**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 understanding of surface emissions through inverse analyses. Variational inverse methods can optimize surface emissions globally at any resolution but do not readily quantify the error and information content of the posterior solution. In fact, the information content of the satellite data may be orders of magnitude 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 method 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 high resolution because it involves perturbing each emission element, typically individual grid cells, in the atmospheric transport model used as the forward model for the inversion. We propose and analyze a reduced-dimension and a reduced-rank method to construct the Jacobian matrix at greatly decreased computational cost while retaining information content. Both methods begin from an initial native-resolution estimate of the Jacobian matrix constructed at no computational cost by making simple transport assumptions. On the basis of this estimate, the reduced-dimension method constructs a Jacobian matrix on a multiscale grid that maintains high resolution in areas with high information content and aggregates grid cells elsewhere. The reduced-rank method constructs the Jacobian matrix at native resolution by perturbing the leading patterns of information content. The resulting matrix optimizes emissions only in areas spanned by these leading patterns and defaults to the prior estimate elsewhere. We demonstrate both methods in an analytic Bayesian inversion of GOSAT methane data with augmented information content over North America in July 2009. We show that both methods reproduce the results of the native-resolution inversion while achieving a factor of four improvement in computational performance. The reduced-dimension method produces an exact solution on the multiscale grid while the reduced-rank method provides a higher resolution solution.

**1. Introduction**

Satellite observations of atmospheric composition provide a powerful resource to improve our knowledge of emissions (Streets et al. 2013). They are, however, subject to large errors, both from the measurements themselves and from the inverse analyses used to infer emissions from the observations (Brasseur and Jacob 2017). Conducting inverse analyses of satellite data to quantify emissions at high resolution 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 optimize the information content of the observations while minimizing computational cost and providing full error statistics.

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 individual observations and the elements of **x** are the emissions optimized by the inversion, 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 inform emission estimates (Jacob et al. 2016). This was first shown with data from the SCIAMACHY satellite instrument (2003-2012) with 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) with 10-km diameter pixels approximately 250 km apart along- and cross-track (Monteil et al. 2013; Alexe et al. 2015; Turner et al. 2015; Maasakkers et al. 2019). The Tropospheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 precursor satellite, 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 construction of the Jacobian matrix is computationally tractable. Bocquet et al. (2011) defined a method to find an optimal multiscale grid by considering 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 the 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 the inversion in the directions of highest information content. Spantini et al. (2015) assumed knowledge of the Jacobian matrix. Bousserez and Henze (2018) avoided explicit construction of the Jacobian matrix by estimating the directions of highest information content, but their approach is effective only if 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 method 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* forward model simulations where *k* is selected by the user based on the information content of the inverse system and the available computational resources. We demonstrate both methods in a 1-month inversion of satellite data.

**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 an initially unknown 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.

*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 constructed by finite difference (see Introduction) and **c** is a constant, then an analytic solution to the cost function minimum exists that yields both 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 represents the sensitivity of the posterior emissions 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 both error covariance matrices **S**Aand **S**O (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).

*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, although the eigenvector corrections generated in the non-discrete, reduced-dimension space (lower right panel) would be difficult to interpret.

We wish 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. Alternatively, we can 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 and represents the signal-to-noise ratio of each eigenvector (Rodgers 2000).

*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**. In the case of high-resolution inversions of chemical transport models, the *n* + 1 forward model simulations needed to construct **K** may be prohibitively expensive. Here we present a two-step approach to construct a reduced-dimension or reduced-rank Jacobian matrix at substantially lower computational cost. We start from a low-cost, native-resolution estimate (below) and calculate the corresponding averaging kernel matrix . In the reduced-dimension method, we use to construct a multiscale grid that maintains resolution in the 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 reduced-rank method, we construct on the basis of the *k* dominant eigenvectors of 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 methods, 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.

