

# Reducing Jacobian Computational Cost

Hannah Nesser

October 2018

## 1 Introduction

All emission reduction pathways that limit global warming to  $1^\circ\text{C}$  with no or limited overshoot involve deep reductions in methane emissions (insert IPCC citation). Global methane emissions of  $550\text{ Tg a}^{-1}$  are well constrained by inference from mass balance with the global methane sink (Jacob 2016) and the biogenic and anthropogenic sources, including wetlands, livestock, oil and gas systems, landfills, coal mines, and rice paddies, are well known. However, the spatial and temporal contributions of different sources and regions are highly uncertain (cite Kirschke 2013, etc). Satellite observations of total methane columns can constrain bottom-up estimates of methane emissions within the framework of a Bayesian inversion. As satellite technology evolves, inversions are able to constrain methane emissions at increasingly high resolutions. However, as the resolution of an analytic inversion increases, the computational cost increases significantly. Here, we develop a method to decrease the computational cost of analytic Bayesian inversions by applying reduced rank methods to the construction of the Jacobian.

Bayesian inversions calculate the spatially- and temporally-resolved emissions that optimize both prior and observational errors, given a forward model for atmospheric chemistry and transport. When the forward model is linear, there is a closed-form solution for the posterior emissions estimate and its error. The information content of the posterior solution is given by the averaging kernel  $\mathbf{A}$  and, in particular, by the trace of  $\mathbf{A}$ , which gives the Degrees Of Freedom for Signal (DOFS), the number of pieces of information the inversion can constrain.

The computational cost of an analytic inversion is limited by the dimension of the emission (state) vector, which defines the cost of (1) generating the Jacobian  $\mathbf{K}$ , which requires  $n + 1$  forward model runs, and of (2) solving the analytic expression for the posterior error covariance, which requires the inversion of a non-sparse  $n \times n$  matrix. Analytic Bayesian inversions of GOSAT observations frequently constrained emissions on a  $4^\circ \times 5^\circ$  grid, generating a posterior estimate for  $n \approx 1,000$  continental grid cells (i.e. Turner, Maasakkers, Zeng, etc.). TROPOMI aboard the Sentinel-5 Precursor, a satellite launched by the European Space Agency in October 2017, will return methane observations at much higher resolution and density than GOSAT: GOSAT provided approximately 2 million methane column measurements from 2009 to 2015 [cite Maassakkers]; TROPOMI could return this quantity of data in four days [cite Hu]. With this increase in data density, Bayesian inversions will be able to constrain emissions on a much finer grid, increasing the dimension of the state vector from  $n \approx 1,000$  to  $n \approx 4,000$  or  $n \approx 16,000$ .

Past work has focused on decreasing the computational cost of calculating the posterior error covariance. [Insert material on variational inference, ensemble Kalman filters, etc.] Bocquet et al. (2011) defined a method to find a reduced-dimension state vector that maximized the DOFS of the system; however, this method required the construction of many aggregated grids, reducing the computational benefit of the dimension reduction. Turner et al. (2015) also reduced the dimension of the state vector, but constrained emissions on a set of Gaussian radial basis functions rather than an aggregated grid. However, this set of Gaussians did not guarantee optimality. Spantini et al. (2015) defined low-rank approximations that maximized the DOFS of the low-rank system by constructing an optimal, low-rank projection. Bousserez et al. (2018) built upon the framework developed by Bocquet et al. and by Spantini et al. and defined additional, improved approximations and closed-form expressions for the error of the reduced-rank solution relative to the full-rank solution. However, none of these methods addressed the computational cost of constructing the Jacobian  $\mathbf{K}$ .

In what follows, I describe a method to reduce the computational cost of constructing the Jacobian  $\mathbf{K}$  that builds upon the reduced-rank framework developed by Bousserez et al. (2018). In section 2, I discuss the theoretical basis of this method. In section 3, I present the results from the application of this method to a high-resolution inversion of GOSAT observations over North America.

