**Reduced Cost Construction of Jacobian Matrices for High-Resolution Inverse Modeling: An Application to Optimizing North American Methane Sources from GOSAT Satellite Data**

**Authors**

Hannah Nesser1, Daniel J. Jacob1, Joannes D. Maasakkers2, Melissa P. Sulprizio1, Yuzhong Zhang1, Tia Scarpelli1

**Affiliations**

**1** Harvard University, Cambridge, Massachusetts, USA.

2 SRON Netherlands Institute for Space Research, Utrecht, the Netherlands.

**Abstract**

Global high-resolution observations of atmospheric trace gas concentrations from satellites can greatly improve our understanding of surface emissions through inverse analyses. For example, the new Tropospheric Monitoring Instrument (TROPOMI) retrieves daily global observations of atmospheric methane concentrations at 7x7 km2 pixel resolution. Variational inverse methods can optimize surface emissions globally at this resolution but do not readily provide error characterization, including information content, for the posterior solution. In fact, the information content of the satellite data may be considerably lower than the data density would suggest because of limited retrieval success rate, instrument noise, and error correlations that propagate through the inversion. This could lead to smoothing errors in variational methods. An analytic inverse solution provides closed-form characterization of the posterior error statistics and information content but requires the construction of the Jacobian matrix relating emissions to atmospheric concentrations. Building the Jacobian matrix is computationally expensive at fine resolution because it involves perturbing each emission element, typically individual grid cells, in the atmospheric transport model. We propose a method to greatly decrease the computational cost of analytic inversions by constructing the Jacobian matrix using only the emission elements with sufficient information content from the observations. Starting from an initial estimate of the Jacobian matrix that assumes simple transport, we iteratively apply perturbations to the leading patterns of information content rather than to the individual model grid cells. The resulting matrix optimizes emissions only in areas spanned by these leading patterns. We demonstrate the method in an analytic Bayesian inversion of GOSAT data over North America in July 2009. We confirm that the estimated Jacobian matrix produces posterior emission estimates and error covariances that are similar to an inversion conducted with the Jacobian matrix for the original model grid. Our method enables computationally efficient, high-resolution analytic inversions of high-density satellite data.

**Section 1: Introduction**

Satellite observations of atmospheric composition provide an increasingly powerful resource to improve our understanding of emissions (Streets et al. 2013; Jacob et al. 2016). Satellites provide dense observations at a global scale but are subject to large errors, both from the measurement retrieval and from the chemical transport models (CTMs) used to relate the observed atmospheric concentrations to emissions (Jacob et al. 2016). Exploiting satellite data to reliably quantify emissions and the associated errors at high resolution is of considerable interest. However, these calculations are computationally expensive and may be limited by the information content of the observations. Here we present a method to achieve high-resolution inversions of satellite data to infer emissions with full error and information content characterization at reduced computational cost, limited by the information content of the observations and by the availability of computational resources.

A standard inversion infers emissions from observations of atmospheric composition by fitting the data to a chemical transport model (CTM) that simulates atmospheric concentrations as a function of emissions. The CTM represents the forward model for the inverse problem. Inversions are often solved by Bayesian optimization, minimizing a cost function regularized by prior information on the emissions. The optimal (posterior) emissions estimate corresponds to the cost function minimum. The minimum is typically found using a numerical (variational) method, often employing the adjoint of the CTM to compute the cost function gradient (Henze et al. 2007). However, the numerical solution provides no explicit characterization of the solution’s error or information content. While ensemble approaches, sensitivity tests, or other numerical approahces can approximate error, these methods are computationally expensive and often fail to fully capture the error of the inverse system (Evensen 2009; Bousserez and Henze 2018).

In the common case where the observed atmospheric concentrations have a linear dependence on emissions and error statistics can be assumed normal, an analytic solution to the Bayesian optimization problem exists that provides closed-form expressions for the posterior emissions estimate, its error statistics, and its information content (Rodgers 2000). This approach can be extended to systems with log-normal error statistics and, through an iterative approach, to non-linear problems (Maasakkers et al. 2019; Rodgers 2000). The analytic solution requires explicit construction of the Jacobian matrix of the forward model, , which represents the sensitivity of the simulated concentrations to the optimized emission state vector (Brasseur and Jacob 2017). The Jacobian is constructed column-wise by conducting CTM simulations that perturb each of the state vector elements to find the corresponding column (Maasakkers et al. 2019). Even on massively parallel computing clusters, the computational cost of conducting these *n* simulations can limit the size of the state vector **x** and therefore the resolution at which inversions can be conducted (Turner and Jacob 2015). However, once the Jacobian matrix is constructed, inversions can be conducted at essentially no additional computational cost to investigate the sensitivity of the solution to changes in the specification of errors, prior assumptions, or the number and type of observations included.