The reduced-dimension and reduced-rank methods rely on accurately characterizing the dominant patterns of information content in the initial estimate of the averaging kernel matrix . Because the averaging kernel matrix depends strongly on the specified prior and observational error covariance matrices and (equation (4)), can provide a good approximation of **A** even if the initial estimate of the Jacobian matrix is crude. In our demonstration case, we generate a native-resolution Jacobian matrix at no cost by assuming that methane emissions [kg m-2 s-1] produce local column mixing ratio enhancements [mol mol-1] dependent on local wind speed and parameterized turbulent diffusion, both assumed constant over the inversion domain. 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 due to gravity, *U* is the local wind speed taken here as 5 km h-1, *W* is the square root of the grid cell area, and *p* is the surface pressure. The coefficients provide a crude mass-conserving representation of turbulent diffusion that reduce the sparsity of . We define for the grid cell where the observation is found with the remaining mass distributed over the three concentric rings surrounding that cell as 0.3, 0.2, and 0.1.

*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, we first define the state vector as a single element that encompasses the inversion domain. We then add the native-resolution grid cells with the highest averaging kernel sensitivities to the state vector one-by-one. For each new element , we calculate the corresponding Jacobian matrix columnn and the resulting increase in DOFS. When the DOFS level off, we add instead clusters of two or more native-resolution grid cells and repeat this procedure. Clusters can be generated by K-means clustering, which aggregates 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.

We apply this approach beginning with our initial estimate in a two-step update that iteratively improves the multiscale grid. The information content for the initial multiscale grid is given by , which identifies the grid cells with the highest sensitivities even given the crude estimate of the Jacobian matrix (Section 2.3). We then construct a multiscale grid and compute the corresponding reduced-dimension Jacobian matrix , introducing 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 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, convergence is rapid and we find no need for further iteration. The analytic inversion can then be solved exactly on the multiscale grid using .

*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 normalized 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 the posterior 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 use the stricter signal-to-noise criterion to account for the errors in the initial estimate of the information content. We calculate the forward model response to each of the eigenvectors using equation (9) and transform the resulting reduced-dimension Jacobian to the full-dimension state space with . 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 . 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.

**3. Results and Discussion**

We demonstrate 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 data 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 demonstrate the applicability of the methods to higher-information observing systems 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 , increasing the native-resolution DOFS from 40 to 216. Because of noise in the GOSAT data, this artificial increase in information content results in an overfit that is inconsequential for our demonstration. All inversions solve for posterior scaling factors defined relative to the prior emissions.

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. The 2098 1º x 1.25º grid cells constitute our native-resolution state vector. 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). The GOSAT observations are from the University of Leicester version 7 CO2 proxy retrieval over land (Parker et al. 2011, 2015) for July 2009. Prior emissions and error covariances are from Maasakkers et al. (2019). This demonstration case is sufficiently short that the native-resolution Jacobian matrix can be explicitly computed with 2099 model runs. After constructing the native-resolution Jacobian matrix, we use 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 (cf. equation (4)), the patterns in the sensitivities are largely driven by the specified error covariance matrices and . In our demonstration, is a diagonal matrix of relative errors on prior emission estimates and is a diagonal matrix of absolute errors on the observations (Maasakkers et al. 2019). As a result, the patterns of the averaging kernel sensitivities given by **A** largely reflect the patterns of the absolute prior errors (lower left panel) and the observation density (lower right panel). Errors on the prior emissions estimate are largest for wetlands along the southeastern coast of the U.S. (Bloom et al. 2017). The variability in the GOSAT observation density is driven by sampling frequency and retrieval success.

Figure 2 (upper right panel) also shows the initial estimate of averaging kernel sensitivities of 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 of the strong dependence on the prior error standard deviation and the observation density.

