**Reduced-Cost Construction of Jacobian Matrices for High-Resolution Inversions of Satellite Observations**

**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 composition from satellites can greatly improve our knowledge of surface emissions through inverse analyses. Variational inverse methods can optimize surface emissions at any grid 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. 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 many runs of an atmospheric transport model perturbing individual emission grid cells. 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 a powerful resource to improve our knowledge of emissions (Streets et al. 2013) but are subject to large errors, both from the measurement and from the inverse analyses used to infer emissions from the observations (Brasseur and Jacob 2017). Conducting high-resolution inverse analyses of the satellite data to quantify emissions at commensurate scales is of considerable interest but may be limited by data quality in ways that are difficult to determine and may compromise the results. Here we present two alternative methods to conduct high-resolution inversions of satellite observations in a manner that maximizes the information content of the observations, minimizes computational cost, and includes error characterization.

Inverse analyses infer emissions by fitting the observed atmospheric concentrations to a chemical transport model (CTM) that simulates atmospheric concentrations as a function of emissions (Brasseur and Jacob 2017). The CTM represents the forward model for the inverse problem. The solution is generally obtained by minimizing a Bayesian cost function regularized by a prior emissions estimate. The optimal (posterior) estimate of emissions corresponds to the minimum of the cost function. This 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 methods of estimating the error exist (Evensen 2009; Bousserez and Henze 2018), these approaches are computationally expensive.

In the frequent case where the observed atmospheric concentrations depend linearly on emissions and the error statistics can be assumed to be normally or log-normally distributed, the Bayesian optimization problem has an analytic solution including closed-form expressions for the posterior emissions estimate, its error statistics, and its information content (Rodgers 2000; )(Maasakkers et al. 2019). 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 elements of **y** are the individual observations and the elements of **x** are the emissions to be optimized, often grid cells in a two-dimensional emissions field. The Jacobian can be 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, allowing study of the sensitivity of the solution to changes in the specification of errors, prior assumptions, and the number and type of observations.

An illustrative example is the inversion of satellite observations to infer methane emissions. Methane is an important greenhouse gas but the spatial and temporal distribution of emissions is highly uncertain (Saunois et al. 2019). Satellite observations of atmospheric methane columns can improve emission estimates (Jacob et al. 2016). This was first shown with data from the SCIAMACHY satellite instrument (2003-2012, nadir pixel resolution of 30 x 60 km2) (Bergamaschi et al. 2009, 2013; Houweling et al. 2014; Wecht et al. 2014). More recent inversions used observations from the TANSO-FTS instrument aboard the GOSAT satellite (2009-present) (Monteil et al. 2013; Alexe et al. 2015; Turner et al. 2015; Maasakkers et al. 2019), which measures column methane concentrations in 10-km diameter pixels approximately 250 km apart along- and cross-track (Kuze et al. 2009). The Tropospheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 precursor, launched in October 2017, now provides daily, global retrievals of atmospheric methane columns at 7 x 7 km2 nadir pixel resolution, increasing coverage by orders of magnitude relative 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). Inversions of TROPOMI data must attempt to capture the high resolution and density of observations where appropriate while recognizing the limitations in information content resulting from data sparsity or errors.

Several methods have been proposed to decrease the computational cost of high-resolution analytic inversions by optimally reducing the dimension or rank of the state vector. Reduced-dimension methods (Bocquet et al. 2011; Turner and Jacob 2015) solve inversions on a multiscale emission grid of dimension *k* < *n* for which the Jacobian matrix can be computed. Bocquet et al. (2011) defined a method to find the optimal multiscale grid from an array of all allowable grids, but this requires a large computational investment. Turner and Jacob (2015) used prior emissions information to group together similar grid cells using a Gaussian mixture model, but the criteria used to define similarity were subjective and did not consider the information content of the forward model or observations. Reduced-rank methods (Spantini et al. 2015; Bousserez and Henze 2018) generate an approximation of the inverse solution at the original dimension *n* by solving for the directions of most information content. Bousserez and Henze (2018) avoided explicit construction of the Jacobian matrix by using randomized methods to estimate these directions. This approach is computationally tractable only for inverse systems where a small number of directions explain the information content.

