Leveraging Uncertainty Quantification to Design Ocean Climate Observing Systems

Nora Loose[1,1] and Patrick Heimbach[2,2]

[1]University of Texas at Austin
[2]The University of Texas at Austin

November 30, 2022

Abstract

Ocean observations are expensive and difficult to collect. Designing effective ocean observing systems therefore warrants deliberate, quantitative strategies. We leverage adjoint modeling and Hessian uncertainty quantification (UQ) within the ECCO (Estimating the Circulation and Climate of the Ocean) framework to explore a new design strategy for ocean climate observing systems. Within this context, an observing system is optimal if it minimizes uncertainty in a set of investigator-defined quantities of interest (QoIs), such as oceanic transports or other key climate indices. We show that Hessian UQ unifies three design concepts. (1) An observing system reduces uncertainty in a target QoI most effectively when it is sensitive to the same dynamical controls as the QoI. The dynamical controls are exposed by the Hessian eigenvector patterns of the model-data misfit function. (2) Orthogonality of the Hessian eigenvectors rigorously accounts for redundancy between distinct members of the observing system. (3) The Hessian eigenvalues determine the overall effectiveness of the observing system, and are controlled by the sensitivity-to-noise ratio of the observational assets (analogous to the statistical signal-to-noise ratio). We illustrate Hessian UQ and its three underlying concepts in a North Atlantic case study. Sea surface temperature observations inform mainly local air-sea fluxes. In contrast, subsurface temperature observations reduce uncertainty over basin-wide scales, and can therefore inform transport QoIs at great distances. This research provides insight into the design of effective observing systems that maximally inform the target QoIs, while being complementary to the existing observational database.

Leveraging Uncertainty Quantification to Design Ocean Climate Observing Systems

Nora Loose[1], Patrick Heimbach[1,2,3]

[1]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX, USA
[2]Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA
[3]Institute for Geophysics, The University of Texas at Austin, Austin, TX, USA

Key Points:

- We apply Hessian uncertainty quantification (UQ) to the global ocean state estimate ECCO, and explore its use for observing system design.
- Hessian UQ elucidates oceanic teleconnections that communicate observational constraints over basin-scale distances.
- Going beyond previous adjoint ocean modeling techniques, Hessian UQ rigorously assesses redundancy and optimality of observing systems.

Corresponding author: Nora Loose, `nora.loose@colorado.edu`

Abstract

Ocean observations are expensive and difficult to collect. Designing effective ocean observing systems therefore warrants deliberate, quantitative strategies. We leverage adjoint modeling and Hessian uncertainty quantification (UQ) within the ECCO (Estimating the Circulation and Climate of the Ocean) framework to explore a new design strategy for ocean climate observing systems. Within this context, an observing system is optimal if it minimizes uncertainty in a set of investigator-defined quantities of interest (QoIs), such as oceanic transports or other key climate indices. We show that Hessian UQ unifies three design concepts. (1) An observing system reduces uncertainty in a target QoI most effectively when it is sensitive to the same dynamical controls as the QoI. The dynamical controls are exposed by the Hessian eigenvector patterns of the model-data misfit function. (2) Orthogonality of the Hessian eigenvectors rigorously accounts for redundancy between distinct members of the observing system. (3) The Hessian eigenvalues determine the overall effectiveness of the observing system, and are controlled by the sensitivity-to-noise ratio of the observational assets (analogous to the statistical signal-to-noise ratio). We illustrate Hessian UQ and its three underlying concepts in a North Atlantic case study. Sea surface temperature observations inform mainly local air-sea fluxes. In contrast, subsurface temperature observations reduce uncertainty over basin-wide scales, and can therefore inform transport QoIs at great distances. This research provides insight into the design of effective observing systems that maximally inform the target QoIs, while being complementary to the existing observational database.

Plain Language Summary

Ocean observing faces multiple challenges: high instrument cost, difficult deployment logistics via ships, harsh environments, and the necessity to sustain observations over long periods of time. Since oceanographers cannot measure the ocean everywhere and at all times, they have to carefully choose the location of their instruments. In an ideal scenario, measurements from a small number of instruments provide maximum information about important ocean metrics, such as poleward ocean heat transport or regional heat content. This paper presents a new method for planning optimal instrument configurations, by combining computer simulations of the global ocean with the mathematics of uncertainty quantification (UQ). As an example, we show that North Atlantic temperature measurements taken below the ocean surface do not only tell us about the ocean properties at the instrument locations themselves, but reduce uncertainty in regions hundreds to thousands of kilometers away. We can therefore use existing ocean observations to extract more information about the ocean than previously appreciated. Our method helps to plan informative observing networks that are complementary to the existing observational database.