Methane illustrates the problem. Methane is an important greenhouse gas but large uncertainties exist in its emissions (Saunois et al. 2019). Inversions of satellite observations of atmospheric methane concentrations can improve emission estimates (Jacob et al. 2016). This approach was first implemented using data from the SCIAMACHY satellite instrument, which provided high-quality observations of column methane concentrations at coarse spatial resolution (30 x 60 km2 nadir pixel resolution) for 2005 – 2007 (Bergamaschi et al. 2009, 2013; Houweling et al. 2014; Wecht et al. 2014). The longest satellite record of methane concentrations (2009 – present) is provided by GOSAT, which measures column methane concentrations with 10-km diameter pixels approximately 250 km apart along- and cross-track (Kuze et al. 2009). Maasakkers et al. (2019) conducted an analytic inversion of GOSAT data from 2010 – 2015 to constrain emissions and trends in 4º x 5º terrestrial grid boxes, commensurate with the information content of the observations. This inversion optimized ~2,000 state vector elements. The launch of the Tropospheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 precursor in October 2017 changes the magnitude of the computation problem. TROPOMI now provides daily, global retrievals of atmospheric methane columns at 7 x 7 km2 nadir pixel resolution, increasing coverage by orders of magnitude compared to GOSAT (Veefkind et al. 2012). However, TROPOMI’s methane retrieval has only a ~3% success rate limited by clouds, high aerosol loadings, and variable surface albedo and topography, resulting in heterogeneously distributed observations (Hu et al. 2018). While analytic solution of an inversion of TROPOMI observations at high resolution is computationally prohibitive, characterizing the error and information content of the posterior solution is necessary given the spatial variability of the retrievals. This observational heterogeneity suggests, however, that one could define a state vector with varying spatial resolution, enabling analytic solution of the inverse problem at decreased computational cost while preserving most of the information content of the observations.

Several methods have been developed to decrease the spatial resolution of the emission state vector while preserving the information content of the inverse system. Bocquet et al. (2011) developed a method to find an optimal multiscale grid for the state vector for a given inverse system. This approach is computationally expensive because it searches through all allowed multiscale grids. Turner and Jacob (2015) reduced the dimension of an analytic inversion using a *k*-member Gaussian mixture model with *k* < *n*, but the Gaussian groupings of the high-resolution grid cells were defined by subjective similarity criteria. Spantini et al. (2015) and Bousserez and Henze (2018) presented low-rank methods to identify the dominant patterns of information in an inverse system to decrease the computational cost of solving for the posterior emissions. However, their approach requires either explicit construction of the Jacobian matrix (Spantini et al. 2015) or expectation that the information content is described by a relatively small number of patterns, as is the case for severely underconstrained inverse systems (Bousserez and Henze 2018).

Here we present two methods to construct the Jacobian matrix for a native *n*-dimensional state vector that preserves the information content of the inverse system using *k* < *n* forward model simulations. These methods allow for analytic inversion of satellite data with limited information content at high resolution without pre-judging information content. The first method uses an initial estimate of the Jacobian matrix to generate a multiscale grid that preserves resolution in the areas with highest information content. The second system populates an initial estimate of the native-resolution Jacobian matrix along the dominant patterns of information content in the system. In both cases, the number of model simulations *k* used to construct the Jacobian is elected by the user and represents a trade-off between computational tractability and information content. We apply both methods to a 1-month inversion of satellite data as demonstration.

**Section 2: Methods**

*Section 2.1: Analytic Solution to the Inverse Problem*

The optimal estimate of a state vector **x** given a prior estimate **xA**, observation vector **y**, and normal error statistics given by error covariance matrices **SA** and **SO**, respectively, is given by the minimization of a Bayesian scalar cost function **J**(**x**):

Here **F**(**x)** represents forward model; i.e. the simulated observations **y** = **F**(**x**) from the CTM. If the forward model is linear so that **F**(**x**) = **Kx** + **c**, where **K** = d**y**/d**x** is the Jacobian matrix and **c** is a constant, then an analytic solution to the cost function minimum exists that yields the posterior emissions estimate and its error covariance matrix :

Comparison of and defines the information content of the inverse system, quantified by the averaging kernel matrix **A**, which represents the sensitivity of the posterior emissions estimate to the true state **x**. **A** can be calculated as or alternatively as (Rodgers 2000)

The trace of **A** measures the number of pieces of information that can be independently constrained by the inverse system, known as the degrees of freedom for signal or DOFS (Rodgers 2000).