## 2 Reduced-Rank Bayesian Inversions and Jacobians

Bayesian inversions calculate the spatially- and temporally-resolved emissions that optimize both prior and observational errors, given a forward model for atmospheric chemistry and transport. If the forward model is linear, then the Jacobian  $\mathbf{K}$  describes the linear map between the emissions space  $E$  with dimension  $n$  and the observations space  $F$  with dimension  $m$ :  $\mathbf{K} : E \rightarrow F$ . In other words, for  $\mathbf{x} \in E$  and  $\mathbf{y} \in F$ ,  $\mathbf{y} = \mathbf{K}\mathbf{x} + \mathbf{c}$ . When using satellite observations, it is often the case that  $m \gg n$ , producing a problem that is, in the classical sense, overconstrained. However, the number of pieces of information  $n_{DOF}$  that can be constrained by an inversion is often less than the number of state vector elements:  $n_{DOF} \ll n$ ; the problem is in fact underconstrained. In the underconstrained linear system, the rank  $k$  of the Jacobian  $\mathbf{K}$  is less than the upper bound  $n$ . Reduced rank inversion schemes, as introduced by Spantini et al. (2015) and elaborated on by Bousserez et al. (2018), solve analytic Bayesian inversions in a reduced rank space, decreasing the computational cost of inverting a non-sparse  $n \times n$  matrix. However, these approaches do not address the computational cost of constructing the Jacobian  $\mathbf{K}$ .

In what follows, we describe the standard approach to Bayesian inversions (Section 2.1) and the reduced rank approach developed by Spantini et al. (2015) and Bousserez et al. (2018) (Section 2.2). We then introduce the reduced rank approach to Jacobian construction (Section 2.3) and develop error characterization for the approach (Section 2.4).

### 2.1 Bayesian Inversions and Jacobians

A Bayesian inversion assumes that prior and observational errors are Gaussian and minimizes a cost function  $J$ :

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{y} - K(\mathbf{x}))^T \mathbf{S}_O^{-1}(\mathbf{y} - K(\mathbf{x})) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_A)^T \mathbf{S}_A^{-1}(\mathbf{x} - \mathbf{x}_A) \quad (1)$$

where  $\mathbf{x}$  is the state vector,  $\mathbf{x}_A$  is the vector of prior emissions,  $\mathbf{y}$  is the vector of observations,  $K$  is the forward model, and  $\mathbf{S}_A$  and  $\mathbf{S}_O$  are the prior and observational error covariance matrices, respectively. When the forward model is linear,  $K(\mathbf{x}) = \mathbf{K}\mathbf{x} + \mathbf{c}$ , and there is an analytic solution that minimizes  $\mathcal{J}(\mathbf{x})$ , giving the posterior  $\hat{\mathbf{x}}$ . There is also an analytic solution for the posterior error  $\hat{\mathbf{S}}$  and for the information content of the system, given by the averaging kernel  $\mathbf{A}$ . The Bayesian linear problem may be written as  $\{\mathcal{B} : \mathbf{y} | \mathbf{x} \sim \mathcal{N}(\mathbf{K}\mathbf{x}, \mathbf{S}_O); \mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_A); \mathbf{x} \in E, \mathbf{y} \in F\}$ . To simplify notation, we use the formalism of Bousserez et al. (2018):  $\mathcal{B} \equiv (E, F, \mathbf{K}, \mathbf{S}_A, \mathbf{S}_O)$ .