Here we present two alternative methods to construct the Jacobian matrix for a native *n*-dimensional state vector that maximize the information content of the inverse analysis using *k* < *n* forward model simulations. The first method generates a multiscale grid that preserves native resolution where information content is highest and aggregates grid boxes elsewhere. The resulting reduced-dimension Jacobian matrix solves the inversion exactly on the multiscale grid. The second approach constructs a reduced-rank Jacobian matrix along the dominant patterns of information content in the system, allowing solution of the inversion at native resolution. In both cases, a low-cost initial estimate of the Jacobian matrix is updated using *k* model simulations where *k* is selected by the user as 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**

This section presents the reduced-dimension and reduced-rank methods of Jacobian matrix construction. Following a review of the standard analytic inverse framework (Section 2.1), we describe optimal reductions in both dimension and rank for an inverse system with a known native-resolution Jacobian matrix (Section 2.2). We then present a two-step approach to approximate that native-resolution matrix, when initially unknown, using either the reduced-dimension or reduced-rank methods (Sections 2.3 through 2.5). For purpose of illustration we take the state vector to be a gridded field of emissions, although the methods apply to any state vector.

*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 **S**Aand **S**O, respectively, is given by the minimization of the Bayesian scalar cost function *J*(**x**):

where **F**(**x)** represents the forward model that simulates the observations **y** (Brasseur and Jacob 2017). The observational error covariance matrix **S**O includes errors from both the measurement and the forward model (collectively the observing system). If the forward model is linear so that **F**(**x**) = **Kx** + **c**, where is the Jacobian matrix calculated via finite difference (see Introduction) and **c** is a constant, then an analytic solution to the cost function minimum exists that yields the posterior estimate and its error covariance matrix :

Comparison of and defines the information content of the inverse system, quantified by the averaging kernel matrix that describes the sensitivity of the posterior estimate to the true state **x**. **A** can be calculated as or equivalently as

Equation (4) expresses the dependence of the averaging kernel matrix on the forward model and on the error covariance matrices **SA** and **SO** (Rodgers 2000). The diagonal elements of **A** are commonly referred to as the averaging kernel sensitivities. The sum of the sensitivities, or the trace of **A**, measures the number of pieces of information that can be independently quantified by the inverse system, known as the degrees of freedom for signal or DOFS (Rodgers 2000).

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

We first consider the problem of optimally reducing the dimension and rank of an inverse system with a known Jacobian matrix . Figure 1 illustrates dimension and rank reductions for a hypothetical 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 matrix 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, in which case the *k* state vector components are themselves *n*-dimensional vectors (lower right panel). A second linear transformation restores 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 four spaces.

We would like to define matrices and that minimize the information loss associated with reducing the dimension or rank of the state vector. Following Bousserez and Henze (2018), the projection that maximizes the probability of restoring the original full dimension state vector **x** given the reduced dimension state vector is given by where . For a projection of this form, they show that information loss is minimized by maximizing where and **A** are the low-rank and original resolution averaging kernel matrices, respectively. Define

where the columns of **W** are the eigenvectors of and the diagonal elements of are the corresponding eigenvalues of ranked in descending order. Bousserez and Henze (2018) show that 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 corresponding optimal projection is then

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

The columns of 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 the sum of the first *k* largest eigenvalues divided by the total DOFS (Bousserez and Henze 2018). We will refer to the rate at which the information content accumulates as the number of eigenvectors increases as the information content spectrum. On the basis of this spectrum, we can select *k* so that most of the information content is explained by the *k* corresponding eigenvectors. We can also select *k* so that all eigenvectors have a sufficiently large signal-to-noise ratio. The diagonal of

gives the singular values of the pre-whitened Jacobian matrix , representing the signal-to-noise ratio of each eigenvector (Rodgers 2000). This criterion is stricter than the information content criterion because even eigenvectors with a low signal-to-noise ratio can contribute to the total information content.

*Section 2.3: Approximating 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 in order to decrease computational cost. Here we present a two-step update to do so by reducing either the dimension or the rank of the native-resolution state vector using the optimal transformations described above. We start from an initial native-resolution estimate and calculate the corresponding averaging kernel matrix . We then consider two methods for updating the Jacobian matrix. In the first method (reduced-dimension), we construct a multiscale grid that maintains resolution in areas of highest information content (upper right panel of Figure 1) and use the forward model to generate the updated, reduced-dimension Jacobian matrix on the resulting grid. In the second method (reduced-rank), we construct on the basis of the *k* dominant eigenvectors by perturbing those patterns in the forward model, generating an approximation of the Jacobian matrix in a reduced-rank state space (lower left panel of Figure 1). 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 either or to conduct a second update and construct the final Jacobian matrix. Although the reduced-rank method could also be applied in *k* space (lower-right panel of Figure 1), the results of the inversion expressed as eigenvector corrections to the prior estimates would be more difficult to interpret.