We aim to reduce the number of forward model runs needed to construct the Jacobian matrix by 75% relative to the native-resolution inversion, from 2099 to 530 simulations. We first apply the reduced-dimension method to construct a reduced-dimension Jacobian matrix on a multiscale grid. We follow the methodology described in Section 2.4, adding 50 state vector elements to the grid at a time to reproduce a parallel computing environment. The resulting multiscale grid and reduced-dimension Jacobian matrix constrains 359 clusters and required 470 model simulations. We disaggregate 16 clusters with a sensitivity increase greater than 0.4, adding 64 native-resolution grid cells. The resulting multiscale grid is shown in the upper right panel of Figure 1. It has dimension 423 and the corresponding reduced-dimension Jacobian matrix required 534 forward model simulations. The grid has 199 native-resolution grid cells and clusters of between 2 and 49 grid cells. The grid maintains native resolution where information content is highest (upper left panel of Figure 2). Grid cells are aggregated elsewhere.

Figure 3 shows the averaging kernel sensitivities (top) and posterior emission scaling factors relative to the prior estimate (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 reduced-dimension averaging kernel sensitivities, which are more uniform than the native-resolution values. The reduced-dimension posterior scaling factors exhibit less variability than the native-resolution solution, which exhibits checkerboard patterns resulting in part from the overfit of the posterior solution to observational noise. The posterior scaling factors agree on regional scales.

We next apply the reduced-rank method (Section 2.5) to construct a reduced-rank approximation of the 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 2.5. This yields = 74 eigenvectors, which account for 37% of the initial-estimate DOFS. We perturb these eigenvectors in the forward model and construct the reduced-rank Jacobian matrix. We then calculate the averaging kernel matrix and its dominant eigenvectors, using the initial eigenvalue spectrum to determine = 462 by requiring that the improved eigenvectors capture 98.5% of the information content defined by . The resulting Jacobian matrix has rank 462 and required 537 forward model simulations, a 75% reduction from the 2099 simulations required for the native-resolution solution. We solve the inversion with and find 155 DOFS compared to the 216 DOFS generated in the native-resolution inversion (results discussed below).

The reduced-rank DOFS have limited sensitivity to the signal-to-noise ratio and information content fraction thresholds chosen in the first and second updates. The lower panel of figure 4 shows the DOFS of the reduced-rank solution as a function of the number of first- and second-update forward model runs. Considering all possible partitions of 537 total model runs (dashed line), we find that the update scheme used above (starred) maximizes the DOFS. Within this set of update configurations, using a signal-to-noise ratio threshold of 1 or 4 instead of 2.5 (dots), decreases the reduced-rank DOFS by 2% and 3% to 150 and 152, respectively. Considering update configurations with a different number of total model runs, we find a weak dependence of DOFS on the number of model runs conducted in the first update after the first ~50 simulations, suggesting that the initial estimate of the Jacobian matrix accurately characterizes only the most significant patterns of information content, a conclusion consistent with our previous analysis (cf. Figure 2). Indeed, the large errors in the tailing eigenvectors of the initial averaging kernel estimate result in decreasing DOFS as the number of first update model runs increases in many cases. We find that applying a signal-to-noise ratio threshold of 2.5 in the first update maximizes the DOFS regardless of the number of model runs conducted in the second update. The top panel of figure 4 shows the Among these optimal model configurations, we find that the reduced-rank DOFS are only moderately sensitive to the total number of forward model runs (top panel). After only 275 simulations, 87% fewer than the native-resolution inversion and 51% fewer than our reduced-rank inversion, the optimal reduced-rank inversion generates 108 DOFS, half of the native-resolution DOFS and 70% of our reduced-rank DOFS. After 1050 model simulations, the optimal reduced-rank inversion generates 181 DOFS, generating 84% of the native-resolution DOFS at half the computational cost.

