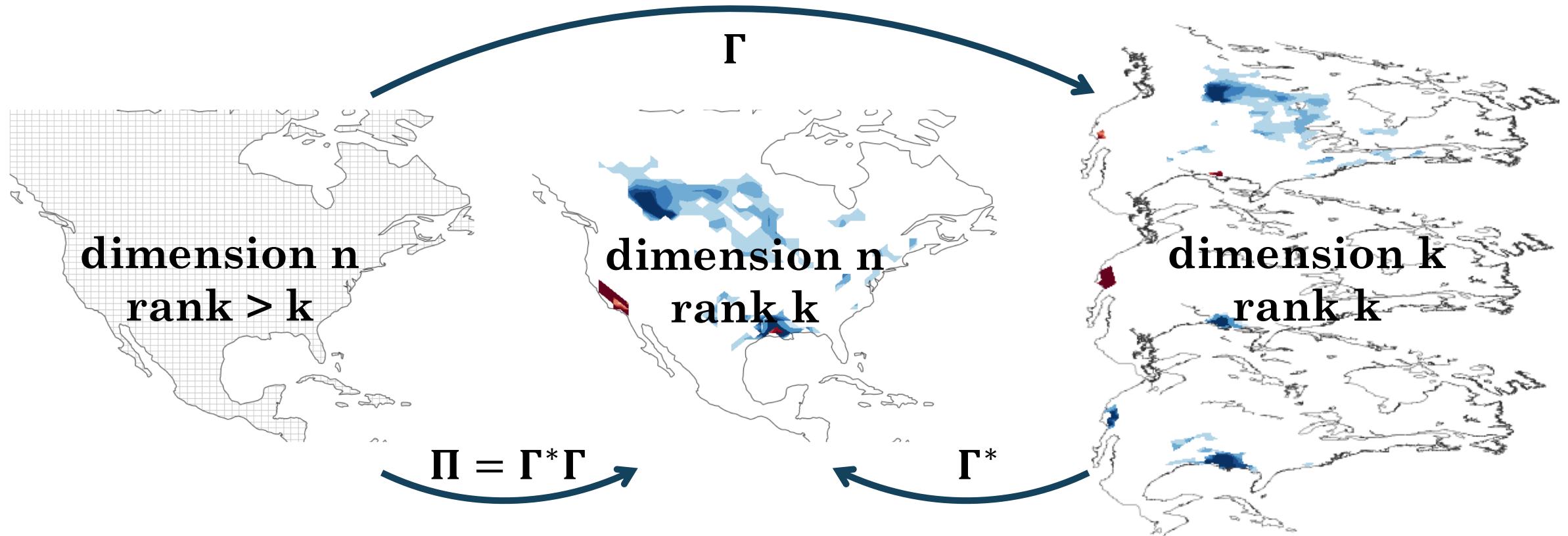
$$\text{Averaging Kernel : } \mathbf{A} = \frac{\partial \hat{x}}{\partial x} = f(\mathbf{S}_A, \mathbf{S}_O, \mathbf{K})$$



An eigendecomposition of $A Q_{DOF} H_P$ gives the patterns of information in the emission-observation system



The inversion can be solved in the directions given by the eigenvectors in a reduced-dimension or reduced-rank space

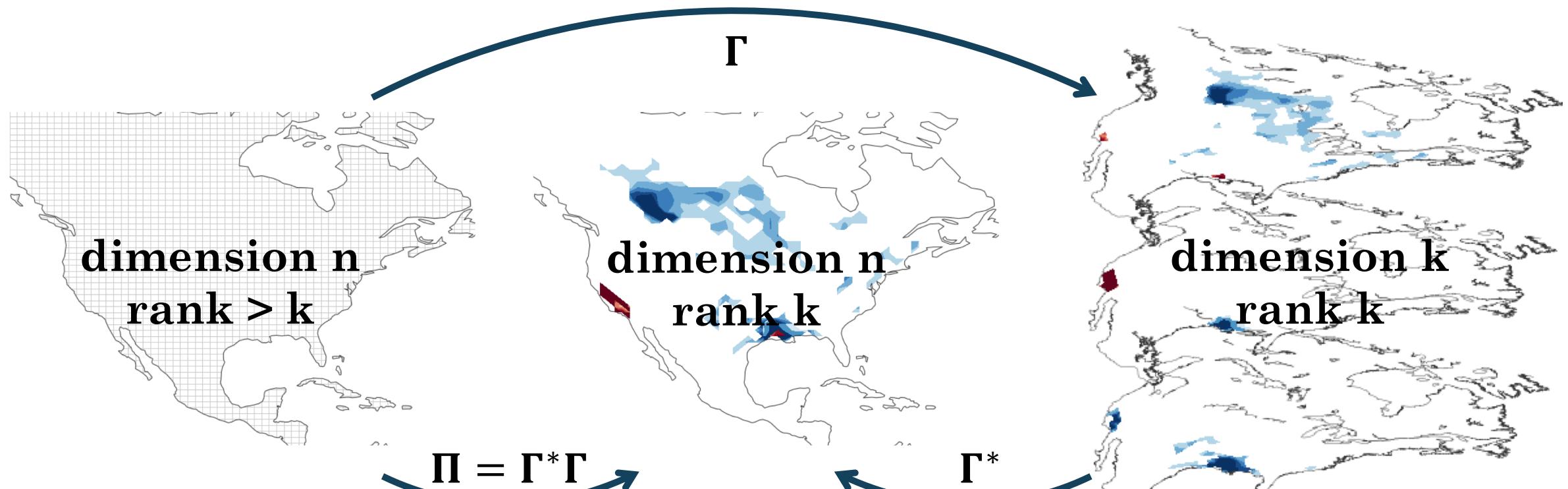


The inversion can be solved in the directions given by the eigenvectors in a reduced-dimension or reduced-rank space

$$\hat{\mathbf{x}}, \hat{\mathbf{S}}, \mathbf{A}$$

$$\hat{\mathbf{x}}_\Pi, \hat{\mathbf{S}}_\Pi, \mathbf{A}_\Pi$$

$$\hat{\mathbf{x}}_\omega, \hat{\mathbf{S}}_\omega, \mathbf{A}_\omega$$

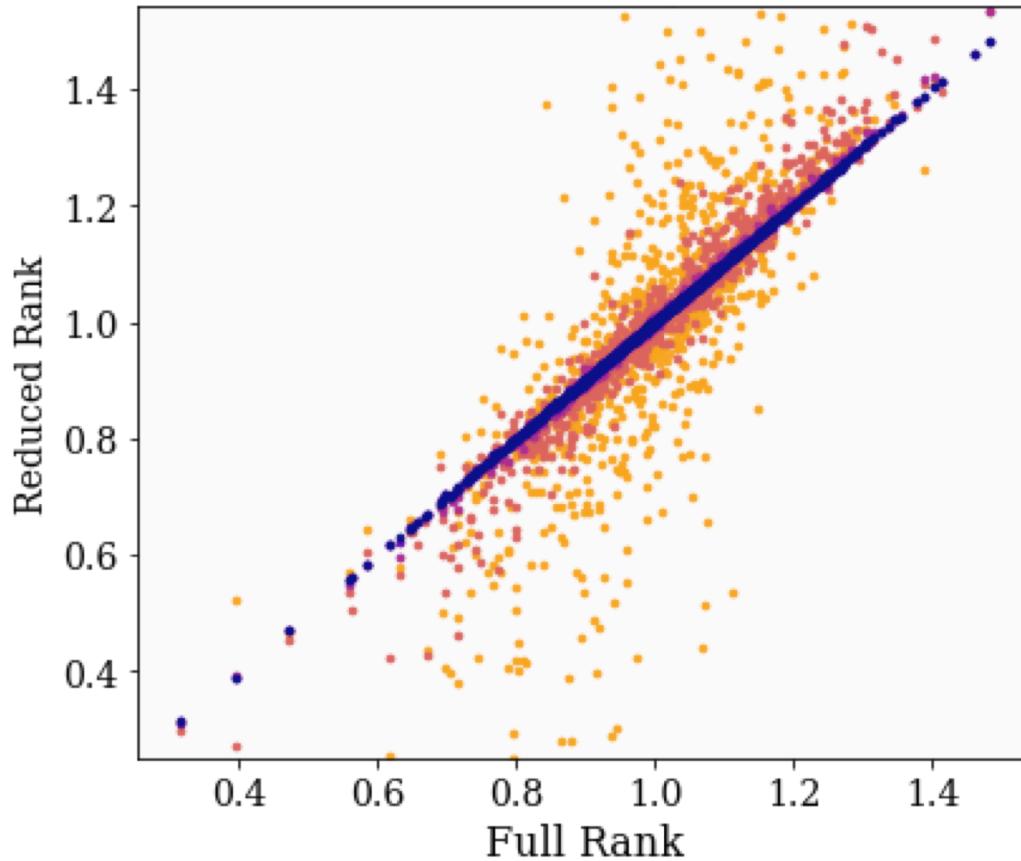


$$\hat{\mathbf{x}}_{K_\Pi}, \hat{\mathbf{S}}_{K_\Pi}, \mathbf{A}_{K_\Pi}$$

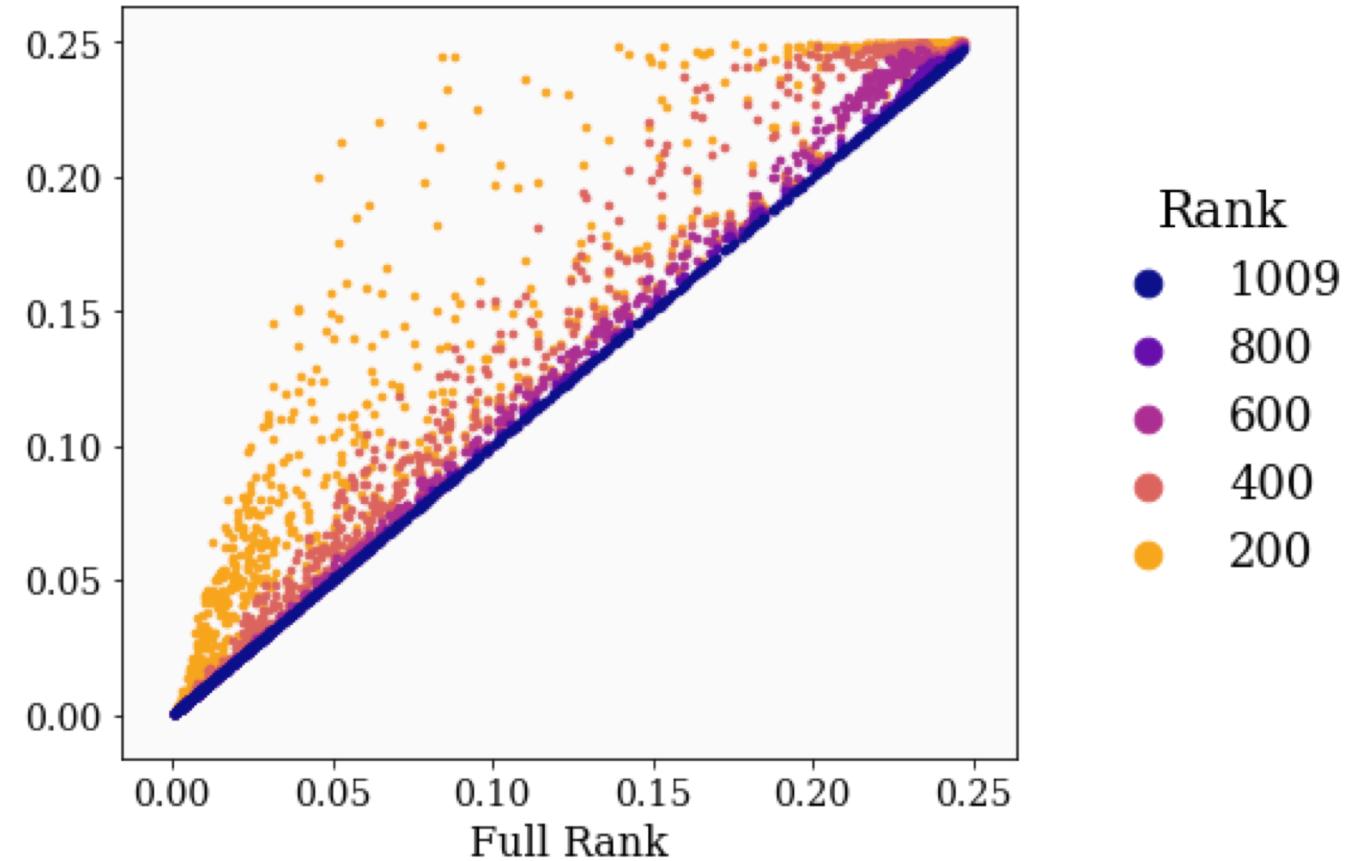
$$\hat{\mathbf{x}}_{FR}$$

The full rank approximation reproduces the full rank solution and converges to the true solution as $k \rightarrow n$

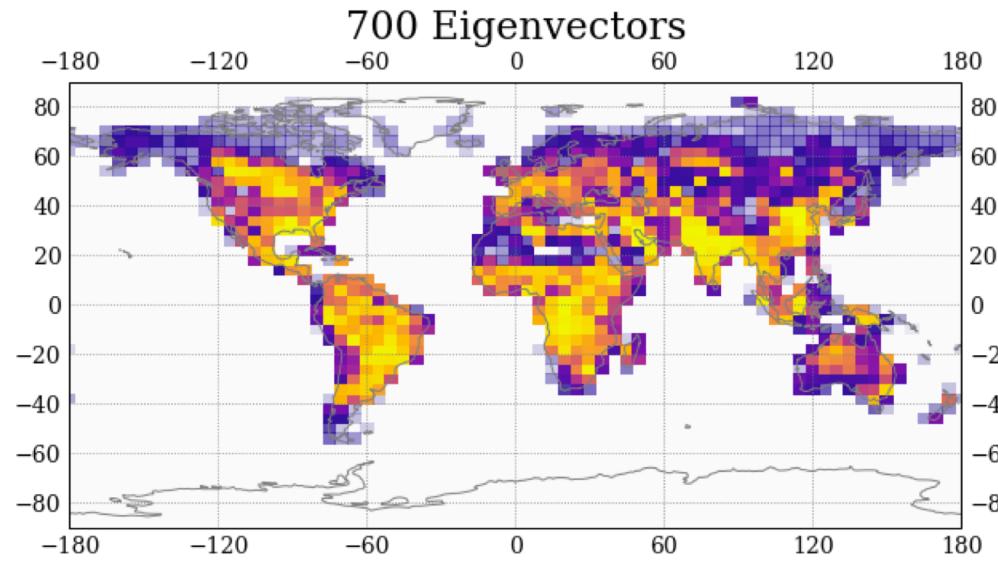
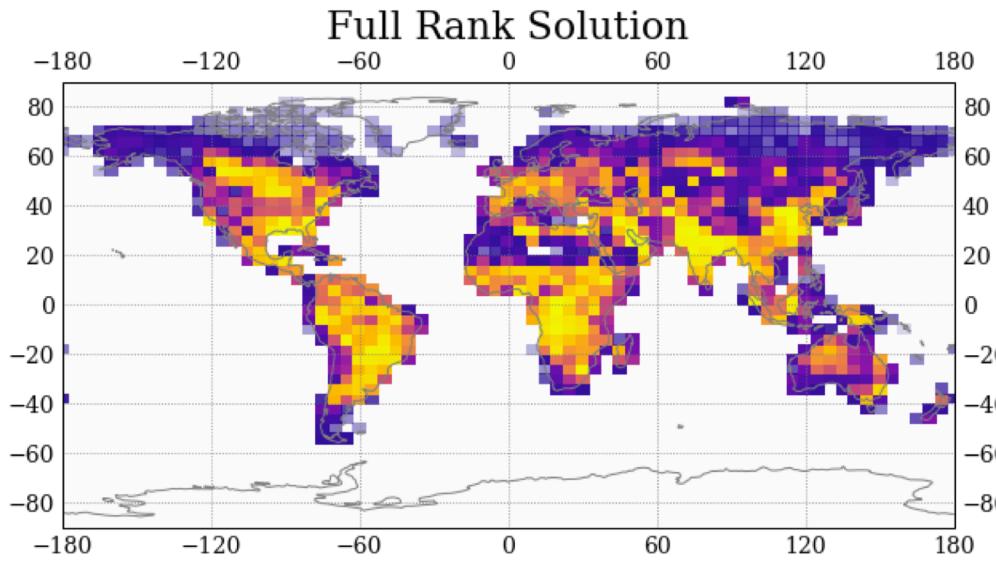
Full Rank Approximation
Posterior Mean