An initial estimate of the native-resolution Jacobian matrix can be generated in our illustrative case at no cost by assuming that methane emissions [kg m-2 s-1] produce local column mixing ratio enhancements [ppb] dependent on local wind speed and parameterized turbulent diffusion. The sensitivity of observation *i* to emissions in grid cell *j*, representing the Jacobian matrix element , is given by

where is a dimensionless coefficient that decreases with the distance of observation *i* from grid box *j*, and are the molecular weights of dry air and methane, respectively, *g* is the acceleration from gravity, *U* is the local 10-m wind speed, *W* is a length scale for the grid cell given by the square root of the grid cell area, and *p* is the surface pressure. We use wind speeds from the NASA MERRA-2 database (Bosilovich). The coefficients provide a crude mass-conserving representation of turbulent diffusion to reduce the sparsity of . We define for the grid box where the observation is found with the remaining mass distributed over the three concentric rings surrounding that box as 0.3, 0.2, and 0.1.

The initial estimate of the Jacobian matrix can be used to calculate the associated averaging kernel matrix . The averaging kernel matrix is a function not only of the Jacobian matrix but of the prior and observational error covariance matrices (equation (4)). Therefore, the initial estimate of already provides a good approximation of the patterns of information content as determined by the prior errors and observational density, and despite the crude but neutral estimate of **K**. This will be illustrated in Section 3.

*Section 2.4: Constructing the Reduced-Dimension Jacobian Matrix*

In an inverse system with a known native-resolution Jacobian matrix **K**, a reduced-dimension Jacobian matrix can be constructed on a multiscale grid that maintains native resolution where information content is highest and aggregates grid cells elsewhere (upper right panel of Figure 1). An optimal multiscale grid maximizes the total DOFS and the averaging kernel sensitivities of each state vector element, referred to here as the DOFS per cluster. To construct this grid, the state vector is first defined as a single element that encompasses the inversion domain. The native-resolution grid cells with the highest averaging kernel sensitivities are then added one-by-one to the state vector. For each new element , we calculate the corresponding Jacobian matrix column and the resulting increase in DOFS. Eventually the DOFS level off, at which point we start adding instead clusters of two or more native-resolution grid cellsand repeat the operation. . Clusters can be generated by K-means clustering, which aggregates together spatially proximate 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. We repeat this process, increasing cluster size, until all native-resolution grid cells are allocated to the multiscale grid and the corresponding reduced-dimension Jacobian matrix is constructed.

Here we use the same approach beginning with our initial estimate in a two-step update process that iteratively improves the multiscale grid. We assume that the fine structure of the averaging kernel sensitivities is largely determined by the prior errors and observational density so that the initial estimate correctly identifies the grid cells with the highest sensitivities. We construct a multiscale grid and apply the forward model to construct the corresponding reduced-dimension Jacobian matrix . This construction introduces information content from the forward model to the inverse system. To identify the state vector elements where the forward model contributes the most information content, we compare the sensitivities given by the updated reduced-dimension averaging kernel matrix to the sensitivities given by the initial estimate regridded onto the multiscale grid. We disaggregate the clusters with the largest difference and update the reduced-dimension Jacobian, generating .

The information content associated with both and includes contributions from prior emissions estimates, the observations, and the forward model. As a result, there is rapid convergence and we find that there is no need for further iteration. We therefore take as our Jacobian matrix. The analytic inversion can then be solved exactly on the multiscale grid.

*Section 2.5: Constructing the Reduced-Rank Jacobian Matrix*

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

where is any scalar sufficiently large 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 to enable physical interpretation of inverse results. Bousserez and Henze (2018) show that the reduced-dimension Jacobian matrix is given by and the reduced-rank Jacobian matrix by Thus, 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 the corresponding averaging kernel matrix and the matrix of its eigenvectors . When calculating , we select the eigenvectors that have a signal-to-noise ratio greater than some threshold We calculate the model response to each of the eigenvectors using equation (9) and transform the resulting reduced-dimension Jacobian to the full-dimension state space with .