The computational cost of an analytic inversion is dominated by the time it takes to (1) define the linear relationship between the observations and emissions and (2) calculate the posterior error covariance. In both cases, the computational cost is limited by the dimension  $n$  of the state vector  $\mathbf{x}$ . In the former case, an  $n \times n$  matrix must be inverted to calculate the posterior error covariance. In the latter case, the Jacobian is constructed column-wise by running a  $n$  perturbation experiments in the forward model. In a given experiment, one state vector element is perturbed, and the model output is compared to a base run. This gives  $d\mathbf{y}/d\mathbf{x}_i$ , a single column of the Jacobian, where  $i = 1, \dots, n$ . Using the adjoint instead of the Jacobian may decrease the computational cost of the first inversion, but the cost increases with subsequent inversions (i.e. sensitivity tests). Once the Jacobian is built, it may be used for a series of inversions, including sensitivity tests. It is therefore the preferred method.

(Insert transition)

### 2.2 Reduced-Rank Bayesian Inversions

A reduced-rank inversion solves the inverse problem in a subspace of the original state space. Bousserez et al. (2018) defined a rank- $k$  projection operator  $\Pi$  that translates the Bayesian inversion to this subspace.

Like any rank- $k$  linear operator,  $\Pi$  may be written as the product of two rank- $k$  matrices, one which is the left inverse of the other:

$$\Pi = \Gamma^* \Gamma \quad (2)$$

where  $\Gamma^*$  has dimension  $n \times k$  and  $\Gamma$  has dimension  $k \times n$ . The projection may therefore be understood as a sequence of two operations: the first,  $\Gamma$ , reduces the dimension of the state vector from  $n$  to  $k$ , while the second,  $\Gamma^*$ , increases the dimension from  $k$  to  $n$ . I will refer to  $\Gamma$  as the reduction operator and  $\Gamma^*$  as the prolongation operator, keeping with the formalism of Bousserez et al. (2018).

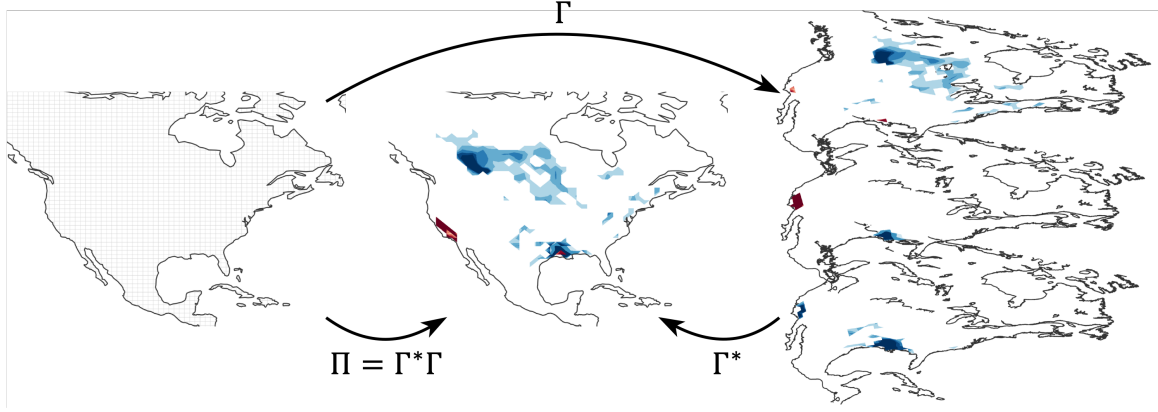


Figure 1: *Note: this chart is adapted from Bocquet et al. (2009). I will modify the chart to more accurately reflect our scheme (i.e. representing non-discrete elements of the subspace  $w$ ) at a later date.*

Figure 1 shows the relationship between the original emission-observation space, the projected subspace, and the reduced dimension subspace. The original space of the problem is depicted by the leftmost panel. The rank  $k$  projection  $\Pi$  transforms the original space to a rank  $k$  subspace, depicted by the center panel. The reduction operator  $\Gamma$  transforms the reduced rank subspace to a reduced dimension subspace, depicted by the right panel. The optimal prolongation operator  $\Gamma^*$  returns the reduced dimension space to the reduced rank subspace. However, there is no operator defining a return to the original space. Qualitatively, reducing the rank of the original space involves a loss of information that cannot be restored. Quantitatively, the projector  $\Pi$  is of dimension  $n \times n$  but rank  $k$ , meaning that it is non-invertible and the transformation cannot be precisely reversed.