In an analytic inversion of satellite observations where the number of observations is much larger than the number of state vector elements, *m* >> *n*, the computational cost is determined by the cost of constructing the Jacobian matrix **K**. The Jacobian matrix is constructed column-wise by computing the model response to a perturbation of each state variable, requiring *n* + 1 forward model simulations covering the observation period (Maasakkers et al. 2019). This construction is computationally expensive. The DOFS of the inverse system may also limit the useful size of *n*. The computational cost of constructing the Jacobian matrix can be reduced by (1) decreasing the dimension *n* by aggregating together state variables or (2) retaining the original dimension but decreasing the number of model simulations to generate a low-rank approximation of the Jacobian. We consider both of these approaches. We discuss first optimal reductions in both dimension and rank for an inverse system with a known Jacobian matrix (Section 2.2). We then discuss a two-step update method to approximate the Jacobian matrix using the defined reductions in dimension and rank (Sections 2.3 through 2.5). As a test case, the state vector is taken to be a gridded field of emissions, where the dimension *n* is defined by the spatial resolution, which may vary across the grid. The method can also be generalized to a temporally resolved state vector.

*Section 2.2: Optimal Reductions in Dimension and Rank of Inverse Systems*

In an inverse system with a known Jacobian matrix, reductions in dimension and rank can decrease the memory needed to store and manipulate the large, non-sparse matrices needed for the analytic solution. Figure 1 illustrates dimension and rank reductions for an emission grid over North America. The upper left panel represents the original *n*-dimensional state space, i.e. the native-resolution grid. A linear transformation reduces the dimension of the state space from *n* to *k*. This transformation may reduce dimension discretely, as in the case of grid cell aggregation (upper right panel), or non-discretely, as in the case of GMM (lower left panel). A second linear transformation can extend the dimension of the state space from *k* back to the original *n*. The resulting space, depicted in the lower left, is a low-rank approximation of the original state space. The projection transforms the original state space to the low-rank subspace. The inverse problem can be solved in any of these three spaces.

We would like to define and matrices to minimize the information loss associated with reducing the dimension or rank of the state vector. To find the optimal values for both and , we consider a projection that maximizes the probability of the full dimension state vector **x** given the reduced dimension state vector in the same way that a Bayesian inversion maximizes the probability of the state vector **x** given the observations **y**. Bousserez and Henze (2018) show that this is achieved when where . Within the set of projections associated with all linear operators , information loss is minimized when the DOFS of the resulting reduced-rank space is maximized. Bousserez and Henze (2018) find that for a projection of this form, where and **A** are the low-rank and original resolution averaging kernel matrices, respectively (equation (4)). Define

where the columns of **W** are the eigenvectors of and the diagonal elements of are the corresponding eigenvalues of ranked in descending order. Because is a symmetric semi-positive definite matrix, the eigenvectors form an orthonormal basis for the space spanned by . As a result, and is maximized when where is the matrix of the first *k* columns of **W**, with *k* corresponding to the rank of the projected subspace. The projection that maximizes the information content of the resulting subspace is then

This projection applies a dimension-reducing transformation followed by a dimension-restoring transformation :

The columns give an eigenvector basis of the averaging kernel matrix while the eigenvalues of **Q** give its eigenvalues. The fraction of information content explained by the first *k* columns of is then the sum of the first *k* largest eigenvalues divided by the total DOFS (Bousserez and Henze 2018). We can then select *k* so that most of the information content is explained by the *k* corresponding eigenvectors. Furthermore, the diagonal of

gives the singular values of the pre-whitened Jacobian matrix , representing the signal-to-noise ratio of each eigenvector (Rodgers 2000). The rank *k* can then be chosen so that the *k* leading eigenvectors explain most of the information content in the inverse system or so that all eigenvectors have a sufficiently large signal-to-noise ratio.