The reduced-rank Jacobian matrix approximation introduces information from the forward model to the inverse system, improving the characterization of the eigenvectors of information content relative to the native-resolution system. We calculate the associated averaging kernel matrix and the matrix of its eigenvectors . Because is a reduced-rank approximation, its spectrum of information content is discontinuous at . To select the rank of the second update and calculate , we therefore use the spectrum of information content associated with the initial, full-rank estimate . While is not informed by the forward model, we expect that it captures with some accuracy the spectrum of information content because much of its structure is determined by the prior error covariance matrix and observational density (Section 2.1). We use the eigenvectors that span most of the information content from the initial estimate to construct an updated reduced-rank Jacobian matrix approximation as above. The resulting Jacobian matrix is a rank approximation that accurately quantifies the forward model where the inverse system has high information content as and loses accuracy in areas with lower information content where the observations are least able to constrain emissions.

**Section 3: Results and Discussion**

We demonstrate both the reduced-dimension and reduced-rank Jacobian matrix construction methods in an analytic Bayesian inversion of atmospheric methane columns observed by the GOSAT satellite over North America in July 2009. Although TROPOMI now provides higher density observations, using GOSAT allows us to follow the framework developed by Maasakkers (2019) to construct a “native-resolution” inverse system at 1º x 1.25º grid cell resolution (*n* = 2098, upper left panel of Figure 1) against which our reduced-dimension and reduced-rank methods can be compared. To make this demonstration more relevant to a higher-information observing system such as TROPOMI, we artificially increase the information content of the GOSAT data by introducing a regularization factor to the cost function that increases the weight of the observational terms relative to the prior terms:

The regularization factor functionally decreases the observational error covariance, increasing the DOFS. We set . In the native-resolution inversion, this increases the DOFS from 40 to 216. Because of noise in the GOSAT data, this artificial increase in information content results in overfit and unphysical solution, but that does not affect our demonstration.

We use the nested North American GEOS-Chem CTM version 12.4.0 as forward model to simulate atmospheric methane column concentrations at 1º x 1.25º resolution for July 2009. This is taken as our native resolution and includes 2098 grid cells for which methane emissions are to be optimized. 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 posterior GEOS-Chem 4º x 5º simulation from Maasakkers et al. (2019). Relative prior emissions, relative prior error covariances, observations, and observational error covariances are also from Maasakkers et al. (2019). In particular, we use the GOSAT data from the University of Leicester version 7 CO2 proxy retrieval over land (Parker et al. 2011, 2015) for July 2009. Unlike Maasakkers et al. (2019), we use observations north of 60ºN; any large errors in these data do not matter for the sake of our demonstration. After constructing the native-resolution Jacobian matrix, we use it as the forward model in lieu of additional GEOS-Chem simulations.

Figure 2 (upper left panel) shows the averaging kernel sensitivities of the native-resolution averaging kernel matrix **A**. As discussed in Section 2 and illustrated in Figure 2, the sensitivity patterns are largely driven by the prior error standard deviation (lower left panel) and the observation density (lower right panel) due to the large variability in these terms. The variability in the prior error standard deviation is driven by the spatial variability of the prior methane emissions because the errors are specified as. The variability in observation density is driven by variability in sampling and retrieval success, which in the GOSAT case is limited mainly by clouds.

Figure 2 (upper right panel) also shows the initial estimate of averaging kernel sensitivities derived from the initial estimate of the Jacobian matrix constructed as described in Section 2.3. While no forward model simulations were conducted to construct this initial estimate, the patterns of information content as given by the sensitivities closely reproduce those given by the native-resolution averaging kernel matrix **A** because they are again largely determined by the prior error standard deviation and the observation density.

We first construct a reduced-dimension Jacobian on a multiscale grid. Figure 3 shows the evolution of the DOFS per cluster as a function of the number of native resolution grid cells added to the state vector as increasingly coarse state vector elements, aggregated by a K-means clustering algorithm and color coded by cluster size, are added to the multiscale grid. We begin with a state vector with one domain-encompassing element. To account for the functional decrease in the observational error covariance resulting from the increase in the number of observations constraining each state vector element, we scale down the regularization factor introduced in equation (10) by the ratio of the new state vector dimension to the native state vector dimension. As result of this scaling, the initial DOFS per cluster is less than one. We add to this single element state vector the native resolution grid cells with the highest estimated averaging kernel sensitivities as given by . We add 50 cells at a time to reproduce the construction of the multiscale grid in a parallel computing environment. The DOFS per cluster decreases initially because the high-resolution grid cells have lower DOFS. After adding 150 native-resolution grid cells (151 model simulations, including the single state vector element simulation), we find that the DOFS per cluster consistently decreases and that the maximum DOFS per cluster occurred at 80 native-resolution grid cells (black shading). We then add to our state vector of 80 native-resolution grid cells clusters of first ~2 (purple shading), then ~4 (maroon), ~8 (dark orange), ~16 (light orange), and ~32 (yellow) native-resolution grid cells, repeating the same procedure as before: we add clusters in batches of about 50 (and then 25) until the DOFS per cluster decreases or stabilizes, at which point we increase the cluster size. The resulting multiscale grid is summarized in the table at the bottom of Figure 3; the reduced-dimension Jacobian matrix constrains 359 clusters and required 470 model simulations to construct.