The projected Bayesian problem corresponds to the center panel of Figure 1. Its posterior solution corresponds to the posterior solution in the original space projected to the reduced rank subspace. Specifically,

$$\hat{\mathbf{x}}_{\Pi_{\text{opt}}} = \Pi \hat{\mathbf{x}} \quad (3)$$

$$\hat{\mathbf{S}}_{\Pi_{\text{opt}}} = \Pi \hat{\mathbf{S}} \Pi^T \quad (4)$$

$$\mathbf{A}_{\Pi_{\text{opt}}} = \Pi \mathbf{A} \Pi \quad (5)$$

The rank  $k$  projection  $\Pi$  necessarily discards information. However, a projection can be found that optimizes some criteria, such as minimal information content loss. Of particular interest is the minimization of the representativeness error, which is the error introduced if the subspace of the observations is not fully included in the range of the projector; it is the difference between the full-rank observational error and the observational error of the reduced-rank system. Bousserez et al. (2018) describe a general form of a projection that minimizes representativeness error; in particular, they show that for any reduction operator  $\Gamma$  there exists a prolongation operator  $\Gamma^*$  that minimizes representativeness error:

$$\Gamma_{\text{opt}}^* = \mathbf{S}_A \Gamma^T (\Gamma \mathbf{S}_A \Gamma^T)^{-1} \quad (6)$$

Within this class of projections, additional optimality criteria, such as maximizing the DOFS or minimizing the posterior variance, can be introduced. Bousserez et al. (2018) find that it is possible to create

representativeness-error minimizing projection that also maximizes the DOFS of the reduced-rank inversion. This projection is given by:

$$\mathbf{\Pi}_{\text{DOF}} = \mathbf{S}_A^{1/2} \mathbf{W}_k \mathbf{W}_k^T \mathbf{S}_A^{-1/2} \quad (7)$$

$\mathbf{W}_k$  is the matrix whose columns are the first  $k$  eigenvectors of the prior-error-weighted averaging kernel  $\mathbf{Q}_{\text{DOF}} \equiv \mathbf{S}_A^{-1/2} \mathbf{A} \mathbf{S}_A^{1/2}$  with eigendecomposition  $\mathbf{Q}_{\text{DOF}} = \mathbf{W}^T \mathbf{\Sigma} \mathbf{W}$ , where  $\mathbf{W}$  is the matrix whose columns are the eigenvectors of  $\mathbf{Q}_{\text{DOF}}$  and  $\mathbf{\Sigma}$  is a diagonal matrix whose elements are the eigenvalues of  $\mathbf{\Sigma}$ . However, this solution requires the inversion of a non-sparse  $n \times n$  matrix, negating any benefit to the computational cost. Bousseret et al. (2018) find that the eigenpairs of  $\mathbf{Q}_{\text{DOF}}$  are linearly related to those of another matrix, the prior-preconditioned Hessian  $\mathbf{H}_p \equiv \mathbf{S}_A^{1/2} \mathbf{K}^T \mathbf{S}_O^{-1} \mathbf{K} \mathbf{S}_A^{1/2}$  with eigendecomposition  $\mathbf{H}_p = \mathbf{V}^T \mathbf{\Lambda} \mathbf{V}$ . The eigenvectors and eigenvalues of the prior-preconditioned Hessian are related to those of  $\mathbf{Q}_{\text{DOF}}$  by the following linear relationships:

$$\mathbf{W} = \mathbf{V} \quad (8)$$

$$\mathbf{\Sigma} = \mathbf{\Lambda} (\mathbf{I} + \mathbf{\Lambda})^{-1} \quad (9)$$

Given the eigendecomposition of the prior-preconditioned Hessian, the optimal reduction and prolongation operators are:

$$\Gamma = \mathbf{V}_k^T \mathbf{S}_A^{-1/2} \quad (10)$$

$$\Gamma^* = \mathbf{S}_A^{1/2} \mathbf{V}_k \quad (11)$$

(Personal note: This is a correction of the result presented by Bousseret et al. (2018).)