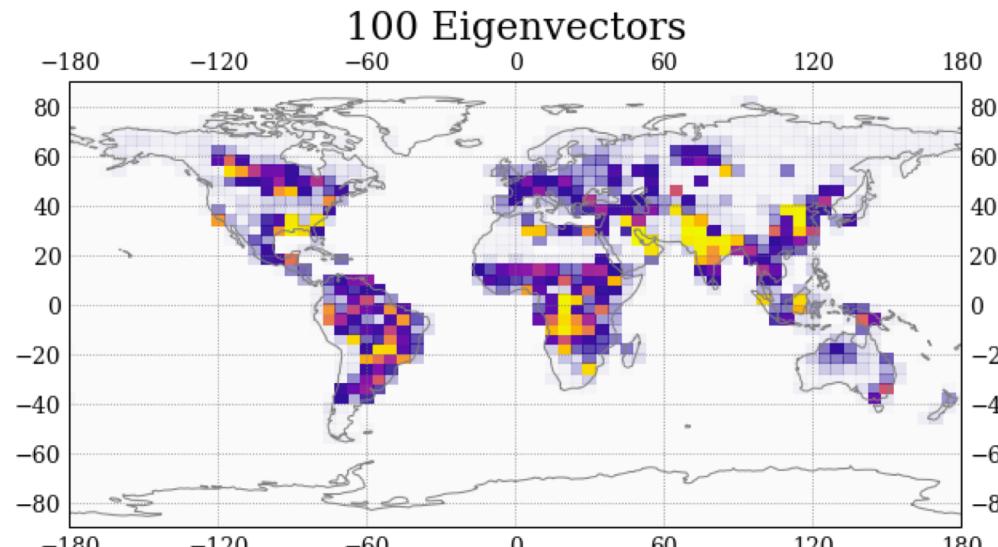
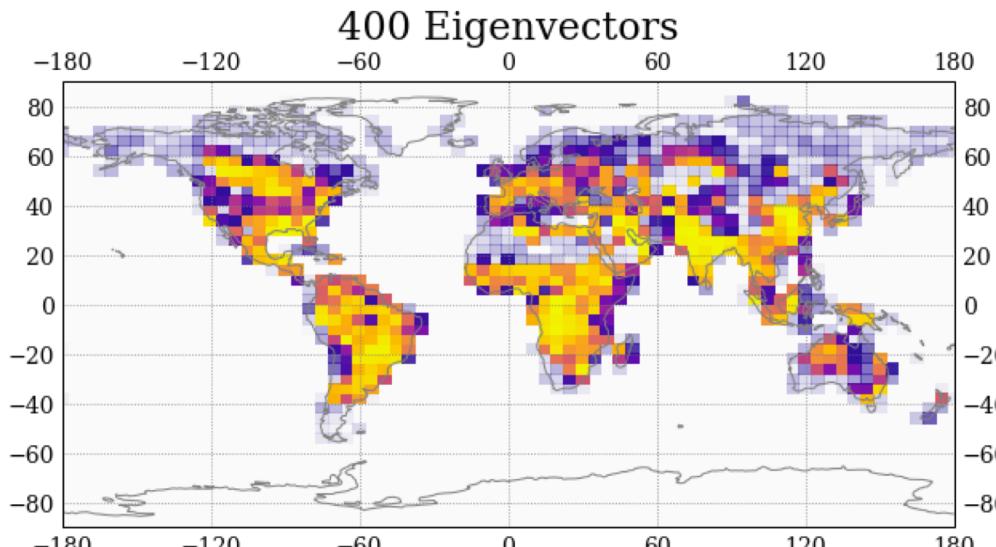
Projected Forward Model
Posterior Error Variance



The “reduced-rank inversion” maximizes the information content of the posterior solution



More Constrained

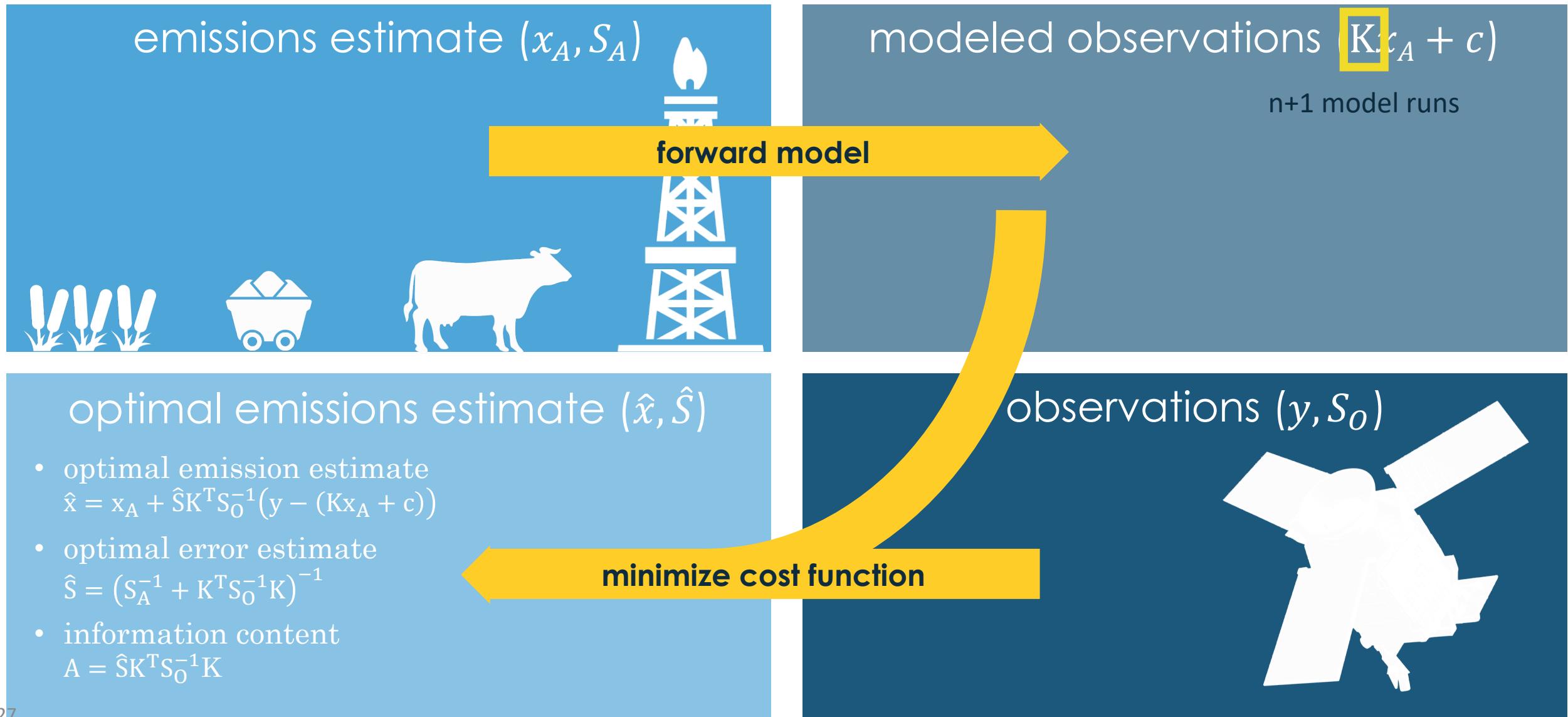


Less Constrained

Reduced-Rank Jacobian Construction

Reduced-rank inversions decrease computational cost without significant loss of information content in the posterior solution.

Increasing inversion resolution increases computational cost, which is limited by the number of grid boxes optimized



Reduced-Rank Jacobian Construction

Reduced-rank inversions decrease computational cost without significant loss of information content in the posterior solution.

Next steps:

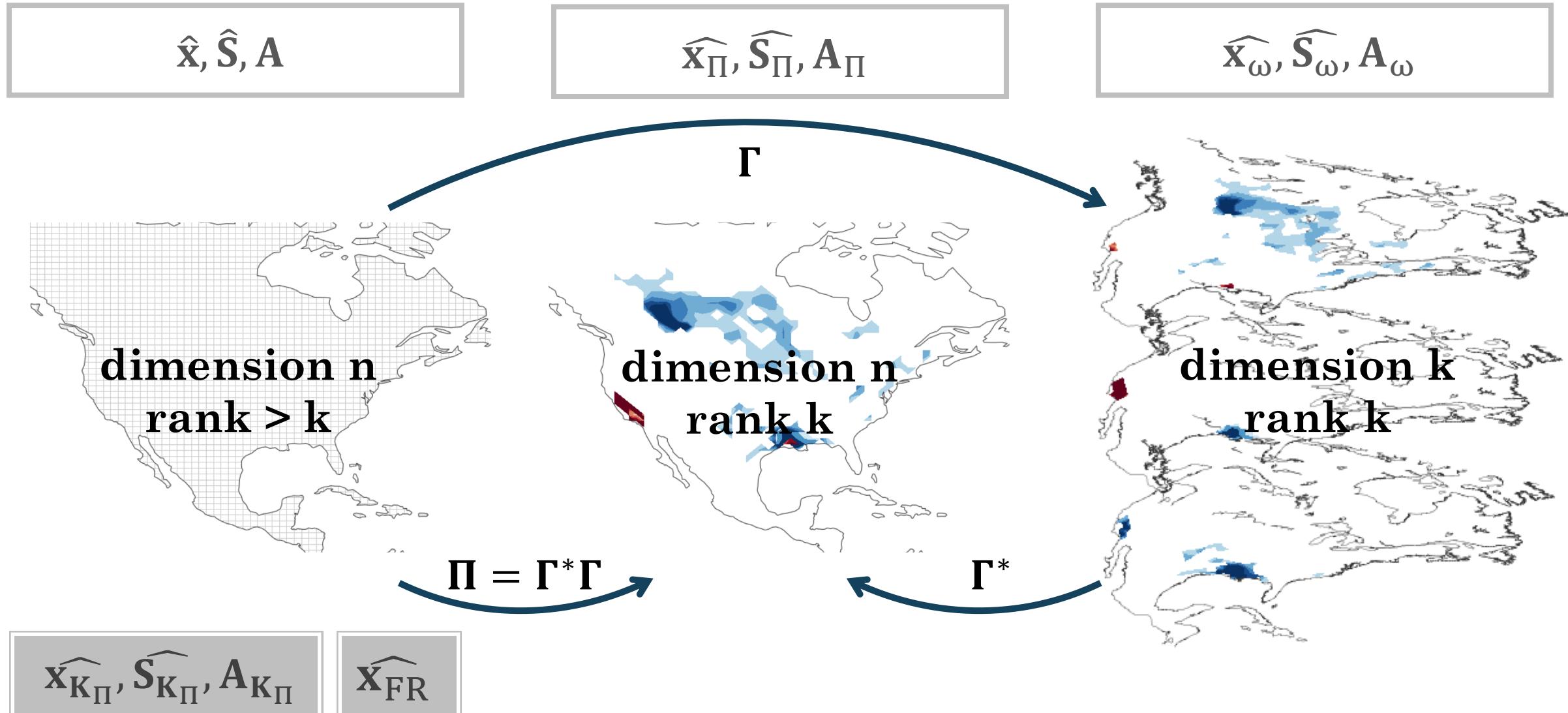
- I. Reduce the computational cost of constructing Jacobians for analytic Bayesian inversions.

Reduced-Rank Jacobian Construction

I. REDUCED-RANK INVERSIONS (BOUSSEREZ ET AL. 2018)

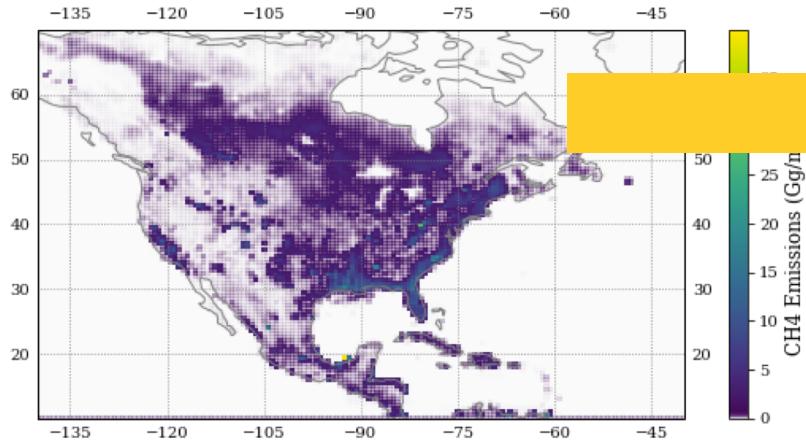
II. REDUCED-RANK JACOBIANS

The approximated posterior implicitly uses a reduced-rank Jacobian



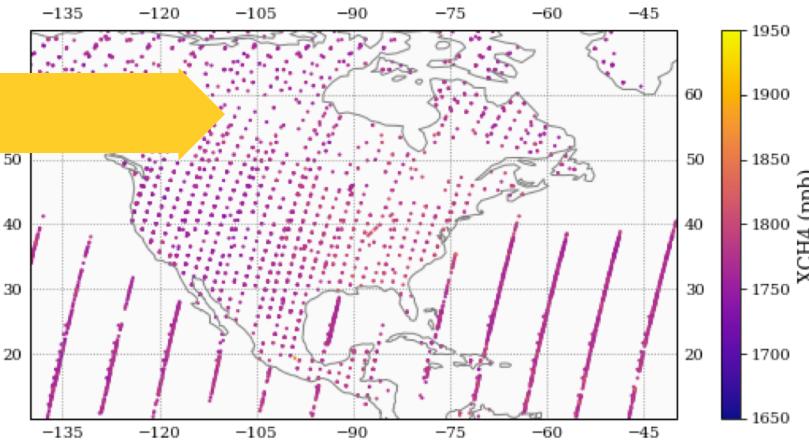
The Jacobian \mathbf{K} represents the sensitivity of observations to emissions: $\mathbf{y} = \mathbf{Kx} + \mathbf{c}$ with $\mathbf{K} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}}$