We then update the multiscale grid and the reduced-dimension Jacobian matrix to reflect the information added by the forward model. We identify the clusters where the forward model contributed the most information by regridding the sensitivities given by to the multiscale grid and calculating the change in the sensitivities. We find that the sensitivity of 165 clusters increased, 162 of which were clusters of 2 or more native-resolution grid cells. We disaggregate 16 clusters with a sensitivity increase greater than 0.4, adding 64 native-resolution grid cells and forward model simulations. Disaggregating additional clusters would increase the number of forward model simulations beyond the computational limit of 600 model runs for this demonstration. The updated multiscale grid increases the resolution of the inverse solution, increases the total DOFS from 79 to 89, and maintains the DOFS per cell at 0.21.

Figure 4 shows the final multiscale grid. The final multiscale grid has dimension 423 and the corresponding reduced-dimension Jacobian matrix required 534 forward model simulations. The final grid has 199 native-resolution grid cells and clusters ranging from size 1 to size 58.

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 and compare the results to the native-resolution solution. Figure 5 shows the averaging kernel sensitivities (top) and posterior emission scaling factors (bottom) for the reduced-dimension solution (center column) compared to the native-resolution solution (left column). Both solutions are exact on the grids used. The reduced-dimension solution generates fewer DOFS (89) than the native-resolution solution (216) but twice as many DOFS per cell (0.21 compared to 0.10), reflecting the consolidation of information content. This is reflected in the plotted sensitivities: the reduced-dimension solution has more uniform sensitivities relative to the native-resolution solution. The reduced-dimension posterior scaling factors exhibit less variability than the native-resolution values. However, some of the variability in the native-resolution solution is attributable to the large regularization factor, which causes “checker-boarding” where the posterior solution fits observational noise. The reduced-dimension scaling factors successfully capture the native-resolution regional patterns.

We next construct a reduced-rank Jacobian matrix. We calculate the dominant eigenvectors of the initial averaging kernel matrix estimate, requiring that the signal-to-noise ratio of all eigenvectors be greater than 2, corresponding to 44% of the DOFS and *k* = 92. We perturb these eigenvectors in the forward model and construct the reduced-rank Jacobian matrix as described in Section 2.5. We then recalculate averaging kernel matrix and its dominant eigenvectors. The eigenvalue spectrum given by the updated information content has a discontinuity at *n* = 113 as expected for a rank 113 system. We use the initial eigenvalue spectrum to determine *k* for the second update, requiring that the improved eigenvectors capture 98% of the information content, corresponding to *k* = 437. The resulting Jacobian matrix has rank ~437 and required 530 forward model simulations including the prior run, a 75% reduction from the 2099 simulations required for the native-resolution solution.

We solve the inversion (equations (2) – (4)) using and compare the information content to the native-resolution solution. The top row of Figure 5 shows the distribution of the reduced-rank averaging kernel sensitivities (right) compared to the native resolution solution (left). Because was constructed on the basis of the dominant patterns of information content, it solves for the posterior scaling factors in the areas of highest information content and defaults to the prior value (a scaling factor of 1) elsewhere. Visual inspection suggests that the reduced-rank averaging kernel sensitivities accurately capture these patterns of information content. The lower right panel of figure 5 confirms this conclusion. We show the reduced-rank averaging kernel sensitivities greater than 0.01 plotted against the corresponding native-resolution values. The reduced-rank sensitivities correlate strongly (R = 0.91) with the native-resolution values. The reduced-rank sensitivities tend to be lower, reflecting the loss of information content in the reduced-rank solution; the reduced-rank DOFS (153) and DOFS per grid cell (0.07) are lower than native-resolution values (216 and 0.10, respectively). If we consider only the optimized grid cells (averaging kernel sensitivities greater than 0.01), we find 152 DOFS across 679 grid cells, or 0.22 DOFS per grid cell. The reduced-rank solution generates only ~70% of the native-resolution DOFS despite constructing the reduced-rank Jacobian matrix on the basis of eigenvectors that capture 98% of the information content; the discrepancy results from errors in the characterization of the tailing eigenvectors.