It is then possible to write the posterior solution of a DOFS maximizing rank  $k$  projection:

$$\hat{\mathbf{x}}_{\mathbf{\Pi}_{\text{DOF}}} = \mathbf{S}_A^{1/2} \left( \sum_{i=1}^k \lambda_i^{1/2} (1 + \lambda_i)^{-1} \mathbf{v}_i \mathbf{w}_i^T \right) \mathbf{S}_O^{-1/2} (\mathbf{y} - (\mathbf{K} \mathbf{x} + \mathbf{c})) \quad (12)$$

$$\hat{\mathbf{S}}_{\mathbf{\Pi}_{\text{DOF}}} = \mathbf{S}_A^{1/2} \left( \sum_{i=1}^k (1 + \lambda_i)^{-1} \mathbf{v}_i \mathbf{v}_i^T \right) \mathbf{S}_A^{1/2} \quad (13)$$

$$\mathbf{A}_{\mathbf{\Pi}_{\text{DOF}}} = \mathbf{S}_A^{1/2} \left( \sum_{i=1}^k \lambda_i (1 + \lambda_i)^{-1} \mathbf{v}_i \mathbf{v}_i^T \right) \mathbf{S}_A^{-1/2} \quad (14)$$

If all eigenvectors and eigenvalues of the prior-preconditioned Hessian are explicitly calculated (and not approximated using random SVD), the posterior mean may be written:

$$\hat{\mathbf{x}}_{\mathbf{\Pi}_{\text{DOF}}} = \hat{\mathbf{S}}_{\mathbf{\Pi}_{\text{DOF}}} \mathbf{K}^T \mathbf{S}_O^{-1} (\mathbf{y} - (\mathbf{K} \mathbf{x} + \mathbf{c})) \quad (15)$$

$$(16)$$

These are solutions to the projected problem  $\mathcal{B}_{\mathbf{\Pi}_{\text{DOF}}} \equiv (E_{\mathbf{\Pi}_{\text{DOF}}}, F, \mathbf{K}, \mathbf{S}_{\mathbf{A}_{\mathbf{\Pi}_{\text{DOF}}}}, \mathbf{S}_{\mathbf{O}_{\mathbf{\Pi}_{\text{DOF}}}})$ . It is also possible to define the problem  $\mathcal{B}_{\mathbf{K}_{\mathbf{\Pi}_{\text{DOF}}}} \equiv (E, F, \mathbf{K} \mathbf{\Pi}_{\text{DOF}}, \mathbf{S}_A, \mathbf{S}_{\mathbf{O}_{\mathbf{\Pi}_{\text{DOF}}}})$ , which corresponds to the initial inverse problem with a projected forward model. The posterior solutions of that problem are:

$$\hat{\mathbf{x}}_{\mathbf{K}_{\mathbf{\Pi}_{\text{DOF}}}} = \hat{\mathbf{x}}_{\mathbf{\Pi}_{\text{DOF}}} \quad (17)$$

$$\hat{\mathbf{S}}_{\mathbf{K}_{\mathbf{\Pi}_{\text{DOF}}}} = \hat{\mathbf{S}}_{\mathbf{\Pi}_{\text{DOF}}} + \mathbf{S}_A^{1/2} \left( \sum_{i=k+1}^n \mathbf{v}_i \mathbf{v}_i^T \right) \mathbf{S}_A^{1/2} \quad (18)$$

$$= \mathbf{S}_A - \mathbf{S}_A^{1/2} \left( \sum_{i=1}^k \lambda_i (1 + \lambda_i)^{-1} \mathbf{v}_i \mathbf{v}_i^T \right) \mathbf{S}_A^{1/2} \quad (19)$$

$$\mathbf{A}_{\mathbf{K}_{\mathbf{\Pi}_{\text{DOF}}}} = \mathbf{A}_{\mathbf{\Pi}_{\text{DOF}}} \quad (20)$$

The posterior error covariance associated with the projected forward model is similar to that of the maximum DOFS projection, but corresponds to "a forward model for which the modes are filtered by the projection  $\mathbf{\Pi}_{\text{DOF}}$ " (Bousseret et al. (2018)). (Edit. Also note which set of solutions is preferable.)