emissions estimate

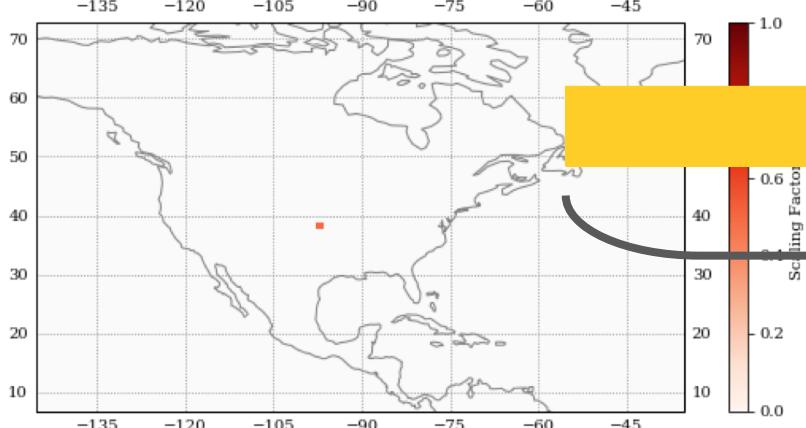


forward model

modeled observations

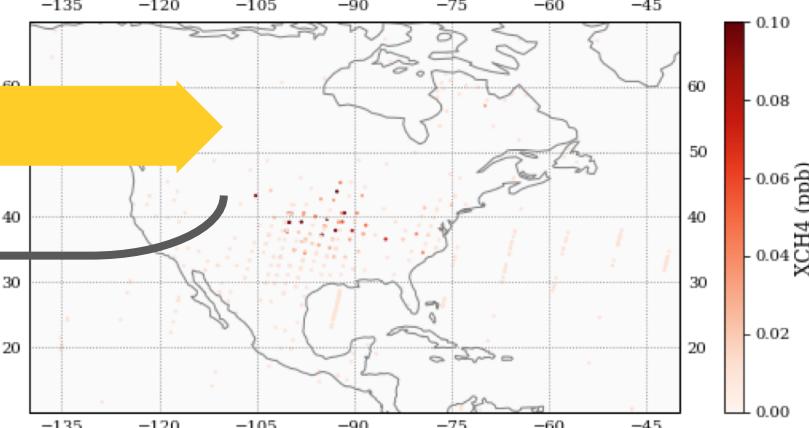


Δx



forward model

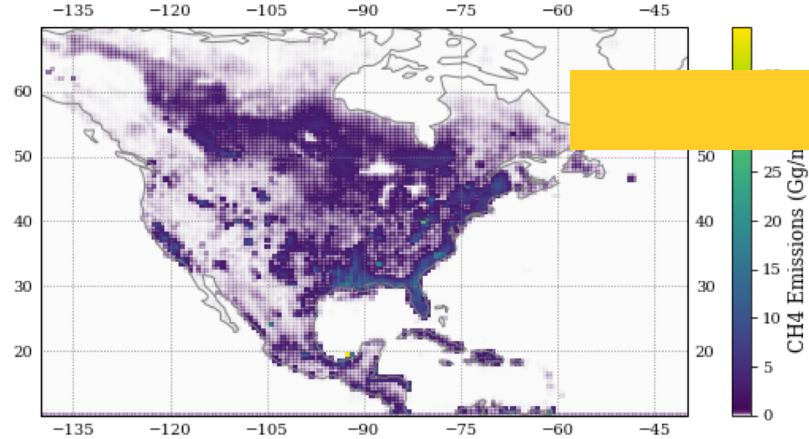
Δy



n perturbations,
 n model runs

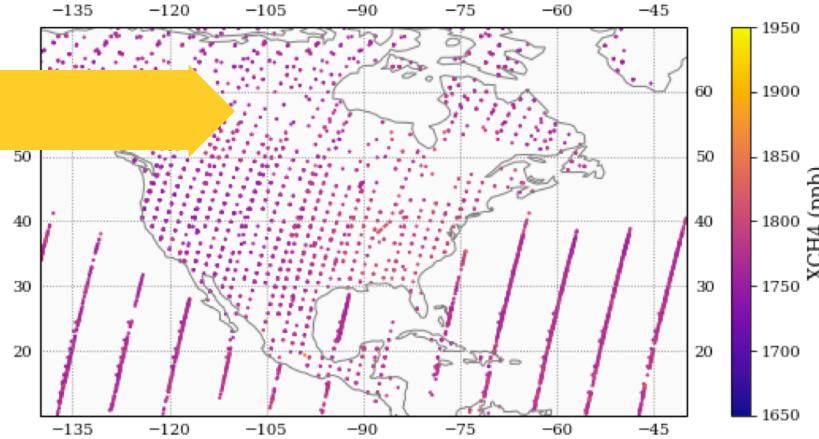
Perturbing eigenvectors would require $k < n$ model runs and yield a reduced-rank Jacobian

emissions estimate

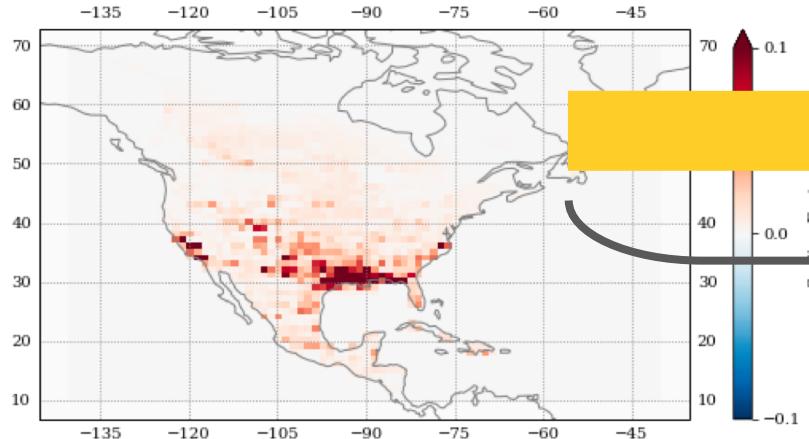


forward model

modeled observations



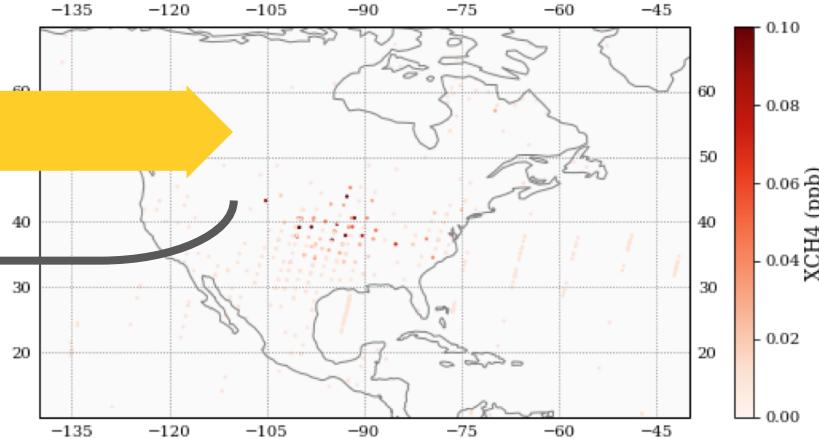
Δx



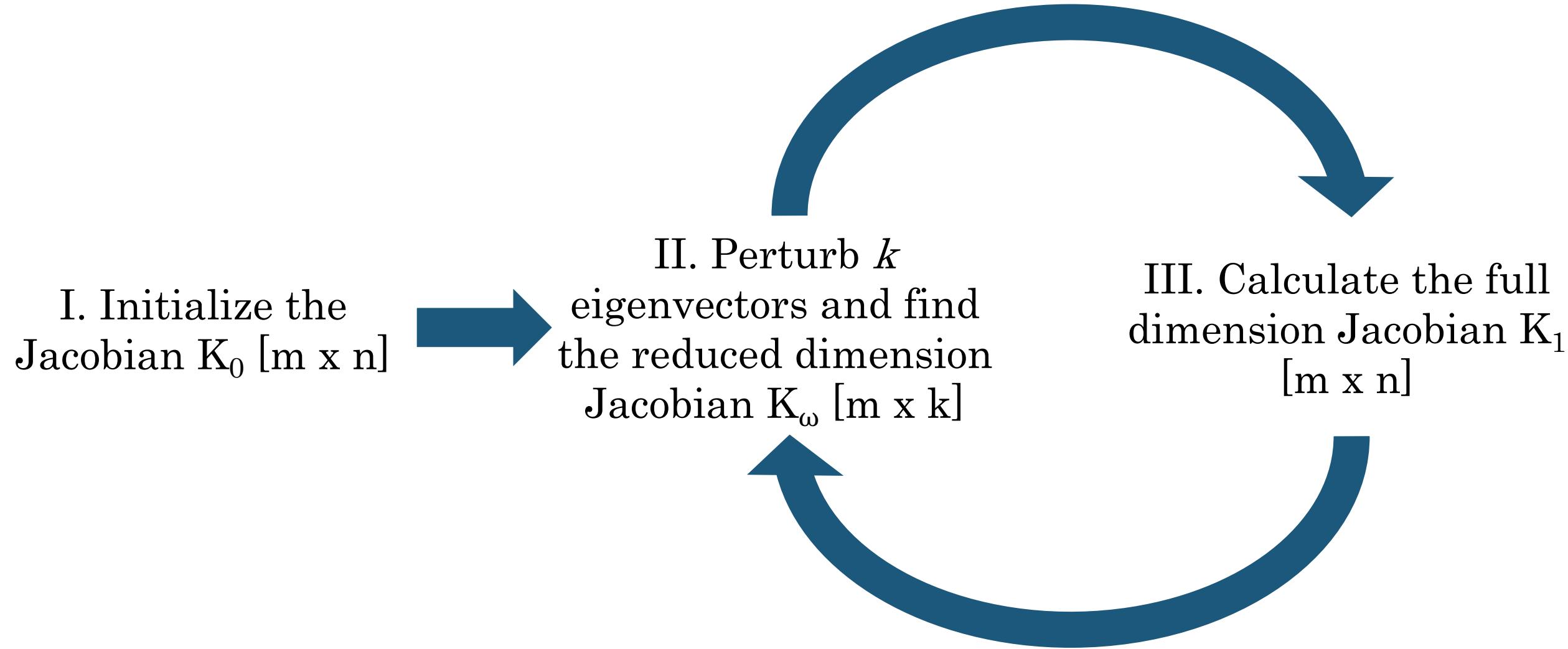
forward model

Δy

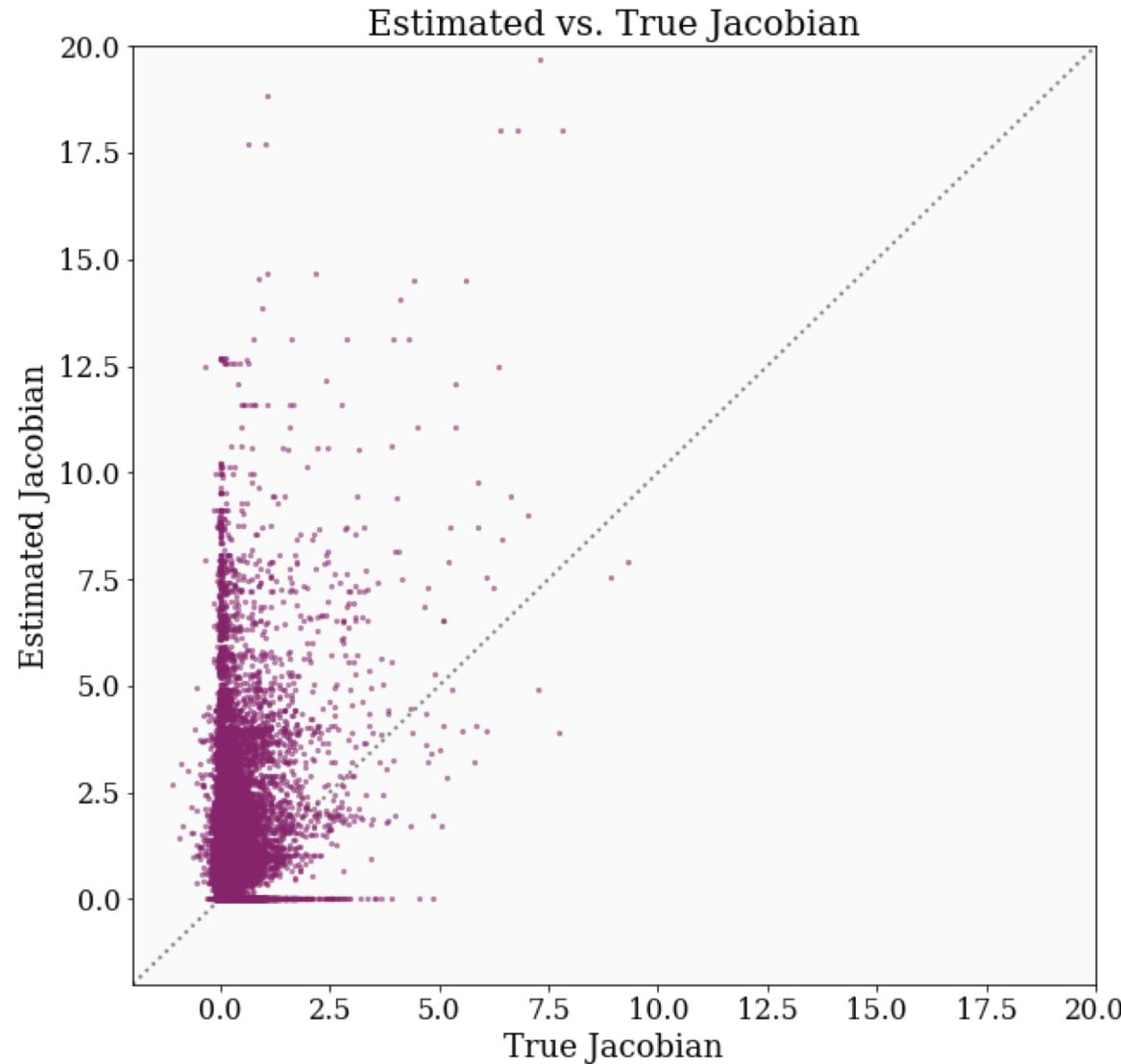
k perturbations,
 k model runs



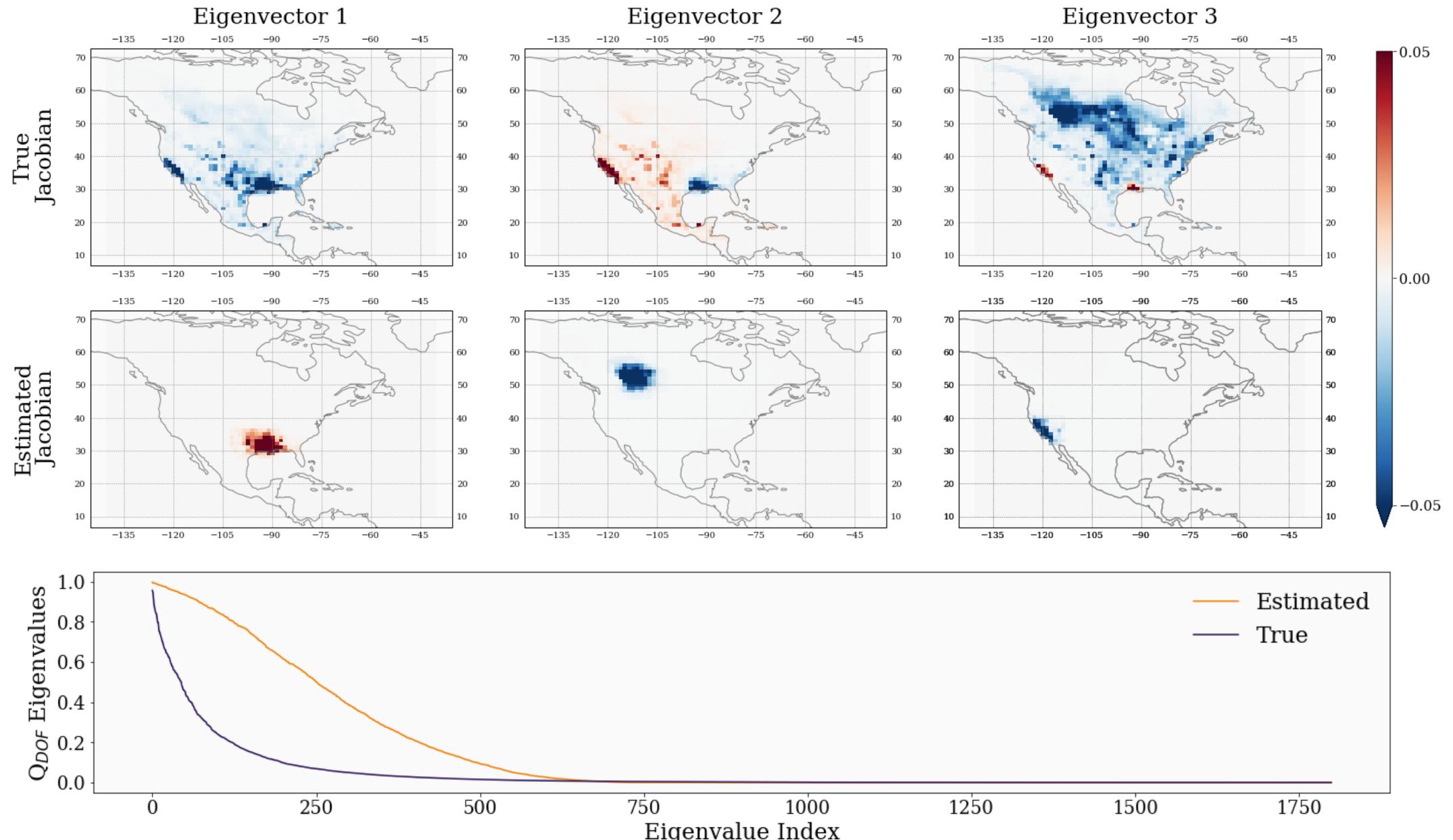
Because the eigenvectors include a contribution from the model, constructing the Jacobian is an iterative process