We also compare the reduced-rank posterior scaling factors to the native-resolution solution. The bottom row of figure 4 shows the distribution of the reduced-rank (right) and native-resolution (left) posterior scaling factors. The reduced-rank solution captures the native-resolution solution where the averaging kernel sensitivities are highest. The upper right panel of figure 5 supports this conclusion: the filtered reduced-rank posterior scaling factors correlate strongly with the native-resolution values (R = 0.87). The correlation is lower than that exhibited by the Jacobian matrix elements (upper left, R = 0.97), posterior variances (lower left, R = 0.98), and averaging kernel sensitivities due to the propagation of errors (equation (2)). While the reduced-rank inversion underestimates the averaging kernel sensitivities and overestimates the posterior variances, reflecting the loss of information content, the error in the posterior scaling factors appears random.

Beyond the total number of forward model runs that can be afforded, one must decide how to partition these runs between the first and second update. Figure 6 summarizes these results for the reduced-rank method as the DOFS plotted against the number of model runs conducted in the first and second iteration. As expected, the DOFS increase with the total number of model runs. We find that the DOFS have 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**

**References**

Alexe, M., P. Bergamaschi, A. Segers, R. Detmers, A. Butz, O. Hasekamp, S. Guerlet, et al. 2015. “Inverse Modelling of CH4 Emissions for 2010-2011 Using Different Satellite Retrieval Products from GOSAT and SCIAMACHY.” *Atmospheric Chemistry and Physics*. https://doi.org/10.5194/acp-15-113-2015.

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.

Bocquet, M., L. Wu, and F. Chevallier. 2011. “Bayesian Design of Control Space for Optimal Assimilation of Observations. Part I: Consistent Multiscale Formalism.” *Quarterly Journal of the Royal Meteorological Society* 137 (658): 1340–56. https://doi.org/10.1002/qj.837.

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.

Buchwitz, M., M. Reuter, O. Schneising, H. Boesch, S. Guerlet, B. Dils, I. Aben, et al. 2015. “The Greenhouse Gas Climate Change Initiative (GHG-CCI): Comparison and Quality Assessment of near-Surface-Sensitive Satellite-Derived CO2 and CH4 Global Data Sets.” *Remote Sensing of Environment* 162: 344–62. https://doi.org/10.1016/j.rse.2013.04.024.

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, Monica Hersher, A. Anthony Bloom, 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.

Maasakkers, Joannes D., Daniel J. Jacob, Melissa P. Sulprizio, Tia R. Scarpelli, Hannah Nesser, Jian Xiong Sheng, Yuzhong Zhang, Monica Hersher, A. Anthony Bloom, et al. 2019. “Global Distribution of Methane Emissions, Emission Trends, and OH Concentrations and Trends Inferred from an Inversion of GOSAT Satellite Data for 2010-2015.” *Atmospheric Chemistry and Physics*. https://doi.org/10.5194/acp-19-7859-2019.

Monteil, Guillaume, Sander Houweling, André Butz, Sandrine Guerlet, Dinand Schepers, Otto Hasekamp, Christian Frankenberg, Remco Scheepmaker, Ilse Aben, and Thomas Röckmann. 2013. “Comparison of CH4 Inversions Based on 15 Months of GOSAT and SCIAMACHY Observations.” *Journal of Geophysical Research Atmospheres*. https://doi.org/10.1002/2013JD019760.

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.

Turner, A. J., and D. J. Jacob. 2015. “Balancing Aggregation and Smoothing Errors in Inverse Models.” *Atmospheric Chemistry and Physics* 15 (12): 7039–48. https://doi.org/10.5194/acp-15-7039-2015.

Turner, A. J., D. J. Jacob, K. J. Wecht, J. D. Maasakkers, E. Lundgren, A. E. Andrews, S. C. Biraud, et al. 2015. “Estimating Global and North American Methane Emissions with High Spatial Resolution Using GOSAT Satellite Data.” *Atmospheric Chemistry and Physics* 15 (12): 7049–69. https://doi.org/10.5194/acp-15-7049-2015.

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.