1 Introduction

Sustaining long-term ocean observations to develop climate-quality observational records is crucial for understanding the ocean's role in climate and for evaluating climate model simulations (National Academies of Sciences, Engineering, and Medicine, 2017). Yet, ocean observing faces multiple challenges: complex deployment operations in frequently rough weather (or ice) conditions, limited instrument lifetime due to corrosive and high-pressure environments, and the necessity of adequate spatial and temporal sampling. The high cost and logistical challenges call for deliberate, quantitative approaches. Here, we leverage adjoint modeling and Hessian uncertainty quantification (UQ) within the ECCO (Estimating the Circulation and Climate of the Ocean) framework to explore a new design strategy for ocean climate observing systems. This approach has two distinguishing features, which, taken together, foster collaboration and system co-design within the oceanographic community. First, it gives insights into the physical mechanisms that govern optimal design strategies; and second, it quantitatively assesses redundancy and optimality of an (existing or future) observing system.

To place our technique into context, we briefly recall existing formal approaches to observing system design. Observing System [Simulation] Experiments (OS[S]Es, Fujii et al., 2019) are the most common computational tools in oceanography to support observing system design (e.g., Balmaseda et al., 2007; Gasparin et al., 2019; Griffa et al., 2006; Halliwell et al., 2017). OSEs are limited to evaluate *existing* observing systems, whereas OSSEs can test the skill of proposed *future* observing systems. The design strategy to be tested in an OSSE has to be specified by the investigator. Once a region is targeted for monitoring, the proposed observing system design (to be tested in the OSSE) is typically guided by best available knowledge of both local hydrographic properties and local dynamical balances (Hirschi et al., 2003; Li et al., 2017; Perez et al., 2011). An example are the Atlantic trans-basin mooring arrays OSNAP (Li et al., 2017), RAPID (Hirschi et al., 2003), and SAMBA (Perez et al., 2011), which target monitoring of the meridional overturning circulation. Key components of each design are western and eastern boundary moorings, which allow geostrophic transport estimates across each trans-basin section. Although these local considerations support practical local monitoring, it is possible that similar constraints could be obtained elsewhere, perhaps with an instrument configuration more easily sustained, at reduced cost, or less susceptible to noise. This opportunity arises from the appreciation that observed variability at any given location is rarely a purely instantaneous response to local forcing. Instead, it is the superposition of phenomena originating in distant regions and at distinct times, communicated through the ocean by advection, diffusion and wave propagation (Heimbach et al., 2011).

Exploring the possibility of remote constraint is essential for truly optimal observing system design. Adjoint models have proven valuable for fully mapping the local and remote origins and pathways of variability in targeted quantities, e.g., meridional overturning at given latitudes (Heimbach et al., 2011; Köhl, 2005; Pillar et al., 2016; Smith & Heimbach, 2019). Exploiting the rich information exposed by an adjoint model, a number of adjoint modeling techniques have previously been used to inform ocean observing system design, for example adjoint sensitivity (Heimbach et al., 2011; Masuda et al., 2010), observation sensitivity (Köhl & Stammer, 2004; Moore et al., 2011), and singular vectors (Fujii et al., 2008; Zanna et al., 2012). Despite giving valuable insight into where observations may be useful, none of these latter techniques provide a measure of redundancy versus complementarity, nor of optimality of an observing system. These obstacles are overcome by Hessian UQ: an adjoint-based technique embedded in a variational data assimilation system. The Hessian matrix (composed of second derivatives) of the cost function J captures the curvature of J with respect to the control variables, and allows one to calculate how much uncertainty is reduced with any changes applied to the observing system (Thacker, 1989). In contrast to the previous adjoint modeling techniques named above, Hessian UQ accounts for data redundancy. It also provides a measure of optimality: the more uncertainty an observing system reduces in a defined target quantity (on a scale of 0% to 100%), the closer it is considered to being optimal for the defined target.