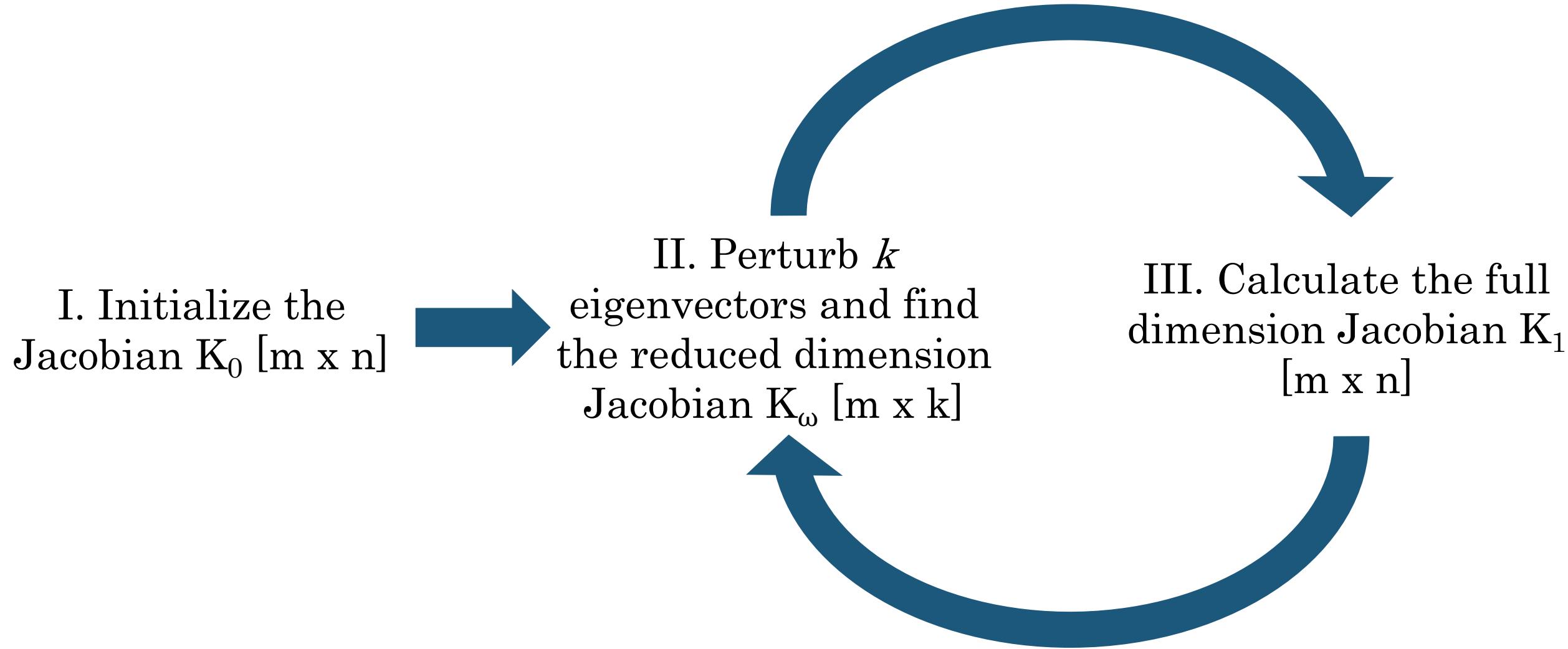
The Jacobian can be initialized using a mass-balance approach



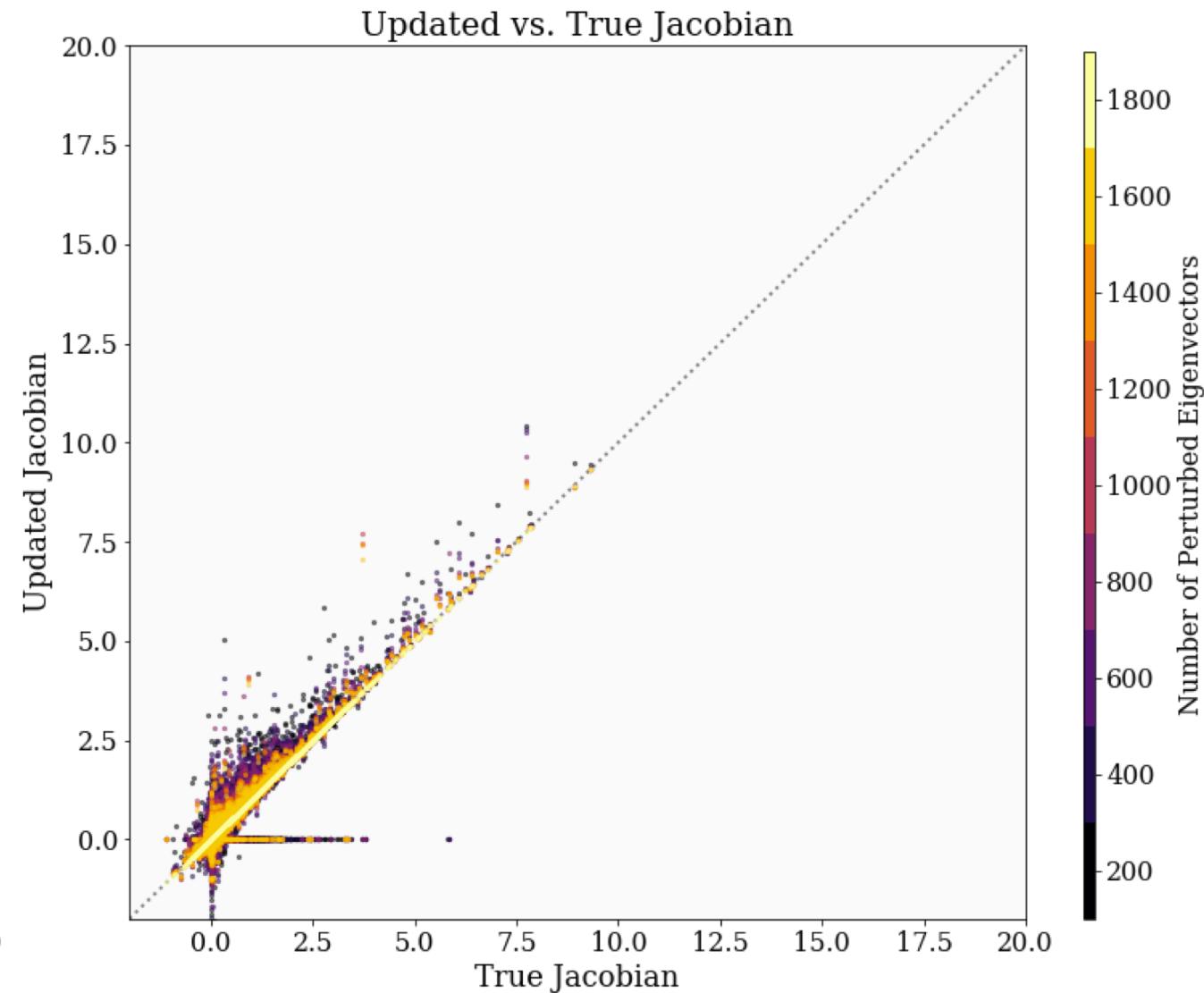
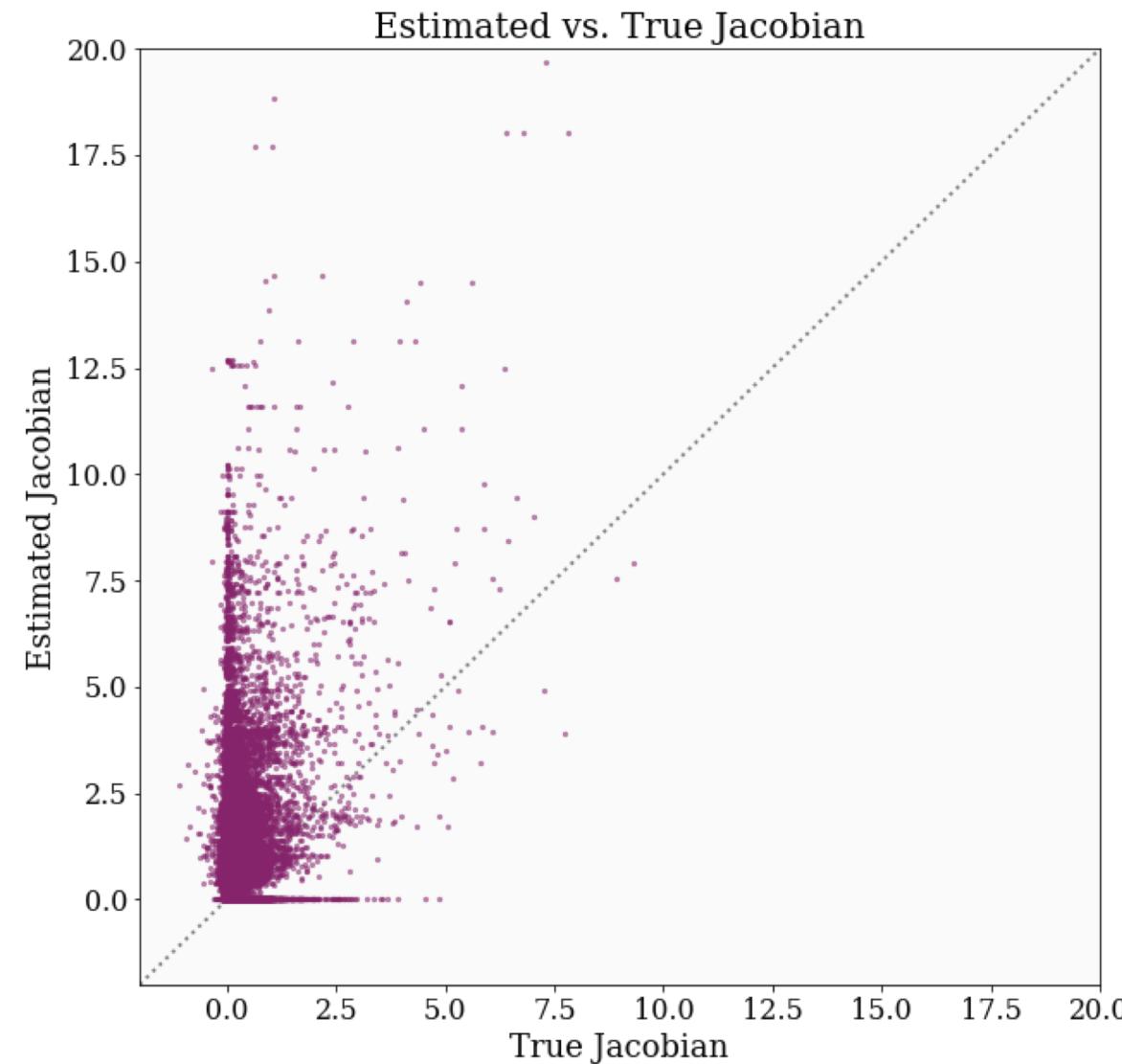
The initial estimate produces sufficiently similar eigenpairs



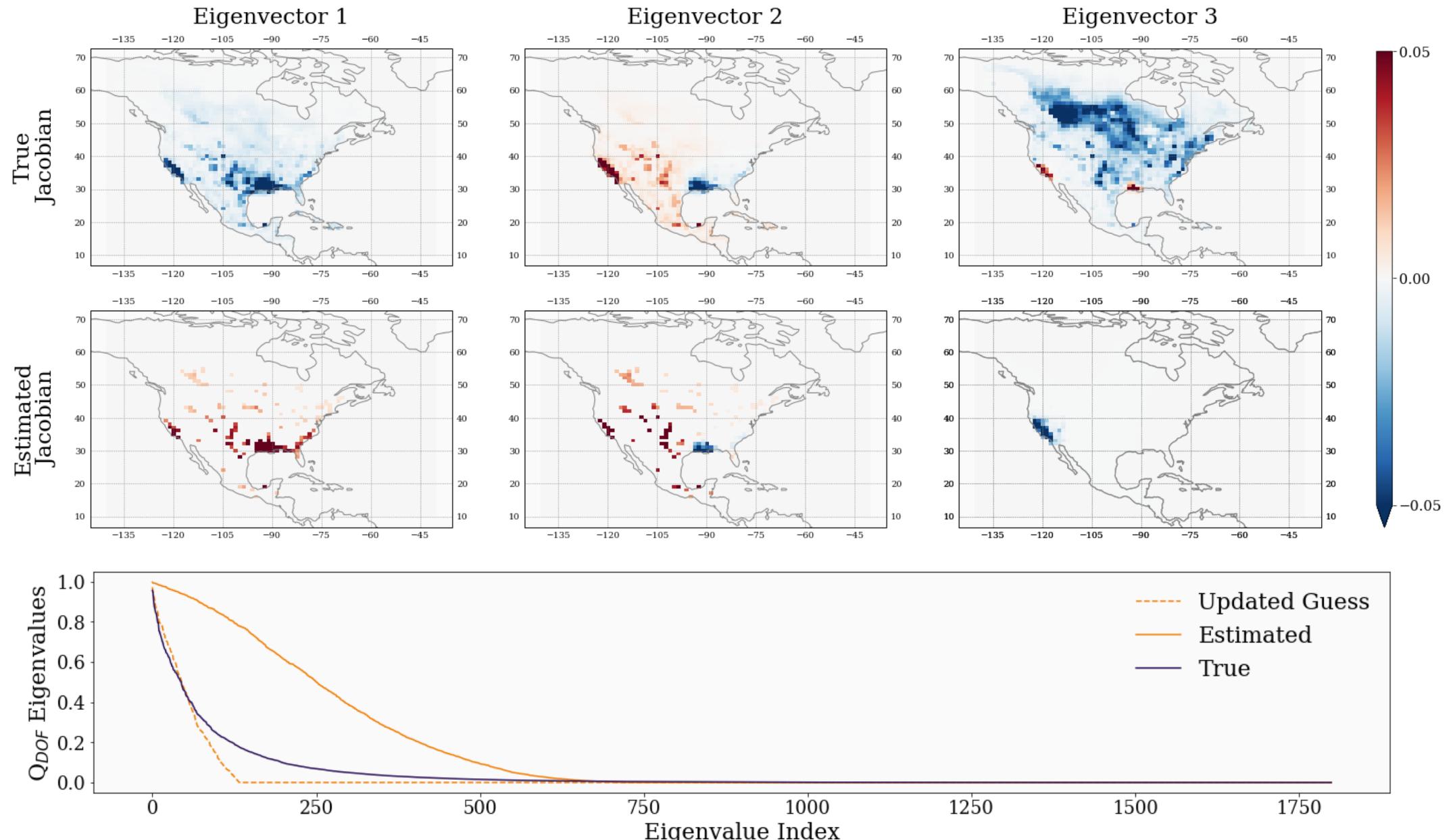
Because the eigenvectors include a contribution from the model, constructing the Jacobian is an iterative process