(Insert transition)

## 2.3 Reduced-Rank Jacobian Construction

In a standard inversion, the Jacobian is constructed column-wise by conducting  $n + 1$  model runs, where in model run  $i + 1$ , the  $i$ th grid cell is perturbed. In a reduced-rank inversion, the elements of the reduced-dimension space  $\omega$  with dimension  $k \ll n$  may be perturbed instead. However, defining the optimal projection and its reduction and prolongation operators requires prior knowledge of the Jacobian (insert equation reference). Therefore, it is necessary to iteratively update the reduced-rank Jacobian estimate. In what follows, we assume  $\mathbf{\Pi} = \mathbf{\Pi}_{\text{DOF}}$ .

The general iteration scheme is as follows: *Note: here I enumerate the general approach, considering all possible iteration routes. I will go into more detail on each of the routes, and on their sticking points, later.)*

1. Initialize the Jacobian  $\mathbf{K}_0$ .
2. Solve for the associated rank- $k$  optimal projection  $\mathbf{\Pi}_0 = \Gamma_0^* \Gamma_0$ .
3. Perturb some  $k$  elements (either eigenvectors or grid cells) and calculate the associated Jacobian. The Jacobian may be in (a) the projected subspace (i.e. the subspace defined by  $\mathbf{\Pi}$ ), (b) the reduced dimension space (i.e.  $\omega$ , the subspace defined by  $\Gamma$ ), or (c) the original space.
4. Return to item 2 and iterate until convergence.
5. Solve the inversion. If the final Jacobian is in the projected subspace or the reduced dimension space, the solution will be within that space.

The Jacobian may be initialized by, for instance, generating a Jacobian at coarser resolution and regridding. Changes in modeled observations at coarse resolution may be allocated to higher resolution grid cells according to the fraction of emissions the smaller grid cell is responsible for relative to the larger grid cell, with the fraction given by the prior estimate. For a coarse grid cell with emissions  $x$  with higher resolution constituents  $x_1$ ,  $x_2$ ,  $x_3$ , and  $x_4$  and an observation vector  $\mathbf{y}$  that we wish to decompose into its higher resolution constituents  $\mathbf{y}_1$ ,  $\mathbf{y}_2$ ,  $\mathbf{y}_3$ , and  $\mathbf{y}_4$ :

$$\frac{\Delta \mathbf{y}}{\Delta x} = \frac{\Delta \mathbf{y}_1 + \Delta \mathbf{y}_2 + \Delta \mathbf{y}_3 + \Delta \mathbf{y}_4}{\Delta x_1 + \Delta x_2 + \Delta x_3 + \Delta x_4}$$

$$\Delta \mathbf{y}_1 + \Delta \mathbf{y}_2 + \Delta \mathbf{y}_3 + \Delta \mathbf{y}_4 = f_1 \Delta \mathbf{y} + f_2 \Delta \mathbf{y} + f_3 \Delta \mathbf{y} + f_4 \Delta \mathbf{y}$$

We then estimate  $\Delta \mathbf{y}_i \approx f_i \Delta \mathbf{y}$  for  $i \in \{1, 2, 3, 4\}$ , where  $f_i = \Delta x_i / \Delta x$ , the fraction of the coarse grid cell emissions attributable to the smaller cell. The approximate changes in observations at fine resolution are then used to calculate an initial estimate of the Jacobian  $\mathbf{K}_0$ .