We solve the inversion (equations (2) – (4)) using and compare the posterior to the native-resolution solution visually and quantitatively. Figure 3 (right column) shows the distribution of the reduced-rank averaging kernel sensitivities (top) and posterior scaling factors (bottom) compared to the native-resolution inversion (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 one) elsewhere. ###Visual inspection suggests that the reduced-rank averaging kernel sensitivities accurately capture the largest native-resolution values. In the high-sensitivity grid cells, the reduced-rank scaling factors appear to reproduce the native-resolution solution. Elsewhere, the reduced-rank averaging kernel sensitivities go to zero and the posterior scaling factors go to one. Due the limited extent of the optimized grid cells, the resulting reduced-rank DOFS (153) are lower than native resolution DOFS (216). 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 679 grid cells. 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 98% of the information content; the discrepancy results from errors in the characterization of the tailing eigenvectors.

Figure 7 quantitatively confirms that the reduced-rank inversion reproduces the native-resolution solution. We show the elements of the final reduced-rank Jacobian matrix approximation (upper left) and the resulting posterior scaling factors (upper right), standard deviations (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. As suggested by Figure 5, the filtered reduced-rank averaging kernel sensitivities correlate well with the native-resolution values (R = 0.91) despite tending to underestimate the native-resolution values due to the loss of information content. The filtered reduced-rank posterior scaling factors and standard deviations correlate well with the native resolution values (R = 0.87 and R = 0.99, respectively). The filtered reduced-rank posterior variances tend to be larger than the native-resolution values, reflecting the error introduced by discarding information content. The posterior scaling factors generate the lowest correlation coefficient because of the propagation of errors from the posterior error covariance and Jacobian matrices (equation (2)).

We finally compare the reduced-dimension and reduced-rank methods. Both methods achieve our goal of a 75% reduction in the 2099 model runs needed to construct the native-resolution Jacobian matrix. The reduced-dimension method generates 42% fewer DOFS but 300% more DOFS per state vector element relative to the reduced-rank method. While the reduced-dimension method has lower resolution than the reduced-rank method, the solution is exact on the multiscale grid. It also provides better spatial coverage than the reduced-rank method, although the posterior scaling factors should only be interpreted where the averaging kernel sensitivities are sufficiently high. The reduced-rank method generates a much higher resolution solution where the averaging kernel sensitivities are large. However, the reduced-rank approximation of the Jacobian matrix introduces additional errors to the inversion. The increase in posterior error is reflected in the reduced-rank posterior error covariance. When deciding between the two methods, this trade-off between resolution and precision should be considered.

**4. Conclusions**

We proposed two computationally tractable methods to conduct analytic inversions to infer high-resolution emissions from satellite observations of atmospheric composition while maximizing information content. The methods decrease computational cost by exploiting the dominant patters of information content in the prior emissions estimate and observations to construct the Jacobian matrix, which represents the sensitivity of simulated atmospheric concentrations to emissions for a given forward model, in a reduced-dimension or reduced-rank emissions space instead of at native resolution. Both methods can be applied more generally to the problem of efficient numerical approximation of high-dimension Jacobian matrices.

We considered inverse problems where the forward model is linear with respect to the optimized state vector elements, in our case an emissions grid, and where the errors can be assumed to be normally distributed. In this case, there exists an analytic solution to a Bayesian cost function that yields closed-form characterization of the optimized (posterior) emissions, errors, and information content, given by the averaging kernel matrix. The sum of the diagonal elements of the averaging kernel matrix, called here sensitivities, yields the degrees of freedom for signal (DOFS), the number of pieces of information that can be independently quantified by the inverse system. The analytic solution requires explicit construction of the Jacobian matrix, which is typically constructed by finite difference, requiring a model simulation for every state vector element. When inferring emissions at high resolution, constructing this matrix can become computationally intractable.

The number of model runs needed to construct the Jacobian matrix can be decreased by reducing either the dimension or rank of the state vector space. An optimal dimension- or rank-reduction preserves information content. However, the averaging kernel matrix is a function of the Jacobian matrix. The methods proposed here use a two-step update to improve an initial estimate of the Jacobian matrix and the corresponding averaging kernel matrix. We suggested generating a low-cost initial estimate of the Jacobian matrix using a mass balance approach. Because the averaging kernel matrix has a strong dependence on the prior and observational error covariance matrices, this initial estimate can accurately quantify the fine structure of information content.