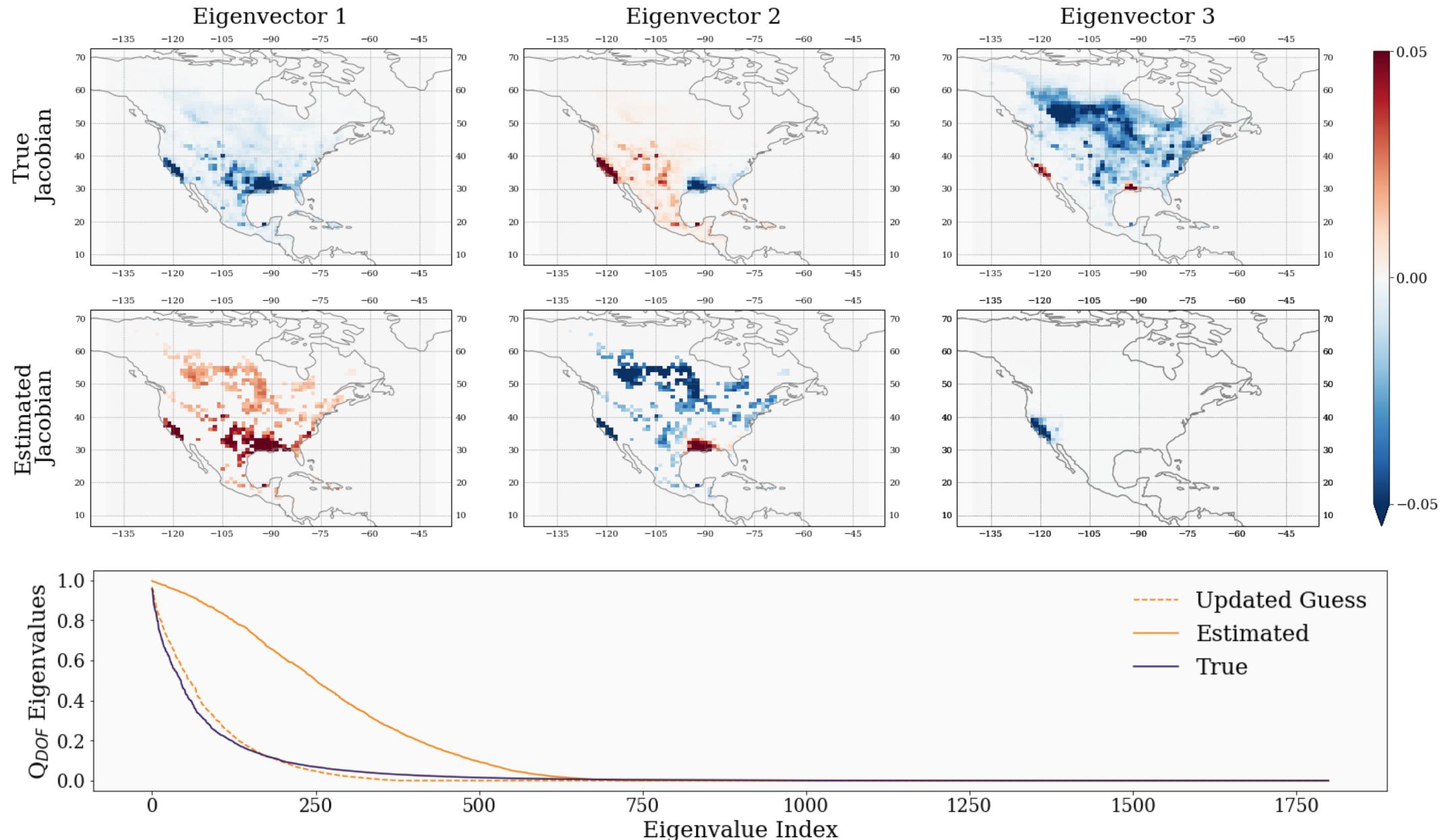
A single iteration significantly improves the estimated Jacobian



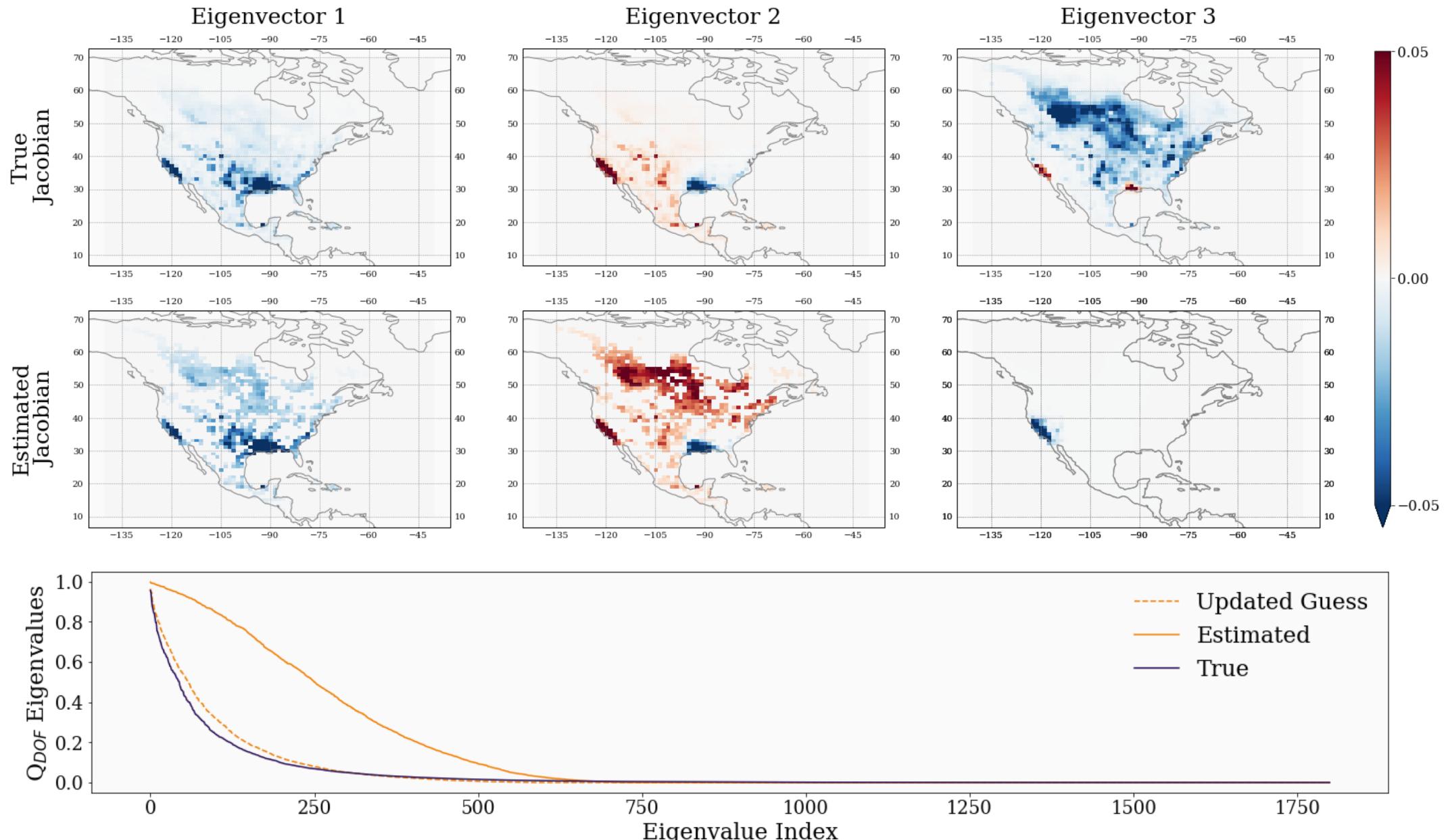
A single iteration with 200 perturbations improves the eigenpairs



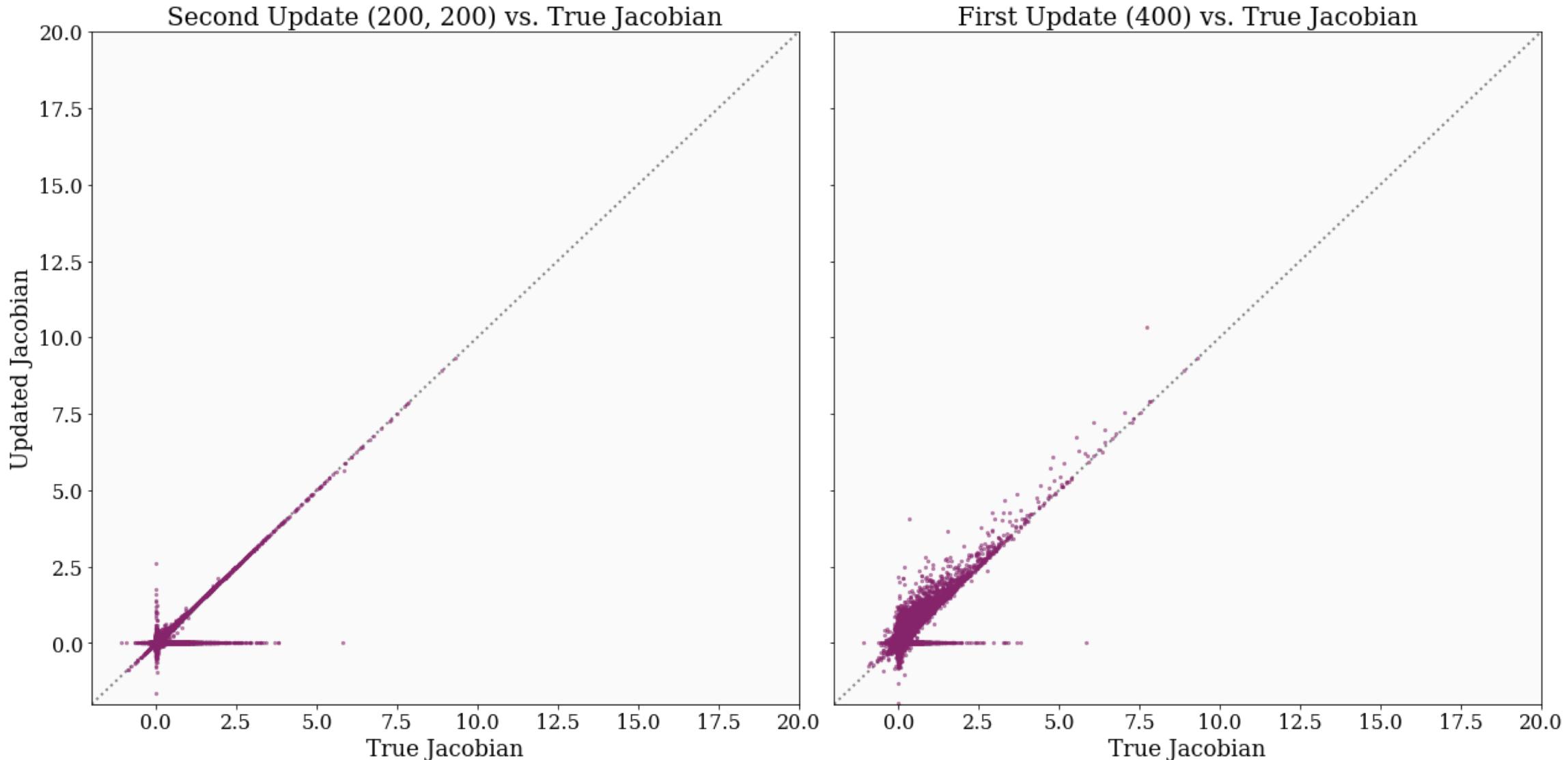
A single iteration with 400 perturbations improves the eigenpairs



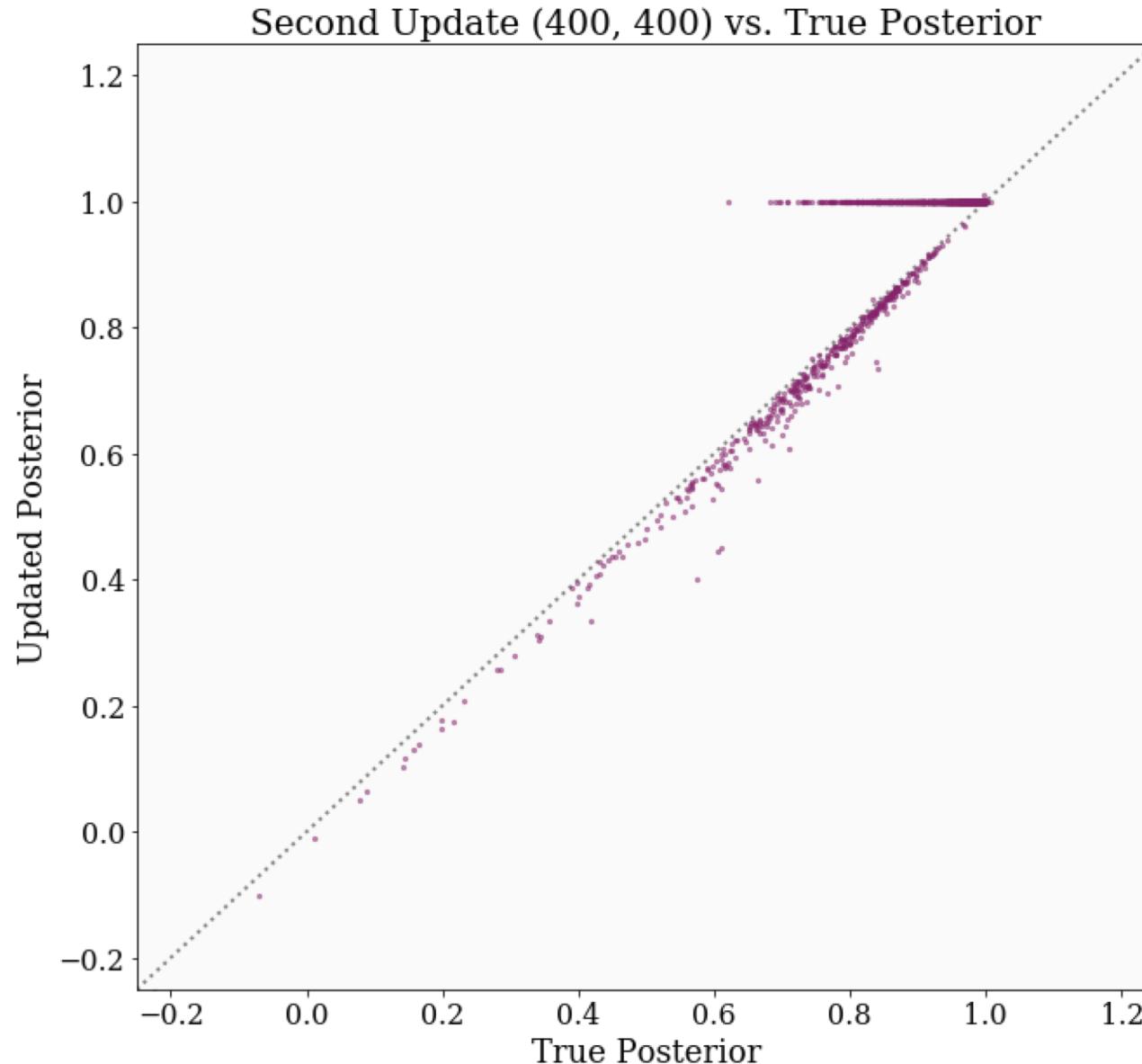
A single iteration with 600 perturbations improves the eigenpairs