*Section 2.3: Dimension and Rank Reductions of the Jacobian Matrix*

Section 2.2. described optimal reductions in dimension and rank of a state vector assuming knowledge of the native-resolution Jacobian matrix **K**. Our goal is to avoid explicit construction of this matrix. We present a two-step update method to construct **K** at decreased computational cost by reducing its dimension and rank. We start from an initial native-resolution estimate at no computational cost using a simple assumption for atmospheric transport, as described in the next paragraph. We calculate the corresponding averaging kernel matrix . We then propose two methods for updating the Jacobian matrix on the basis of these eigenvectors. First, we construct a multiscale grid that maintains resolution in areas of highest information content and use the forward model to generate the updated Jacobian matrix on the resulting grid, illustrated by the upper right panel of Figure 1 (Section 2.4). Second, we construct on the basis the dominant eigenvectors by perturbing those patterns in the forward model, generating a reduced-rank approximation, illustrated by the lower left panel of Figure 1 (Section 2.5). In both cases, the updated Jacobian matrix improves the estimate of the averaging kernel matrix and its eigenvectors by incorporating information content from forward model. We use these improved eigenvectors to calculate a final Jacobian matrix update.

An initial estimate of the native-resolution Jacobian matrix is generated by assuming that emissions produce local concentration enhancements defined by wind speed. In the case where the observations are column concentrations [molecules cm-2], as with satellite retrievals, the sensitivity of observation *i* to emissions in grid box *j*, representing the Jacobian matrix element , is given by

where is a factor that decreases with the distance of observation *i* from grid box *j*, *M*air and *M*x are the molecular weights of dry air and the optimized species, respectively, *g* is gravity, *U* is the wind speed, *W* is a length scale for the grid box given by the square root of the grid box area, and *P* is the surface pressure.

*Section 2.4: Constructing the Jacobian Matrix on a Multiscale Grid* # EDIT

In an inverse system with a known native-resolution Jacobian matrix **K**, a multiscale grid can be constructed that maintains the native state vector resolution where the information content, as given by the diagonal elements of the corresponding averaging kernel **A**, is highest and aggregates together grid cells elsewhere. An example of such a grid is shown by the upper right panel of Figure 1. The number of grid cells where native or near-native resolution is maintained is determined by the computational resources available but the fraction of DOFS explained should not be less than . Grid cells can be consolidated by, for example, K-means clustering, which aggregates together spatially proximate grid cells that are likely to yield similar model responses but neglects the variation in emissions magnitudes or profiles in different grid cells. An algorithm that considers the similarity of emissions, such as the Gaussian mixture model implemented by Turner and Jacob (2015), could also be used. A reduced-dimension Jacobian matrix can then be constructed on the multiscale grid by running the forward model for perturbations of each grid element.

Without a fully characterized Jacobian matrix, the multiscale grid can be constructed in a two-step update that iteratively improves an initial estimate of the inverse system’s information content. We use the initial estimate of the Jacobian matrix to calculate an averaging kernel matrix . We then construct an initial multiscale grid and the associated reduced-dimension Jacobian by running the forward model for each element of the grid. This introduces information content from the forward model to the inverse system. We then disaggregate the clusters with highest information content in the initial multiscale grid and update the reduced-dimension Jacobian, generating . Alternatively, we could regrid to the native resolution and repeat the process, adjusting the multiscale grid only where the relative difference between the diagonal elements of and is sufficiently large. However, regridding requires the introduction of additional information regarding the distribution of sensitivities.

The information content associated with both and includes contributions from prior emissions estimates, the observations, and the forward model. As a result, the distribution of information content is largely the same for both updates. Further iterations would improve the characterization of information content in grid boxes with small diagonal averaging kernel matrix values, but these grid boxes would not change the multiscale grid. We therefore take as our Jacobian matrix. The analytic inversion can then be solved exactly on the multiscale grid. In order to interpret the inverse results at the original state vector resolution, additional information must be introduced to allocate the posterior solution to the original grid. The weights could, for example, be given by the prior emissions estimate. Or, they could be determined by the optimal dimension restoring transformation (equation 8) corresponding to the dimension reducing transformation associated with the grid cell aggregation. This suggests the possibility of constructing a Jacobian matrix at the native resolution using the optimal transformations and , generating a reduced-rank approximation.