However, building a coarse resolution Jacobian reduces the computational benefit of the reduced rank approach since it requires at maximum  $n/4$  and down to  $n/8$  or  $n/16$  model runs. Instead, we may assume that the fine structure of the Jacobian is given by the prior emissions estimate and use that vector to construct a sparse Jacobian that we iterate on. (*This is a work in progress; more elaboration will come in time.*)

Once the Jacobian is initialized, the optimal prolongation and reduction operators and the resulting projection may be calculated by equations (7), (10), and (11). Given these operators, it is possible to calculate an improved Jacobian estimate in any of three spaces: the reduced dimension subspace, the reduced rank subspace, or the original, full rank state space. Consider each in turn. (*Note: The consideration of each of these is mainly an attempt to explore each avenue to see what possibilities exist.*)

**Reduced dimension perturbation:** A perturbation of the reduced-dimension prior  $\mathbf{x}_{Aw}$  is given

by

$$\Delta \mathbf{x}_{A\omega_i} = \begin{bmatrix} 0 & & & & 0 \\ & \ddots & & & \\ & & 0.5 & & \\ & & & \ddots & \\ 0 & & & & 0 \end{bmatrix} \mathbf{x}_{A\omega} \quad (21)$$

$$= \mathbf{P}_i \mathbf{x}_{A\omega} \quad (22)$$

$$= \mathbf{P}_i \Gamma \mathbf{x}_A \quad (23)$$

where  $\mathbf{P}$  represents a  $k \times k$  matrix with  $P_{jj} = 0.5$  for  $j = i$ ,  $P_{jj} = 0$  otherwise.

To calculate the Jacobian, the model response to the perturbation is calculated. However, most atmospheric transport and chemistry models do not support flexible, non-discrete grids. To transform the reduced dimension perturbation to the standard grid (i.e. the full dimension space), consider a model, denoted by  $\mathbf{K}_{true}$ , and its hypothetical reduced dimension equivalent, denoted by  $\mathbf{K}_{true,\omega} = \mathbf{K}_{true} \Gamma^*$ . In this system, the observed response to a perturbation of the  $i$ th element of the reduced dimension space  $\omega$  is given by

$$\begin{aligned} \Delta \mathbf{y}_i &= \mathbf{K}_{true,\omega} \Delta \mathbf{x}_{A\omega_i} \\ &= \mathbf{K}_{true} \Gamma^* \mathbf{P}_i \Gamma \mathbf{x}_A \\ &= 0.5 \mathbf{K}_{true} \Gamma_{:,i}^* \Gamma_{i,:} \mathbf{x}_A \\ &= 0.5 \mathbf{K}_{true} \mathbf{S}_A^{-1/2} v_i v_i^T \mathbf{S}_A^{1/2} \mathbf{x}_A \end{aligned}$$

where  $\Gamma_{:,i}^*$  represents the  $i$ th column of  $\Gamma^*$ ,  $\Gamma_{i,:}$  the  $i$ th row of  $\Gamma$ , and  $v_i$  the  $i$ th eigenvector of the prior-preconditioned Hessian. Notice that the term  $\mathbf{S}_A^{-1/2} v_i v_i^T \mathbf{S}_A^{1/2}$  incrementally constructs the maximum DOFS projection; that is,  $\mathbf{\Pi} = \sum_{i=1}^k \mathbf{S}_A^{-1/2} v_i v_i^T \mathbf{S}_A^{1/2}$ . The sum of the model responses is therefore proportional to the modeled observations given a reduced rank state space:  $\sum_{i=1}^k \Delta \mathbf{y}_i = 0.5 \mathbf{K}_{true} \mathbf{\Pi} \mathbf{x}_A$ .