Iterating improves the Jacobian estimate

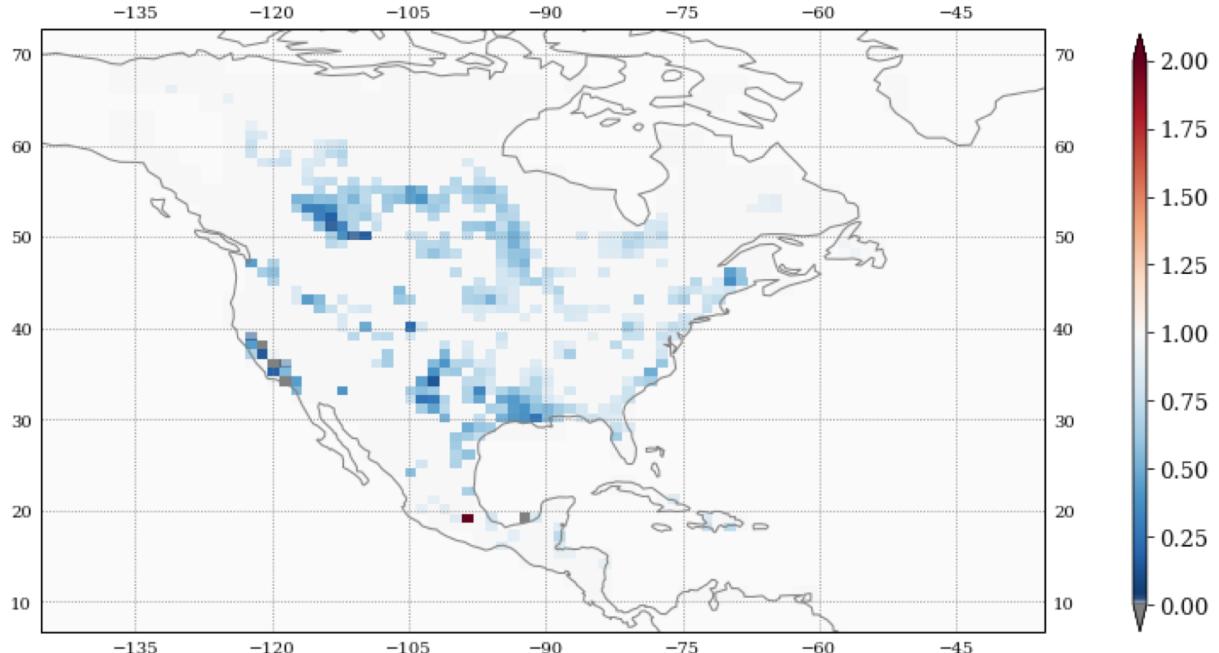


The resulting posterior approximates the true posterior with fewer than half the model runs

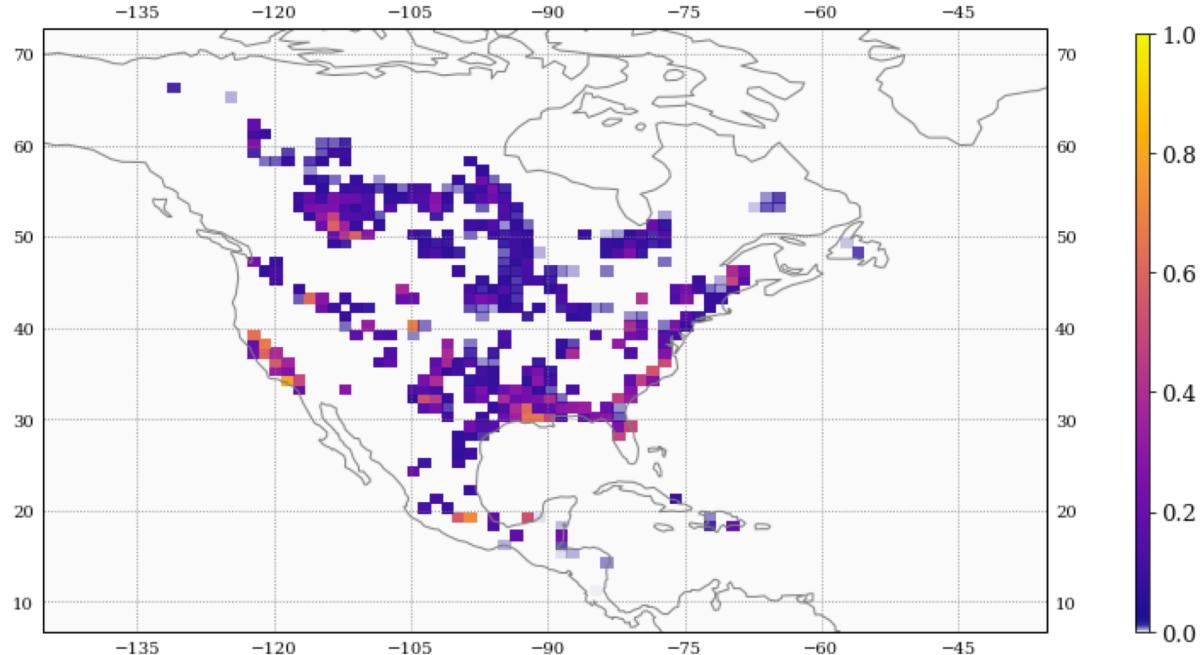


The resulting posterior approximates the true posterior best in areas with high information content

Estimated Posterior (400, 400)



Estimated Averaging Kernel (400, 400)



Reduced-Rank Jacobian Construction

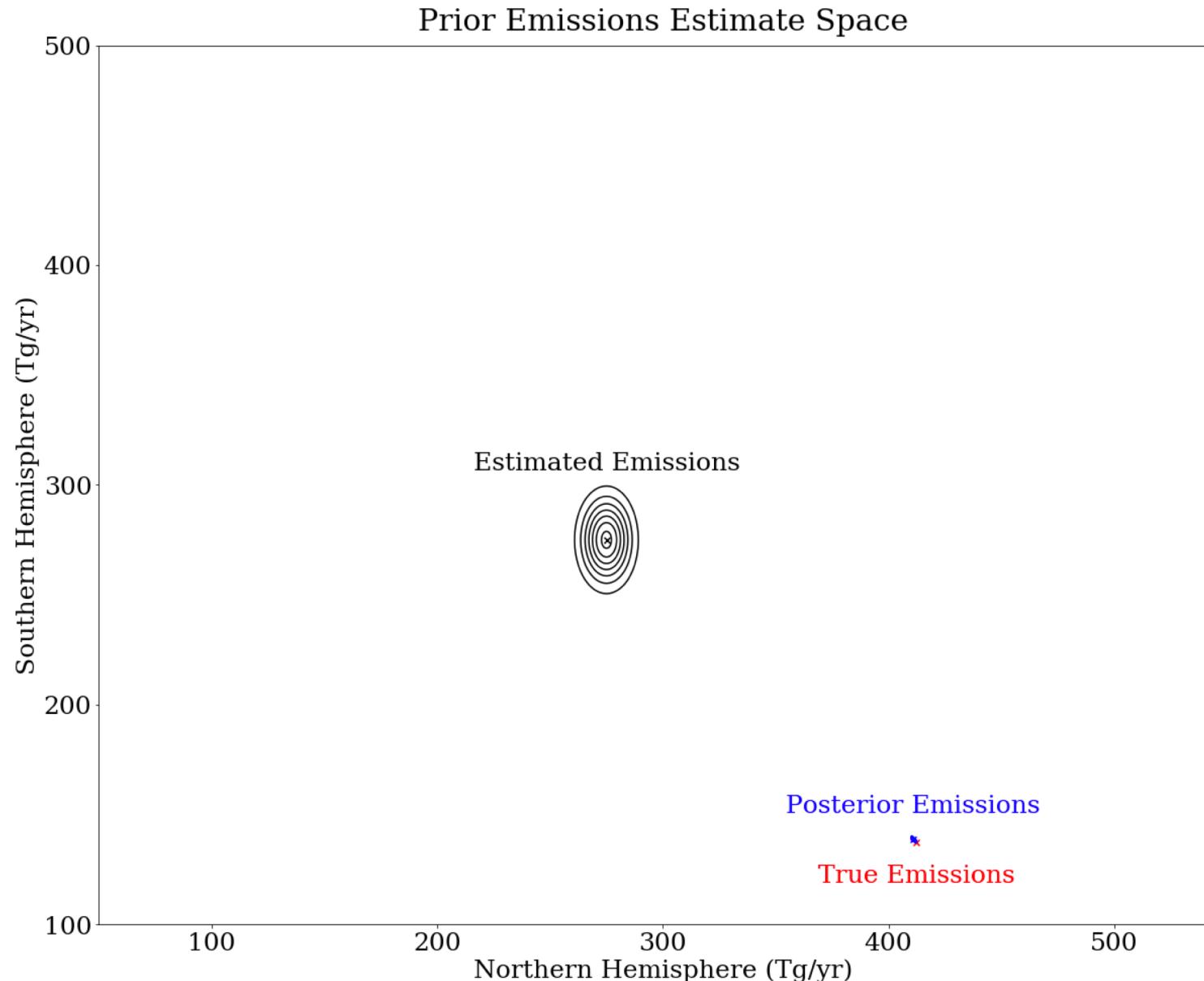
Reduced-rank Jacobians optimize methane emissions in areas with high information content while significantly decreasing computational cost.

Next steps:

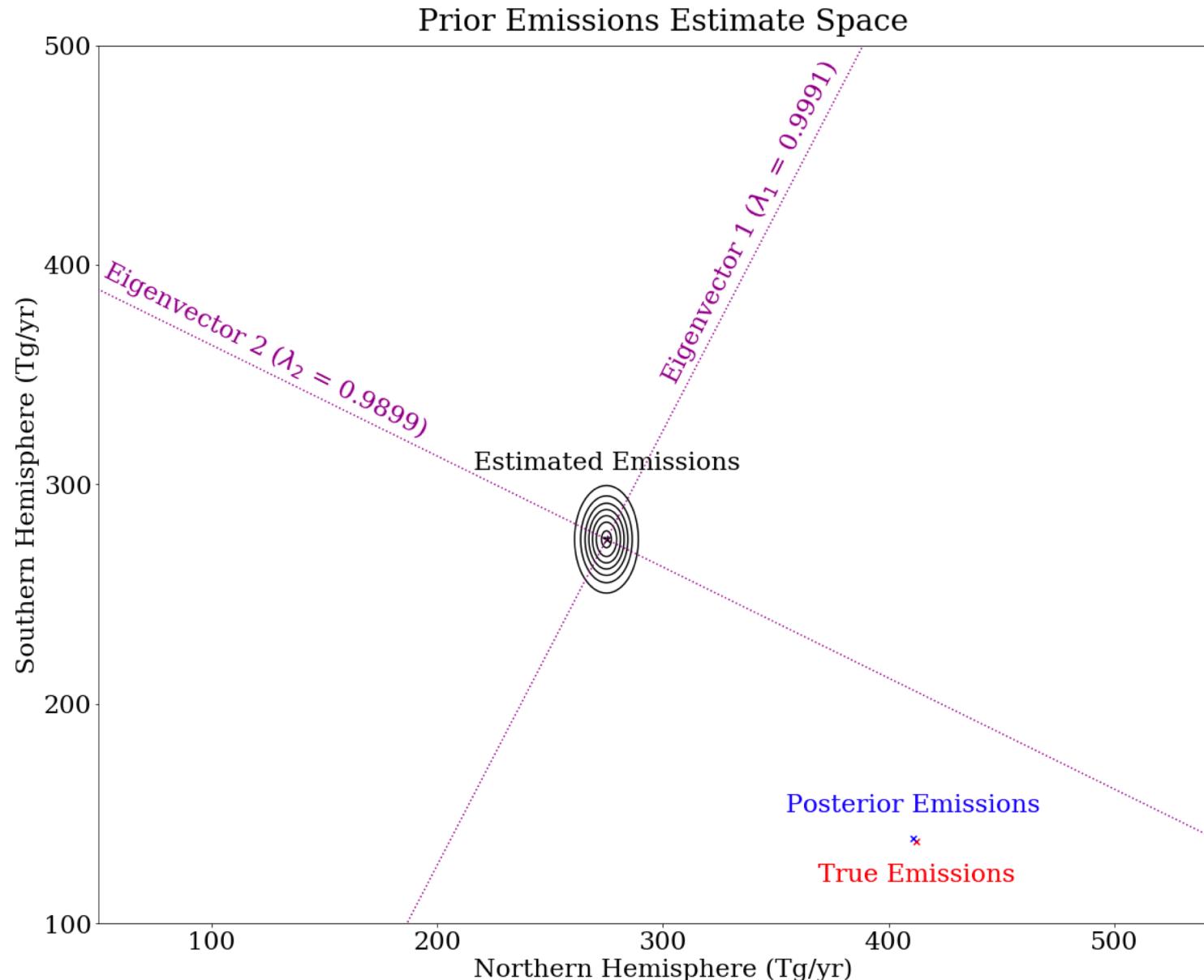
- I. Define and justify an optimal iteration scheme, including convergence criteria;
- II. Quantify the error associated with the reduced-rank Jacobian and the resulting posterior solutions.

Back-up slides

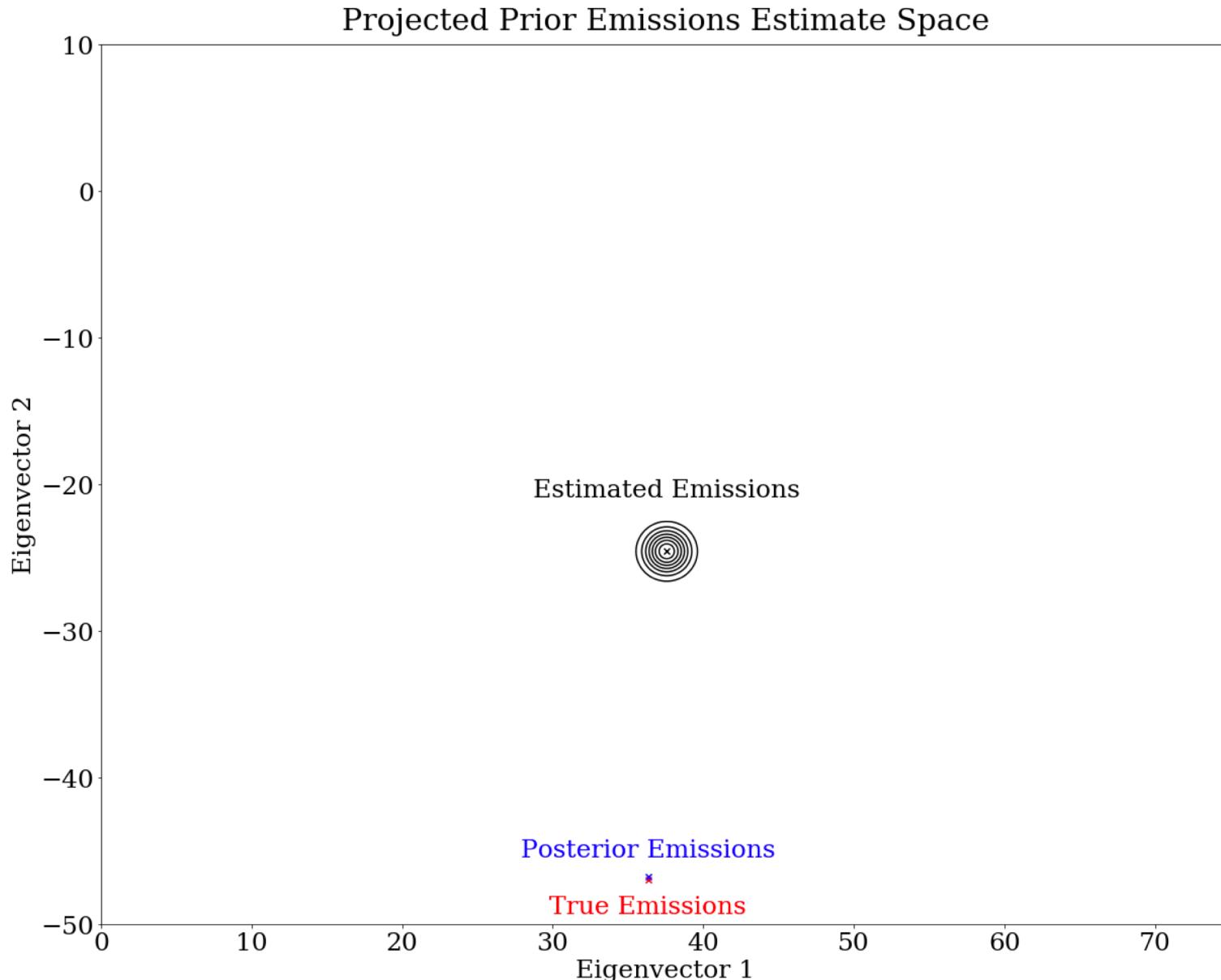
Consider the two dimensional example of constraining methane emissions in the Northern and Southern hemisphere