*Section 2.5: Constructing a Reduced-Rank Approximation of the Jacobian Matrix*

In an inverse system with a known native-resolution Jacobian matrix **K**, a reduced-rank approximation of the Jacobian can be constructed by calculating the linear relationship between emissions and observations for the most important patterns of information content rather than grid cells, as in all previous examples. A low-rank Jacobian corresponds to the state space shown in the lower left panel of Figure 1. We showed earlier that the most important patterns of information content are given by the columns of the dimension-restoring transformation (equation 8). For any *k*, the *k* leading patterns span a rank *k*, dimension *n* subspace of the original information content space. A Jacobian matrix can be constructed within this space by calculating the model response to perturbations of these patterns. The response of a forward model **F** to the *j*th eigenvector , given by the *j*th column of , is

where is a scaling factor applied to ensure numerical stability. The model responses form the columns of the matrix, which is the Jacobian for an inverse system with a reduced-dimension state space spanned by the first *k* eigenvectors of the information content, illustrated by the lower right panel of Figure 1. This reduced-dimension Jacobian must be transformed to the original state dimension for use in analytic inversions. If the forward model is linear, F can be written as and as . Bousserez et al. (2018) show that the reduced-rank Jacobian is given by Then, the reduced-rank Jacobian can be calculated from the reduced-dimension Jacobian by . The resulting Jacobian has dimension and rank .

In an inverse system without a known Jacobian matrix, the reduced-rank Jacobian matrix approximation can be constructed in a two-step update that iteratively improves the patterns of information content used as perturbations. We use the initial estimate of the Jacobian matrix to calculate an averaging kernel matrix and its eigenvectors . We select the eigenvectors that have a signal-to-noise ratio greater than or equal to one and calculate the model response to each of the eigenvectors using equation (9). We transform the resulting reduced-dimension Jacobian to the full-dimension state space with , generating a rank approximation of the Jacobian on the basis of the initial eigenvectors of information content.

The reduced-rank Jacobian matrix approximation introduces information from the forward model to the inverse system. We calculate the associated averaging kernel matrix and its eigenvectors . We use the eigenvectors that span most of the information content from the initial estimate, as defined by the eigenvalues of . We cannot use the updated information content to determine *k* because we reduced the rank of the inverse system, so most of the information content described by is contained in the first eigenvectors. However, underestimates the true rank of the system because information is contained in eigenvectors even when their signal-to-noise ratio is less than one. In contrast, the eigenvalues of may underestimate the DOFS relative to the “true” inverse system but are likely to accurately capture the spectrum of information content; information content from the forward model is unlikely to change the rate at which information content decreases with increasing eigenvector index. Then, on the basis of the first eigenvectors of information content, we construct an updated reduced-rank Jacobian matrix approximation as above.

The resulting Jacobian matrix is a rank approximation of the linear forward model. It accurately quantifies the forward model where the inverse system has high information content as and loses accuracy in areas with lower information content. The resulting posterior emissions and error are similarly biased. In areas with lower information content, the posterior emissions estimate tends toward the prior emissions estimate. To eliminate outliers and reduce error in the inverse solution, we set the posterior emissions estimate to the prior value in grid cells where the trace of the averaging kernel matrix is small. We discuss the selection of the threshold for information content in Section 3, where we demonstrate the reduced-dimension and reduced-rank approaches as they apply to an inversion of atmospheric methane columns observed by the GOSAT satellite over North America in July 2009.

**Section 3: Results and Discussion**

We demonstrate both the reduced-dimension and reduced-rank Jacobian matrix construction approaches in an analytic Bayesian inversion of atmospheric methane columns observed by the GOSAT satellite over North America in July 2009. We construct a true Jacobian matrix for 2,098 grid boxes at 1º x 1.25º resolution, a reduced-dimension Jacobian matrix with ~300 state vector elements, and a rank 200 Jacobian matrix. We also construct a reduced-dimension Jacobian matrix following the Gaussian mixture model (GMM) method described by Turner and Jacob (2015). We use those Jacobian matrices within the inverse framework described by Maasakkers et al. (2019), adapted from the global system to the North American domain and with artificially increased information content. We solve for and compare posterior scaling factors, error covariances, and averaging kernels.

In all inversions, we use the nested North American GEOS-Chem CTM version 12.4.0 as forward model to simulate atmospheric methane column concentrations at 0.5º x 0.625º resolution. We aggregate native resolution grid boxes to generate a state vector composed of 2,098 1º x 1.25º grid boxes. The model is driven with MERRA-2 meteorological fields (Bosilovich et al., 2016) from the NASA Global Modeling and Assimilation Office (GMAO). We use boundary conditions and initial conditions from a global GEOS-Chem 4º x 5º simulation for July 2009 driven by prior emissions with posterior scaling factors applied as described by Maasakkers et al. (2019). All inversions also use the prior emissions, prior error covariances, observations, and observational error covariances as described by Maasakkers et al. (2019). In particular, we use the University of Leicester version 7 CO2 proxy retrieval over land (Parker et al. 2011, 2015) for July 2009, excluding glint data. Unlike Maasakkers et al. (2019), we use observations north of 60ºN. Any bias introduced will exist equally in both the true inverse system and the systems with estimated Jacobians.