Given a perturbation  $\Delta \mathbf{x}_{A\omega}$  and its model response  $\Delta \mathbf{y}_i$ , the reduced dimension Jacobian can be constructed columnwise with the  $i$ th column  $\mathbf{K}_{\omega,:i}$ :

$$\mathbf{K}_{\omega,:i} = \frac{\Delta \mathbf{y}_i}{\|\Delta \mathbf{x}_{A\omega}\|}$$

**Reduced rank perturbation:** A perturbation of the reduced-rank prior  $\mathbf{x}_{A\Pi}$  is given by:

$$\Delta \mathbf{x}_{A\Pi_i} = \begin{bmatrix} 0 & & & & 0 \\ & \ddots & & & \\ & & 0.5 & & \\ & & & \ddots & \\ 0 & & & & 0 \end{bmatrix} \mathbf{x}_{A\Pi} \quad (24)$$

$$= \mathbf{P}_i \mathbf{x}_{A\Pi} \quad (25)$$

$$= \mathbf{P}_i \mathbf{\Pi} \mathbf{x}_A \quad (26)$$

To calculate the model response to a perturbation in this reduced rank space, consider the reduced rank equivalent of the true Jacobian, denoted by  $\mathbf{K}_{true,\Pi} = \mathbf{K}_{true} \mathbf{\Pi}$ . In this system, the observed response to a perturbation of the  $i$ th element of the reduced rank space  $\Pi$  is given by

$$\begin{aligned} \Delta \mathbf{y}_i &= \mathbf{K}_{true,\Pi} \Delta \mathbf{x}_{A\Pi_i} \\ &= \mathbf{K}_{true} \mathbf{\Pi} \mathbf{P}_i \mathbf{\Pi} \mathbf{x}_A \\ &= 0.5 \mathbf{K}_{true} \mathbf{\Pi}_{:,i} \mathbf{\Pi}_{i,:} \mathbf{x}_A \end{aligned} \quad =$$

Given a perturbation  $\Delta \mathbf{x}_{A\Pi}$  and its model response  $\Delta \mathbf{y}_i$ , the reduced rank Jacobian can be constructed columnwise with the  $i$ th column  $\mathbf{K}_{\Pi,:i}$ :

$$\mathbf{K}_{\Pi,:i} = \frac{\Delta \mathbf{y}_i}{\|\Delta \mathbf{x}_{A\Pi}\|}$$

**Full rank perturbation:** A perturbation in the full rank space consists of perturbing individual grid cells rather than eigenvectors. The eigenvectors inform the grid cells that are perturbed; since the eigenvectors form a basis of the most informed subspace, perturbing the grid cells that contribute the most to those eigenvectors may maximize the information content of the Jacobian. Given an initial estimate of the Jacobian, it is possible to calculate an initial estimate of the degrees of freedom for signal, the number of pieces of information that may be constrained by the inversion. Since the eigenvectors give the most informed basis of the inverse system, we use the first DOFS eigenvectors to define the grid cells associated with maximum grid cells. The eigenvectors are normalized by default, so we take the sum of the absolute values of the first DOFS eigenvectors for each grid cell and apply some threshold criteria to eliminate those grid cells with the lowest absolute sum. (*This approach has been attempted; however, we are not pursuing it at the moment.*)

## 2.4 Error Characterization and Convergence Criteria

### 3 Application to a High-Resolution North American Jacobian

To demonstrate the method described above, we construct a Jacobian over the North American domain for a month of observations (insert month) at  $4^\circ \times 5^\circ$  and at  $2^\circ \times 2.5^\circ$ . The Jacobian is constructed using the nested GEOS-Chem methane simulation at  $0.5^\circ \times 0.625^\circ$  and perturbing the 16 grid boxes equivalent to a single  $2^\circ \times 2.5^\circ$  grid box for each of the  $2^\circ \times 2.5^\circ$  grid boxes in the North American domain. The simulated observations can then be aggregated to calculate both Jacobians.

Once constructed, the  $4^\circ \times 5^\circ$  Jacobian is regridded to  $2^\circ \times 2.5^\circ$  using (blank). This is used as the initial