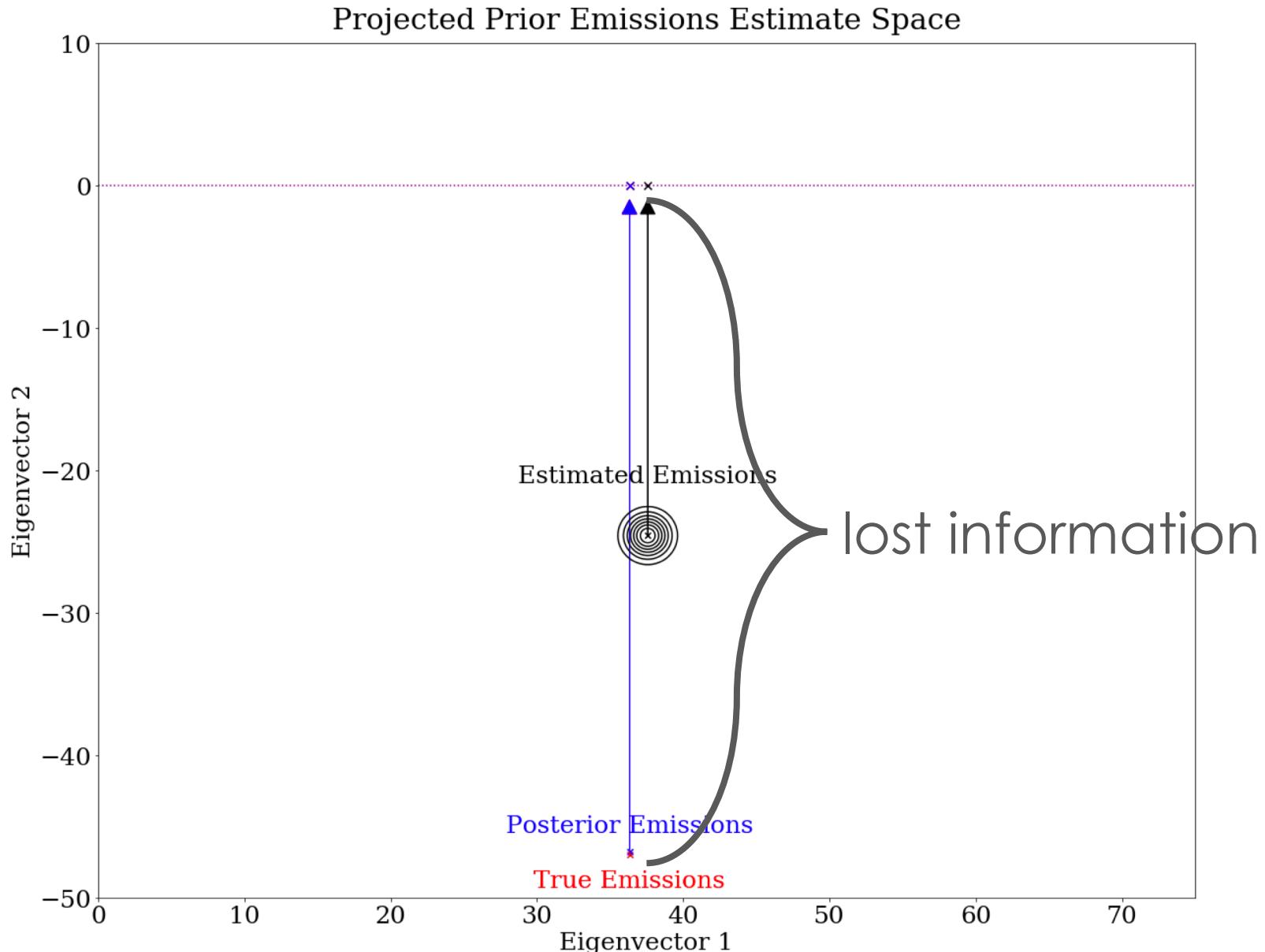
The eigenvectors of H_P form an orthogonal basis of state space;
the eigenvalues represent the significance of each vector



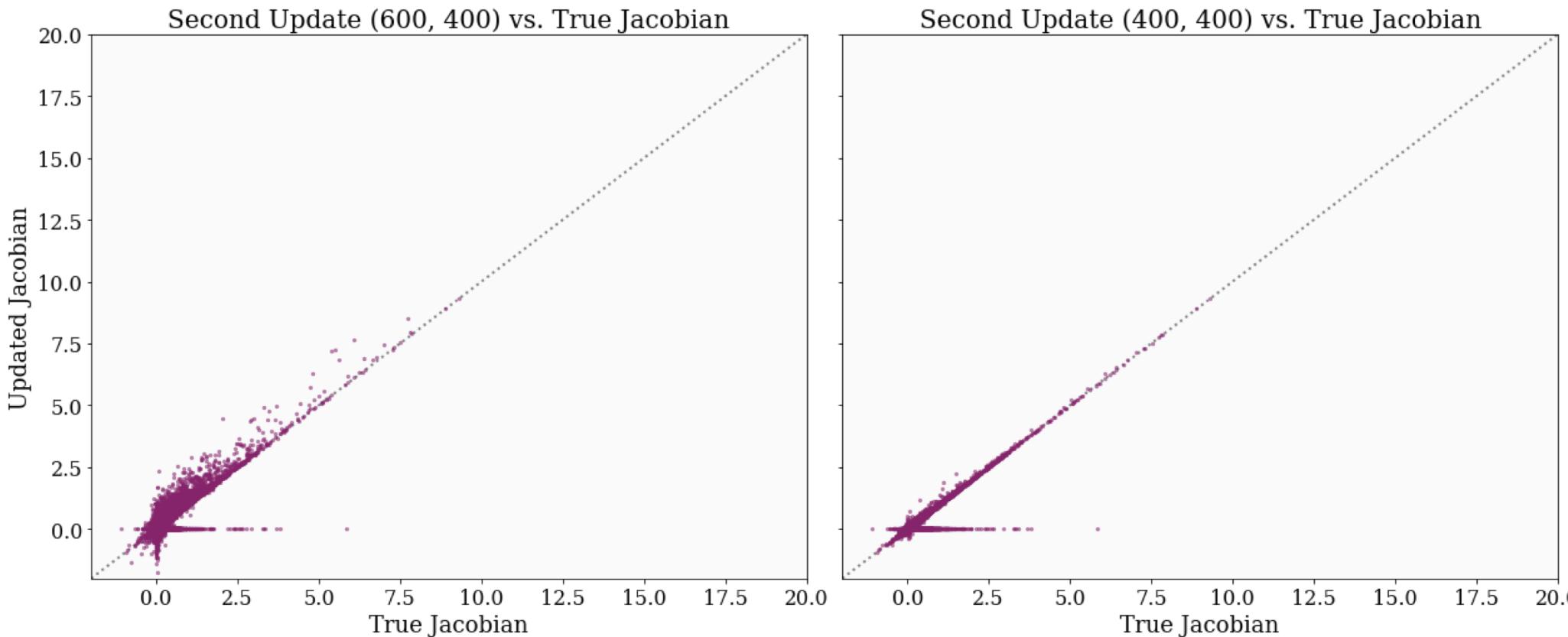
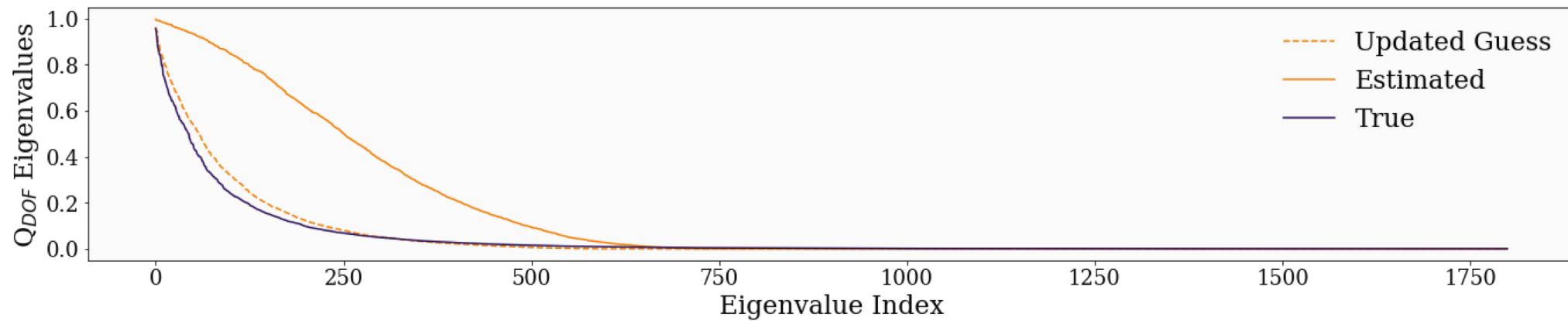
The matrix of eigenvectors V can be used to linearly transform the original state space; this transformation is easily reversed



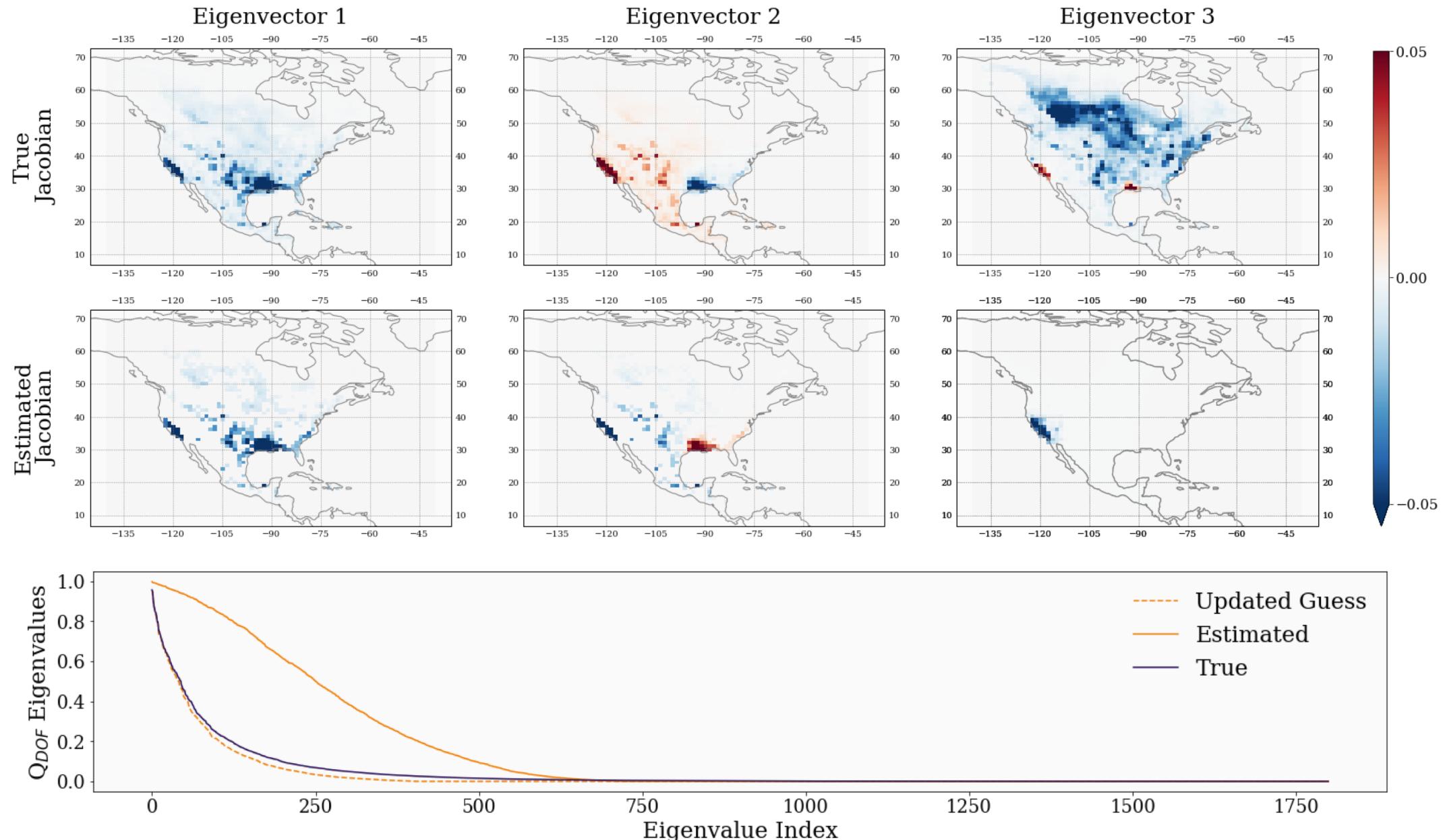
Reducing the number of eigenvectors used reduces the state space dimension



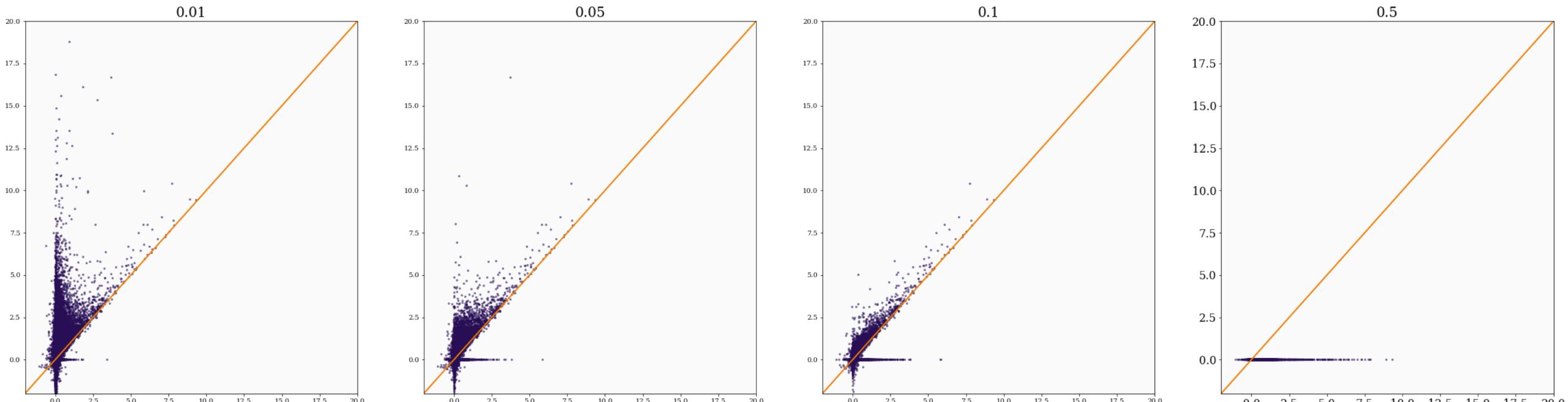
Eigenvalue spectra can define an iteration scheme