Solving the analytic inversion with the true Jacobian yields 40 DOFS for the 2,098 constrained grid boxes. In order to demonstrate that the proposed methods are both computationally affordable and accurate in inverse systems with higher information content, we artificially increase the DOFS in our system. We introduce a regularizing factor that increases the weight of the observational term relative to the emissions term in the cost function:

The regularizing factor functionally decreases the observational error covariance, increasing the DOFS. We set , which increases the DOFS in the true inverse system from 40 to 216. The true posterior emissions associated with this regularization factor are not physical, but produce patterns of scaling factors that are consistent with typical inverse results and that should be reproduced by the reduced-dimension and -rank approaches proposed above.

We first construct an initial estimate for the Jacobian matrix following the approach described in Section 2.3 and generate the associated averaging kernel matrix . Because the averaging kernel matrix is a function of the prior and observational error covariance and the forward model (equation 4), reproduces the patterns of information content attributable to the prior errors and observational density, but has errors generated by the initial estimate of the Jacobian matrix. Figure 2 shows the dependence of the diagonal elements of the native-resolution averaging kernel matrix (upper left panel) on the prior error standard deviation (lower left panel) and observation density (lower right panel). The diagonal elements of are shown in the upper right panel; this initial estimate correctly reproduces the patterns of highest information content. Figure 3 shows that the initial estimate of the averaging kernel matrix reproduces the spectrum of information content. The left panel shows the spectrum of information content given by the DOFS for a series of multiscale grids, while the right panel shows the spectrum given by the sum of eigenvalues of **Q** (equation 5) for a series of reduced-rank Jacobian matrices. In both plots the native resolution spectrum is shown in black and the initial estimate spectrum in orange. Both measures of the spectrum show that the initial estimate of the averaging kernel underestimates the native resolution spectrum while reproducing the rate of information content accumulation.

We use the initial estimate of the averaging kernel matrix first to construct a multiscale grid. The left panel of Figure 3 shows that almost 50% of the initial estimate of information content as measured by DOFS is contained in 100 native resolution grid boxes, 90% in 400 grid boxes, and 99% in 900 grid boxes. We construct a multiscale grid following Section 2.4 with a K-means clustering aggregation scheme. Our initial grid contains ~100 native resolution grid boxes, 100 clusters of ~3 native resolution grid boxes, 100 clusters of ~5, and 150 clusters of ~8. This reduces the total dimension of the state vector to ~450 with the largest cluster containing ~15 native resolution grid boxes. We construct a multiscale Jacobian on this multiscale grid. The updated spectrum of information content is shown by the purple line; information content accumulation has sharply increased due to the reduction in dimension. To account for the added information content generated by the forward model, we disaggregate to native resolution the clusters with the largest DOFS values. We add an additional ~100 native resolution grid boxes by disaggregating 36 grid cells, corresponding to ~25% of the information content. The final multiscale grid has dimension ~550 and the corresponding multiscale Jacobian requires ~550 model runs in total.

Figure 4 shows the final multiscale grid. The grid maintains native resolution where information content is highest (upper left panel of Figure 2), notably over the Gulf Coast and Southeast coast, California’s Central Valley, and South-central Canada. Grid cells are aggregated elsewhere, notably over the American West and western, northern, and eastern Canada, where there is low observation density (lower right panel of Figure 2). We solve the analytic inversion on this grid. Because the regularization factor introduced in equation (10) reflects the balance of the state vector dimension to the observational dimension, we adjust the regularization by the ratio of the new state vector dimension (550) to the native state vector dimension (2,098). Figure 5 shows the posterior emission scaling factors (top) and diagonal averaging kernel values (bottom) for the multiscale solution (center column) compared to the native resolution solution (left column). Both solutions are exact on the grids used. The multiscale solution generates 95 DOFS compared to the native resolution 216 DOFS, reflecting the loss of information content incurred by reducing the dimension.

