**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 trace gas concesntrations 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 a resource to improve our understanding of emissions (Streets et al. 2013). However, satellite data 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 is of considerable interest but may be limited by data quality in ways that are difficult to quantify and that may compromise the results. Here we present two methods to conduct high-resolution inversions of satellite observations that maximize the information content of the observations, minimize computational cost, and provide full 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 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). This approach can be extended to non-linear problems and to log-normal error statistics (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 elements of **y** correspond to individual observations and the elements of **x** correspond to the optimized emission components, 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 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, requiring 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 a reduced-rank approximation of the inverse solution at the original dimension *n* by solving the inversion in the directions that best explain its information content. Bousserez and Henze (2018) avoided explicit construction of the Jacobian matrix by using randomized methods to estimate these directions. However, this approach is computationally tractable only for inverse systems where a small number of directions explain the information content.

Here we present two 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 the 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 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**

This section presents both 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 update method to approximate the Jacobian matrix using the specified reductions in dimension and rank (Sections 2.3 through 2.5). For the purposes 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 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 , the sensitivity of the posterior emissions estimate to the true state **x**. **A** can be calculated as or as

which expresses the dependence of the averaging kernel matrix on the forward model and both error covariance matrices (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 constrained 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 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 (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. Bousserez and Henze (2018) further 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. 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 all eigenvectors have a sufficiently large signal-to-noise ratio. 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 at native resolution. Here we present a two-step update method to decrease the computational cost of constructing the Jacobian matrix by reducing the dimension and rank of the 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, 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, 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.

An initial estimate of the native-resolution Jacobian matrix can be generated in our test case by assuming that methane emissions [kg s-1] produce local column concentration enhancements [ppb] dependent on local wind speed and turbulent diffusion. 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*, and are the molecular weights of dry air and methane, respectively, *g* is gravity, *U* is the local 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. The coefficients provide a simple mass-conserving representation of turbulent diffusion and reduce the sparsity of , which is necessary to efficiently generate an updated Jacobian matrix following the two proposed methods. 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 . Because the averaging kernel matrix is a function not only of the Jacobian matrix but of the prior and observational error covariance matrices (equation (4)), this initial estimate will reproduce the patterns of information content determined by the prior errors and observational density despite the crude representation of transport.

*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). We define information content by the averaging kernel sensitivities. An optimal multiscale grid will maximize the sensitivities of each state vector element, referred to here as the DOFS per cluster. To construct this grid, we consider first a state vector with a single element corresponding to the full domain. We introduce to the state vector the native-resolution grid cells with the highest averaging kernel sensitivities. For each new state vector, we construct the Jacobian matrix by calculating the sensitivity of the forward model to a perturbation in each state vector element. We then solve the inversion and calculate the DOFS per cluster. When the DOFS per cluster remains constant or decreases, we add clusters of two or more native-resolution grid cells to the state vector. We repeat this process, increasing cluster size, until all native-resolution grid cells are allocated to the new multiscale grid. Clusters are generated 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. Once the multiscale grid is constructed, the reduced-dimension Jacobian matrix can be constructed by running the forward model for perturbations of each state.

Here we use the same approach beginning with our initial estimate in a two-step update process that iteratively improves an initial estimate of the inverse system’s information content. We construct an initial multiscale grid using and generate 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 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, 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. 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.

*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 a 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 or equal to about 0.5 (equation (9)). We use a strict signal-to-noise threshold to account for the errors in the initial estimate of the information content. 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 since 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 approaches in an analytic Bayesian inversion of atmospheric methane columns observed by the GOSAT satellite over North America in July 2009. We use GOSAT rather than TROPOMI because the data is currently better validated (Buchwitz et al. 2015). It also 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* = 2,098, upper left panel of Figure 1) against which our reduced-dimension and reduced-rank methods can be compared. To demonstrate the applicability of the methods to inversions of TROPOMI data, which we anticipate will have higher information content than the sparse GOSAT data, we artificially increase the information content of our GOSAT inversion 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 all inversions. In the native-resolution inversion, this increases the DOFS from 40 to 216.

All inversions 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 that may have large errors that do not matter for the sake of our comparison. After constructing the native-resolution Jacobian matrix, we use it to conduct all subsequent model simulations in lieu of additional GEOS-Chem simulations.

Figure 2 (upper left panel) shows the native-resolution 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 variability in the distribution of prior methane emissions because we define prior errors relative to the prior emissions. The variability in observation density is driven by variability in sampling and retrieval success, which depends in part on surface reflectivity and the presence of a clear sky.

Figure 2 (upper right panel) also shows the initial estimate of averaging kernel sensitivities (**)** derived from the initial estimate of the Jacobian matrix following the approach 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 of the strong dependence on the prior error standard deviation and observation density.

### This paragraph has not been edited ### We first construct the Jacobian on a multiscale grid. We find that half of the DOFS in the initial estimate are contained in 100 native resolution grid boxes, 90% in 400 grid boxes, and 99% in 900 grid boxes. We construct a multiscale grid on the basis of this distribution of information content following Section 2.4 with a K-means clustering aggregation scheme. Our initial grid contains 98 native resolution grid boxes (50% of DOFS), 100 clusters of ~3 native resolution grid boxes (40% of DOFS), 100 clusters of ~5 (9% of DOFS), and 150 clusters of ~8 (1% of DOFS). This reduces the total dimension of the state vector to 449 with the largest cluster containing ~15 native resolution grid boxes, where the variations in cluster size result from the K-means clustering algorithm. We construct a multiscale Jacobian on this multiscale grid using the native resolution Jacobian matrix as the forward model. To account for the 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 553 and the corresponding multiscale Jacobian requires ~550 model runs in total.