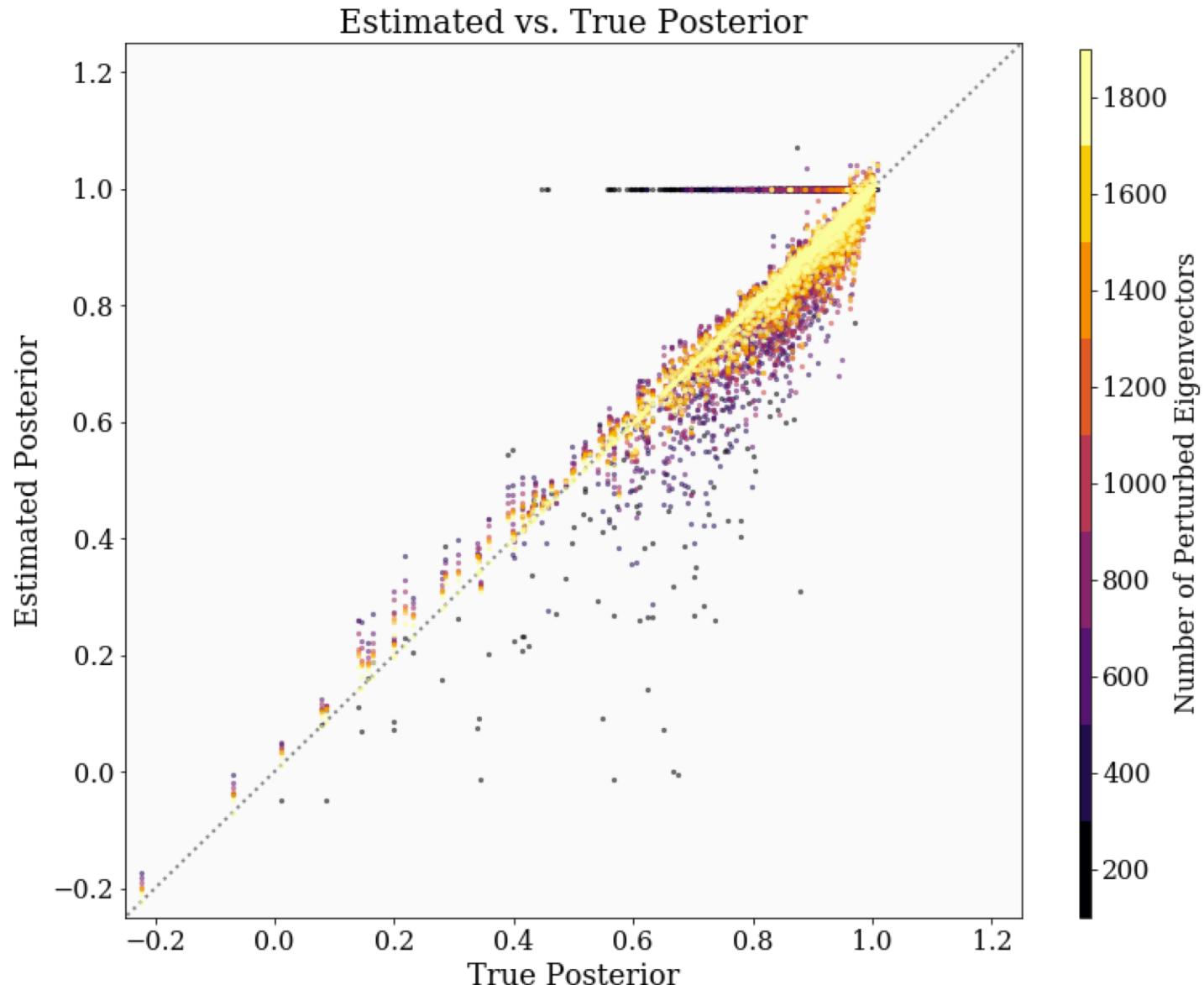
Iterating in this manner improves the eigenpairs



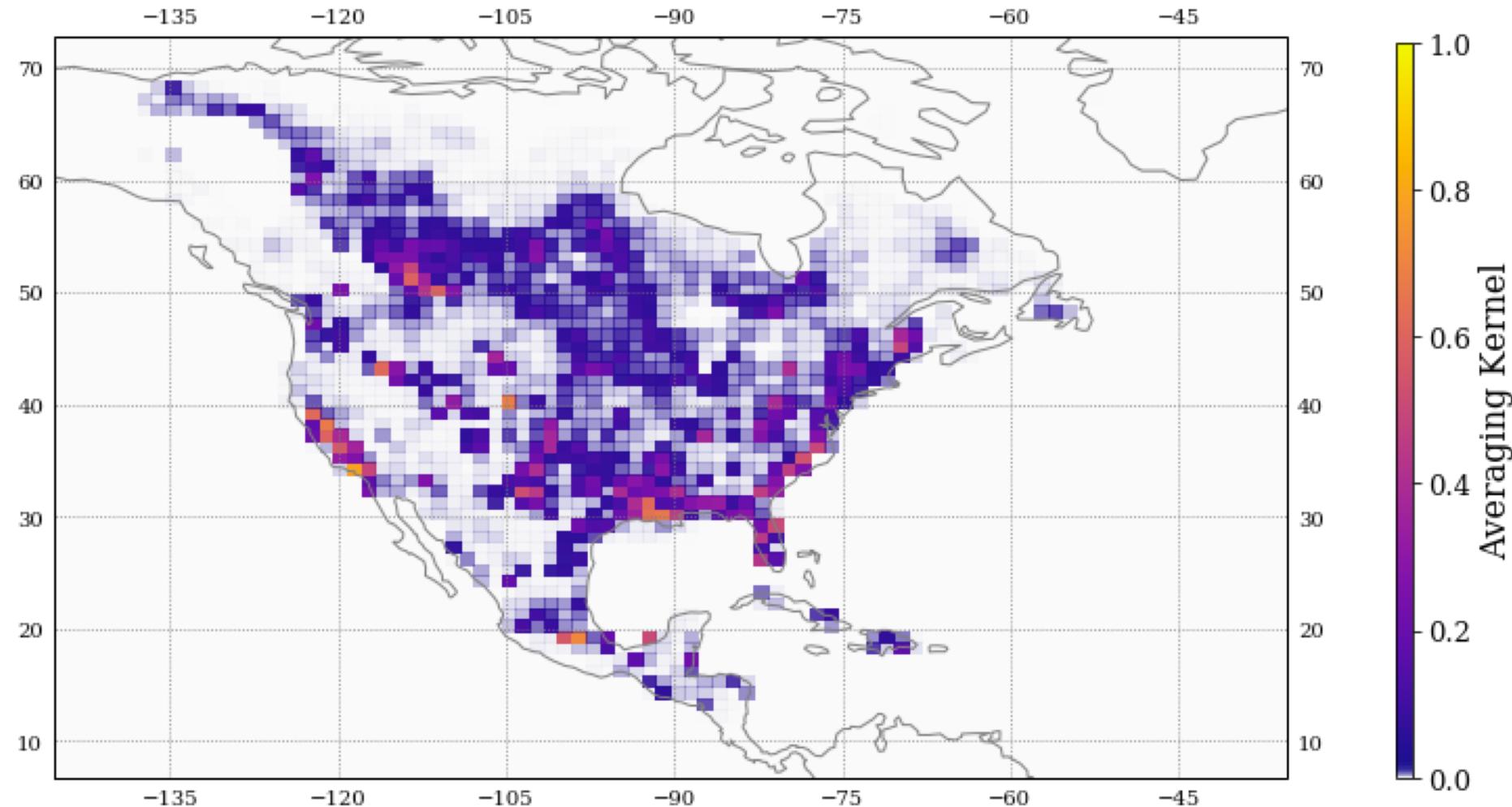
Threshold selection



Reduced-rank Jacobian inversions solve for the posterior in areas with high information content and use the prior elsewhere

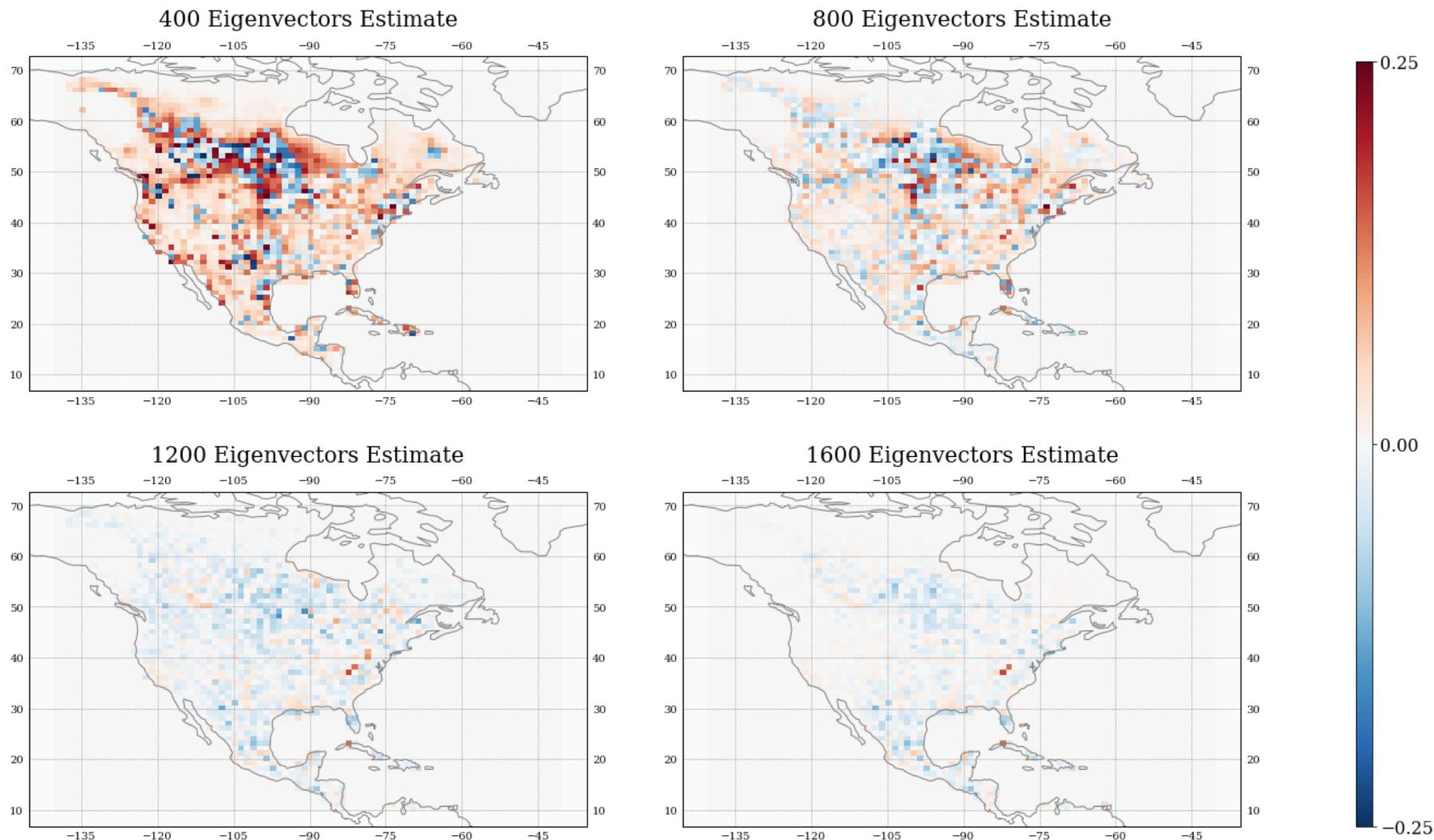


Reduced-rank Jacobian inversions solve for the posterior in areas with high information content and use the prior elsewhere

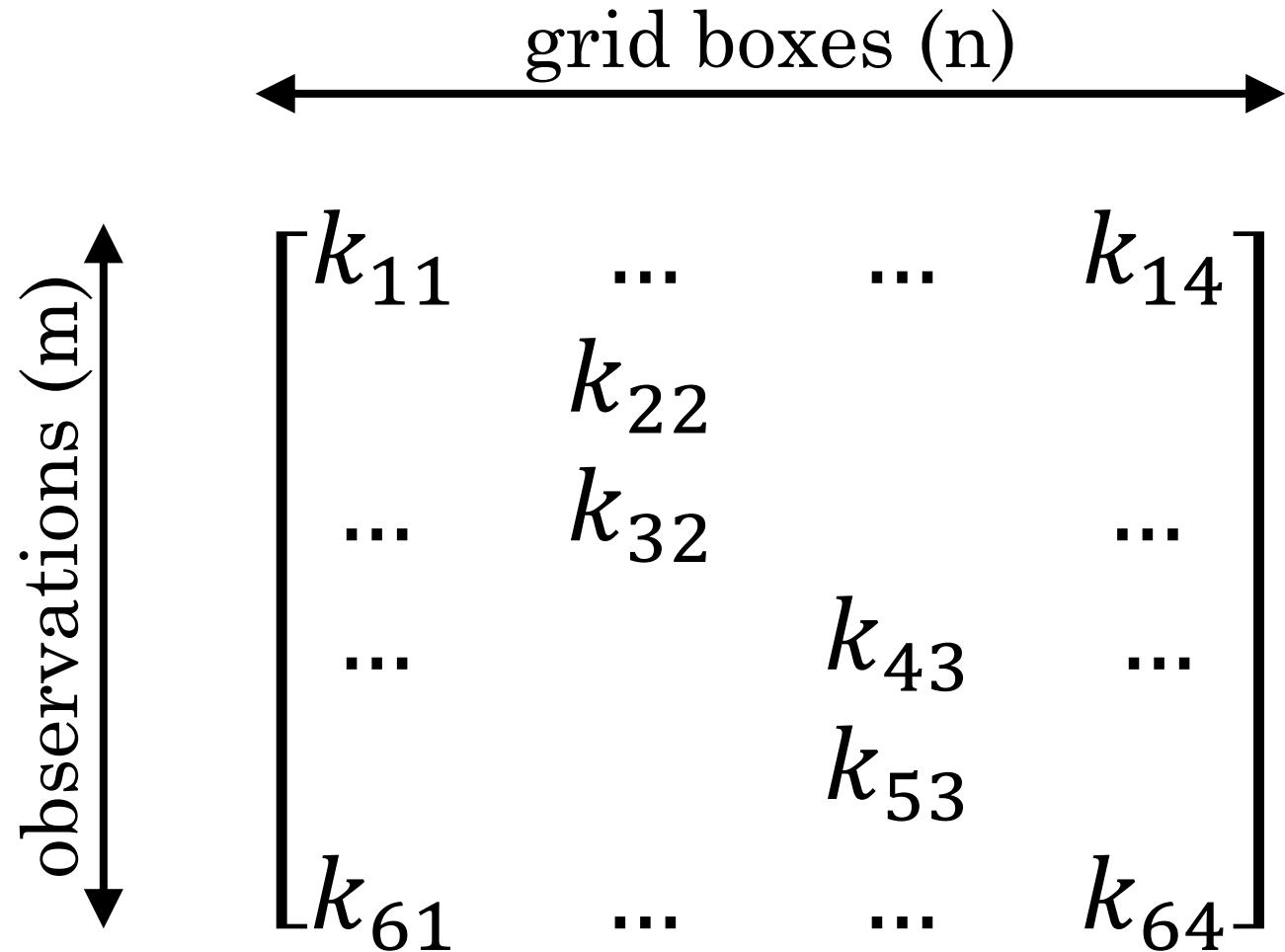


Reduced-rank Jacobian inversions solve for the posterior in areas with high information content and use the prior elsewhere

Estimated Posterior - True Posterior



Even linear forward models are not easily inverted due to (a) non-square matrices ($m \gg n$) and (b) rank deficiencies



Even linear forward models are not easily inverted due to (a) non-square matrices ($m \gg n$) and (b) rank deficiencies