We also use the initial estimate of the averaging kernel matrix to construct a reduced-rank Jacobian matrix. The right panel of Figure 3 shows the information content spectrum as given by the eigenvalues of **Q** (equation 5). As previously noted, the initial estimate underestimates the information content spectrum. Introducing additional information by running the forward model is unlikely to significantly change the rate at which information content accumulates when the prior and observational error covariances remain constant. We therefore use the spectrum of information content associated with the initial estimate to choose the *k* that determines the rank of updated Jacobian matrix approximations. Because only the dominant patterns of information content are captured by , we require first that the signal-to-noise ratio of all patterns be greater than 1.25, corresponding to *k* = 125. We perturb the leading 125 patterns of information content in the forward model and construct the reduced-rank Jacobian matrix as described in Section 2.5. The eigenvalue spectrum given by the updated information content (purple) has a discontinuity at *n* = 125 as expected for a rank 125 system. We use the initial eigenvalue spectrum to determine *k* for the second update, requiring that the improved patterns capture 97.5% of the information content, corresponding to *k* = 417. The resulting Jacobian matrix has rank ~417 and required ~540 model runs to construct.

We solve the inversion using the resulting reduced-rank Jacobian matrix. Figure 5 (right column) shows the distribution posterior scaling factors (top) and diagonal averaging kernel values (bottom) compared to the native resolution solution (left column). Because was constructed on the basis of the dominant patterns of information content, it solves for the posterior scaling factors accurately in the areas of highest information content and defaults to the prior value (a scaling factor or 1) elsewhere. The averaging kernel accurately captures these patterns of information content. Despite the limited extent of the optimized grid cells, the resulting DOFS (153) are still higher than those given by the multiscale grid solution (95) because the solution optimizes the solution in the areas with highest information content.

Figure 6 shows the final reduced-rank Jacobian matrix approximation (upper left) and the resulting posterior solution, including scaling factors (upper right), variance (lower left), and diagonal averaging kernel values (lower right), plotted against the corresponding native-resolution values. To account for the grid boxes that are not optimized by the inversion, we exclude grid boxes where the diagonal values of are smaller than 0.01, leaving 711 grid boxes. Increasing the threshold will increase the accuracy of the posterior solution relative to the truth but will decrease the number of constrained grid cells. Decreasing the threshold will increase the number of grid boxes where the posterior scaling factor is not optimized and defaults to the prior value. Figure 6 shows that the filtered reduced-rank Jacobian matrix correlates well with the native-resolution values (r2 = 0.93). The reduced-rank posterior variance overestimates the true posterior variance, reflecting the error introduced by discarding information content. Similarly, the diagonal elements of the reduced-rank averaging kernel underestimate the true information content. The posterior scaling factors generate the lowest correlation coefficient (r2 = 0.69) because they are a function of the posterior error covariance and the Jacobian matrix (equation 2), resulting in propagated errors.

We conduct a series of sensitivity tests on the number of model runs used in both the first and second update. Figure 7 summarizes these results. We show the r2 of the posterior emissions scaling factor for the estimated compared to the native-resolution Jacobian matrix plotted against the number of model runs conducted in the first and second iteration. As expected, the correlation coefficient increases with the total number of model runs. We find that the correlation coefficient has a stronger dependence on the number of model runs conducted in the second iteration, consistent with the improved characterization of the information content after the first iteration and confirming that it is preferable to increase the number of model runs in the second iteration if additional computational resources are available.

**Section 4: Conclusions**

I’ll add a conclusion next iteration.

**References**

Bergamaschi, P., S. Houweling, A. Segers, M. Krol, C. Frankenberg, R. A. Scheepmaker, E. Dlugokencky, et al. 2013. “Atmospheric CH4 in the First Decade of the 21st Century: Inverse Modeling Analysis Using SCIAMACHY Satellite Retrievals and NOAA Surface Measurements.” *Journal of Geophysical Research Atmospheres*. https://doi.org/10.1002/jgrd.50480.

Bergamaschi, Peter, Christian Frankenberg, Jan Fokke Meirink, Maarten Krol, M. Gabriella Villani, Sander Houweling, Dentener Frank, et al. 2009. “Inverse Modeling of Global and Regional CH4 Emissions Using SCIAMACHY Satellite Retrievals.” *Journal of Geophysical Research Atmospheres*. https://doi.org/10.1029/2009JD012287.

Bousserez, N., and Daven K. Henze. 2018. “Optimal and Scalable Methods to Approximate the Solutions of Large-Scale Bayesian Problems: Theory and Application to Atmospheric Inversion and Data Assimilation.” *Quarterly Journal of the Royal Meteorological Society* 144 (711): 365–90. https://doi.org/10.1002/qj.3209.

Brasseur, Guy P., and Daniel J. Jacob. 2017. *Modeling of Atmospheric Chemistry*. Cambridge University Press.