### This paragraph has not been edited ### Figure 3 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. Decreasing the dimension of the state vector increases the number of observations constraining each state vector element, decreasing the observational error covariance. As a result, we scale down the regularization factor introduced in equation (10) by the ratio of the new state vector dimension (553) to the native state vector dimension (2,098). Figure 4 shows the posterior emission scaling factors (top) and averaging kernel sensitivities (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 only 98 DOFS compared to the native resolution 216 DOFS. Reducing the state vector dimension necessarily reduces the DOFS of the inversion since the state vector dimension defines the upper limit on the DOFS. When normalized for the number of grid cells optimized, the multiscale solution generates 0.18 DOFS per cell compared to the native resolution 0.10 DOFS per cell, reflecting the consolidation of information content.

We next use the initial estimate of the averaging kernel matrix to construct a reduced-rank Jacobian matrix. We first require that the signal-to-noise ratio of all patterns be greater than 1.25, corresponding to *k* = 125. We perturb the leading 125 eigenvectors in the forward model and construct the reduced-rank Jacobian matrix as described in Section 2.5 using the native resolution Jacobian matrix as the forward model. The eigenvalue spectrum given by the updated information content (purple) has a discontinuity at *n* = 125 as expected for a rank 125 system. Because the additional information generated by 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 use the initial eigenvalue spectrum to determine *k* for the second update. We require next that the improved eigenvectors capture 97.5% of the information content, corresponding to *k* = 417. The resulting Jacobian matrix has rank ~417 and required 542 model runs to construct.

We solve the inversion using the resulting reduced-rank Jacobian matrix. Figure 4 (right column) shows the distribution of the posterior scaling factors (top) and averaging kernel sensitivities (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 of 1) elsewhere. The averaging kernel accurately captures these patterns of information content. Due the limited extent of the optimized grid cells, the resulting 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 by excluding grid cells where the averaging kernel sensitivities are less than 0.01, we find 152 DOFS across 711 grid cells, or 0.21 DOFS per grid cell. The reduced rank solution generates only ~70% of the native resolution DOFS despite constructing the reduced-rank Jacobian on the basis of eigenvectors that capture 97.5% of the information content; the discrepancy results from errors in the characterization of the tailing eigenvectors.

Figure 5 shows the elements of the final reduced-rank Jacobian matrix approximation (upper left) and the resulting posterior solution, including scaling factors (upper right), variance (lower left), and averaging kernel sensitivities (lower right), subjected to the 0.01 averaging kernel sensitivity threshold and plotted against the corresponding native-resolution values. Decreasing the threshold will increase the accuracy of the posterior solution relative to the truth but will decrease the number of constrained grid cells. Increasing the threshold will increase the number of grid boxes where the posterior scaling factor is not optimized and defaults to the prior value. Figure 5 shows that the filtered reduced-rank Jacobian matrix correlates well with the native-resolution values (r2 = 0.93). The filtered reduced-rank posterior variance is higher than the native-resolution posterior variance, reflecting the error introduced by discarding information content. Similarly, the filtered reduced-rank averaging kernel sensitivities underestimate the native-resolution values. 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 6 summarizes these results. We show 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**

We proposed two methods for decreasing the computational cost of constructing Jacobian matrices for analytic inversions of satellite observations of atmospheric constituents and demonstrated these methods in an inversion of GOSAT column methane observations for July 2009 over the North American domain. We found that both methods are capable of substantially reducing the computational cost of constructing the Jacobian matrix while accurately constraining posterior emissions where information content is highest. In an era where satellites provide increasingly high-resolution, dense observations of atmospheric constituents, analytic inversions of linear systems can improve constraints on emission sources and characterize the associated errors and information content. Previously, the analytic approach was limited by the computational cost of constructing the Jacobian matrix, which required a forward model simulation for every state vector element constrained by the inversion. Our methods allow analytic inversion of satellite observations at high-resolution with many fewer forward model simulations.

We considered inverse problems where the forward model is linear with respect to the state vector elements. In this case, there exists an analytic solution for the minimum of a Bayesian cost function that is a function of the prior estimate and its errors, the observations and their errors, and the forward model, as expressed by the Jacobian matrix. The analytic solution yields closed-form characterization of the posterior emissions, their errors, and their information content. In systems with a known Jacobian matrix, optimal reductions in the dimension or rank of the inversion maximize the information content. In systems without a known Jacobian matrix, an initial estimate of the Jacobian matrix can be constructed at no computational cost using a mass balance approach. This initial estimate together with the prior and observational errors can accurately quantify the fine structure of information content in the inversion. We use this initial estimate of information content to define optimal reductions in the dimension and rank of the state vector, decreasing the computational cost of constructing the Jacobian matrix.

**Paragraph 3: Reduced dimension construction**

Our dimension reduction method uses the initial estimate of the averaging kernel

**Paragraph 4: Reduced rank construction**

Our rank reduction method uses the initial estimate of the averaging kernel matrix

**Paragraph 5: Results summary (reduced dimension)**

**Paragraph 6: Results summary (reduced rank)**

**Paragraph 7: Next steps**

**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.