The reduced-dimension method uses the initial estimate of the averaging kernel matrix to build the Jacobian matrix on a multiscale grid that maintains native resolution where information content is highest and aggregates grid cells elsewhere. The method constructs the multiscale grid and corresponding Jacobian matrix element-by-element, introducing first the native-resolution grid cells with the highest averaging kernel sensitivities. For each added state vector element, the forward model response is calculated and the inversion solved. When the total DOFS stabilize, the process is repeated with clusters of grid cells generated by, for example, K-means clustering. When all native-resolution grid cells are allocated to the multiscale grid and the reduced-dimension Jacobian is constructed, the clusters where the forward model contributed the most information content are disaggregated, generating a second update of the multiscale grid and reduced-dimension Jacobian matrix. The information content from both the first and second update contains contributions from the prior emissions estimate, the observations and the forward model; no further iteration is needed.

The reduced-rank method constructs an approximation of the Jacobian matrix by calculating the linear relationship between emissions and observations in the forward model for the most important patterns of information content rather than individual or aggregate grid cells. These patterns are given by dimension-reducing and -restoring transformations that best preserve information content introduced by Bousserez and Henze (2018). The number of patterns used can be determined by requiring that the signal-to-noise ratio of each pattern or the fraction of the DOFS explained by the patterns exceed a given threshold. The initial estimate of the averaging kernel matrix is used to calculate those patterns, which are then perturbed in the forward model. The resulting forward model response is transformed back to the native-resolution dimension using the optimal dimension-restoring transformation, generating a reduced-rank approximation of the Jacobian matrix. This process is then repeated using the updated Jacobian matrix and corresponding averaging kernel matrix. As with the reduced-dimension method, we find rapid convergence and no need for further iteration.

We applied both methods in a demonstration inversion of GOSAT column methane observations for July 2009 at 1º x 1.25º resolution over the North American domain. We constructed a native-resolution, reduced-dimension, and reduced-rank Jacobian matrix, solved each inversion, and compared the results. We found that both methods were capable of reducing the computational cost of constructing the Jacobian matrix by 75% while accurately constraining posterior emissions where information content is highest. The reduced-dimension Jacobian solved the inversion exactly on a multiscale grid with dimension 423, including clusters of between 1 and 58 native-resolution grid boxes. While the reduced-dimension inversion produced fewer than half of the native-resolution DOFS, it generated twice the DOFS per state vector element. The reduced-dimension Jacobian solved the inversion accurately in the native-resolution grid boxes with the highest information content, defaulting to the prior emissions estimate elsewhere and preserving 70% of the native-resolution DOFS. Because of the loss of information content, the error of the optimized emissions also increases. The difference between the two methods is characterized by this trade-off between resolution and precision.

Satellite observations of atmospheric methane concentrations are providing an increasingly powerful resource to improve constraints on emission sources, beginning with the launch of TROPOMI in October 2017. High-resolution analytic inversions are needed to interpret these data, including a rigorous interpretation of the information content of the observations. Our methods make such analyses computationally tractable. They also provide a framework for similar inversions of other long-lived atmospheric tracers and for the more general problem of constructing high-dimensionality Jacobian matrices.

**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.

Bloom, A. Anthony, Kevin Bowman, Meemong Lee, Alexander J. Turner, Ronny Schroeder, John R. Worden, Richard Weidner, Kyle C. McDonald, and Daniel J. Jacob. 2017. “A Global Wetland Methane Emissions and Uncertainty Dataset for Atmospheric Chemical Transport Models.” *Geoscientific Model Development Discussions* 10: 2141–56. https://doi.org/10.5194/gmd-2016-224.

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.

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.

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.