Evensen, Geir. 2009. *Data Assimilation: The Ensemble Kalman Filter*. *Data Assimilation (Second Edition): The Ensemble Kalman Filter*. https://doi.org/10.1007/978-3-642-03711-5.

Henze, D. K., A. Hakami, and J. H. Seinfeld. 2007. “Development of the Adjoint of GEOS-Chem.” *Atmospheric Chemistry and Physics Discussions* 7 (9): 2413–33. https://doi.org/10.5194/acpd-6-10591-2006.

Houweling, S., M. Krol, P. Bergamaschi, C. Frankenberg, E. J. Dlugokencky, I. Morino, J. Notholt, et al. 2014. “A Multi-Year Methane Inversion Using SCIAMACHY, Accounting for Systematic Errors Using TCCON Measurements.” *Atmospheric Chemistry and Physics*. https://doi.org/10.5194/acp-14-3991-2014.

Hu, Haili, Jochen Landgraf, Rob Detmers, Tobias Borsdorff, Joost Aan de Brugh, Ilse Aben, Andre Butz, and Otto Hasekamp. 2018. “Toward Global Mapping of Methane With TROPOMI: First Results and Intersatellite Comparison to GOSAT.” *Geophysical Research Letters* 45 (8): 3682–89. https://doi.org/10.1002/2018gl077259.

Jacob, Daniel J., Alexander J. Turner, Joannes D. Maasakkers, Jianxiong Sheng, Kang Sun, Xiong Liu, Kelly Chance, Ilse Aben, Jason McKeever, and Christian Frankenberg. 2016. “Satellite Observations of Atmospheric Methane and Their Value for Quantifying Methane Emissions.” *Atmospheric Chemistry and Physics* 16 (22): 14371–96. https://doi.org/10.5194/acp-16-14371-2016.

Kuze, Akihiko, Hiroshi Suto, Masakatsu Nakajima, and Takashi Hamazaki. 2009. “Thermal and near Infrared Sensor for Carbon Observation Fourier-Transform Spectrometer on the Greenhouse Gases Observing Satellite for Greenhouse Gases Monitoring.” *Applied Optics*. https://doi.org/10.1364/AO.48.006716.

Maasakkers, Joannes D., Daniel J. Jacob, Melissa P. Sulprizio, Tia R. Scarpelli, Hannah Nesser, Jian-Xiong Sheng, Yuzhong Zhang, et al. 2019. “Global Distribution of Methane Emissions, Emission Trends, and OH Concentrations and Trends Inferred from an Inversion of GOSAT Satellite Data for 2010&amp;Ndash;2015.” *Atmospheric Chemistry and Physics Discussions*, no. January: 1–36. https://doi.org/10.5194/acp-2018-1365.

Rodgers, Clive D. 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. *World Scientific Publishing Co.Pte.Ltd.*

Saunois, Marielle, Ann R. Stavert, Ben Poulter, Philippe Bousquet, Joseph G. Canadell, Robert B. Jackson, Peter A. Raymond, et al. 2019. “The Global Methane Budget 2000&ndash;2017.” *Earth System Science Data Discussions*. https://doi.org/10.5194/essd-2019-128.

Spantini, Alessio, Antti Solonen, Tiangang Cui, James Martin, Luis Tenorio, and Youssef Marzouk. 2015. “Optimal Low-Rank Approximations of Bayesian Linear Inverse Problems.” *SIAM Journal on Scientific Computing* 37 (6): 2451–87.

Streets, David G., Timothy Canty, Gregory R. Carmichael, Benjamin De Foy, Russell R. Dickerson, Bryan N. Duncan, David P. Edwards, et al. 2013. “Emissions Estimation from Satellite Retrievals: A Review of Current Capability.” *Atmospheric Environment*. https://doi.org/10.1016/j.atmosenv.2013.05.051.

Veefkind, J. P., I. Aben, K. McMullan, H. Förster, J. de Vries, G. Otter, J. Claas, et al. 2012. “TROPOMI on the ESA Sentinel-5 Precursor: A GMES Mission for Global Observations of the Atmospheric Composition for Climate, Air Quality and Ozone Layer Applications.” *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2011.09.027.

Wecht, Kevin J., Daniel J. Jacob, Christian Frankenberg, Zhe Jiang, and Donald R Blake. 2014. “Mapping of North American Methane Emissions with High Spatial Resolution by Inversion of SCIAMACHY Satellite Data.” *J. Geophys. Res. Atmos. Res.*, 7741–56. https://doi.org/10.1002/2014JD021551.Received.