Hessian UQ has been routinely applied in numerical weather prediction (NWP, Leutbecher, 2003) and, more broadly, in computational science and engineering (CSE, Bui-Thanh et al., 2012), but it has only seen limited use in the oceanographic community. Previous studies have applied Hessian UQ after severely reducing the dimension of the space of uncertain parameters in an ad-hoc manner (Kaminski et al., 2015, 2018), or in the dual form of 'representers' (Bennett, 1985; Moore et al., 2017; Zhang et al., 2010). These examples have focused on regional ocean settings and on daily to monthly time scales. In this study, we take a step toward fully exploiting Hessian UQ to design ocean observing systems that are targeted at climate monitoring in a global context. To this aim, we apply Hessian UQ within the global ocean state estimation framework of the Estimating the Circulation and Climate of the Ocean (ECCO) consortium (Heimbach et al., 2019), and elucidate oceanic teleconnections that communicate observational constraints over basin-scale distances and monthly to interannual time scales.

In ocean climate research, the goal of an observing system is usually to accurately estimate certain quantities of interest (QoIs): forecasts or climate indices that are difficult or impossible to observe directly. Examples of QoIs include transports across certain oceanographic passages, ocean heat content near the polar ice sheets, regional sea level anomalies, or future sea-ice extent. We therefore focus on the information that an observing system contains about a given QoI, here referred to as the observing system's 'proxy potential' for the QoI on a scale of 0% to 100% (Loose et al., 2020). Proxy potential is defined by way of Hessian UQ, as the reduction in QoI uncertainty that would be achieved if the observing system was added to the ocean state estimate. Importantly, proxy potential can be assessed not only for existing but also for future observing systems, because it does not require the actual measurement values of the observations (only their locations, times, types, and uncertainties).

Loose et al. (2020) provided interpretations of Hessian UQ and proxy potential for idealized cases, in which an observing 'system' consists of only a single and noise-free observation. Then, the observation's proxy potential for a QoI reflects the degree to which adjustment mechanisms are shared between the observation and QoI. In this simple case, proxy potential can be understood as the dynamical analogue of statistical correlation (squared) between observation and QoI, with the important distinction that proxy potential accounts only for covariability that has dynamical underpinnings. The goal of this study is to leverage Hessian UQ to generalize the notion of proxy potential introduced by Loose et al. (2020) in three important ways (section 2): first, by extending this concept from a single observational asset to full observing systems; second, by quantifying observational redundancy versus complementarity; and third, by accounting for observational noise.

We illustrate the concepts of Hessian UQ and proxy potential in a North Atlantic case study (section 3). To provide a clear understanding of Hessian UQ, our case study focuses on observing systems that are comprised of only a few observations. We then discuss how our approach and the dynamical insights obtained generalize to the design of full-fledged observing systems, including thousands to millions of observations (section 4).

2 Uncertainty Quantification and Proxy Potential

2.1 Ocean state estimation

Ocean state estimation optimally fits an ocean general circulation model (GCM) to the available observations in a dynamically and kinematically consistent way. For this, one solves an inverse problem: given an observing system (gray box, Fig. 1), one adjusts the control vector $\mathbf{u} = (u_1, \ldots, u_N)^T$ (green box, Fig. 1), such as to minimize the scalar cost function

$$J(\mathbf{u}) = \underbrace{\frac{1}{2}(\mathbf{y} - \mathbf{Obs}(\mathbf{u}))^T \, \mathbf{R}^{-1} \, (\mathbf{y} - \mathbf{Obs}(\mathbf{u}))}_{J_{\text{misfit}}(\mathbf{u})} + \underbrace{\frac{1}{2}(\mathbf{u} - \mathbf{u}_0)^T \, \mathbf{B}^{-1} \, (\mathbf{u} - \mathbf{u}_0)}_{J_{\text{prior}}(\mathbf{u})}. \qquad (1)$$

The control variables $u_1, \ldots, u_N$ (i.e., the elements of the control vector $\mathbf{u}$) are the uncertain input variables of the model, and consist not only of initial conditions (as common in NWP), but also of atmospheric forcing variables and uncertain model parameters (green box, Fig. 1). The function $J_{\text{misfit}}(\mathbf{u})$ measures the misfit between the vector of actual observations, $\mathbf{y} = (y_1, \ldots, y_M)^T$ (gray box, Fig. 1), and the vector of simulated observations, $\mathbf{Obs}(\mathbf{u}) = (\text{Obs}_1(\mathbf{u}), \ldots, \text{Obs}_M(\mathbf{u}))^T$ (pink box, Fig. 1), given the input variables $\mathbf{u}$. The function $J_{\text{prior}}(\mathbf{u})$ penalizes deviations from a first-guess $\mathbf{u}_0$ of uncertain inputs. The $M \times M$ matrix $\mathbf{R}$ and $N \times N$ matrix $\mathbf{B}$ are chosen error covari-

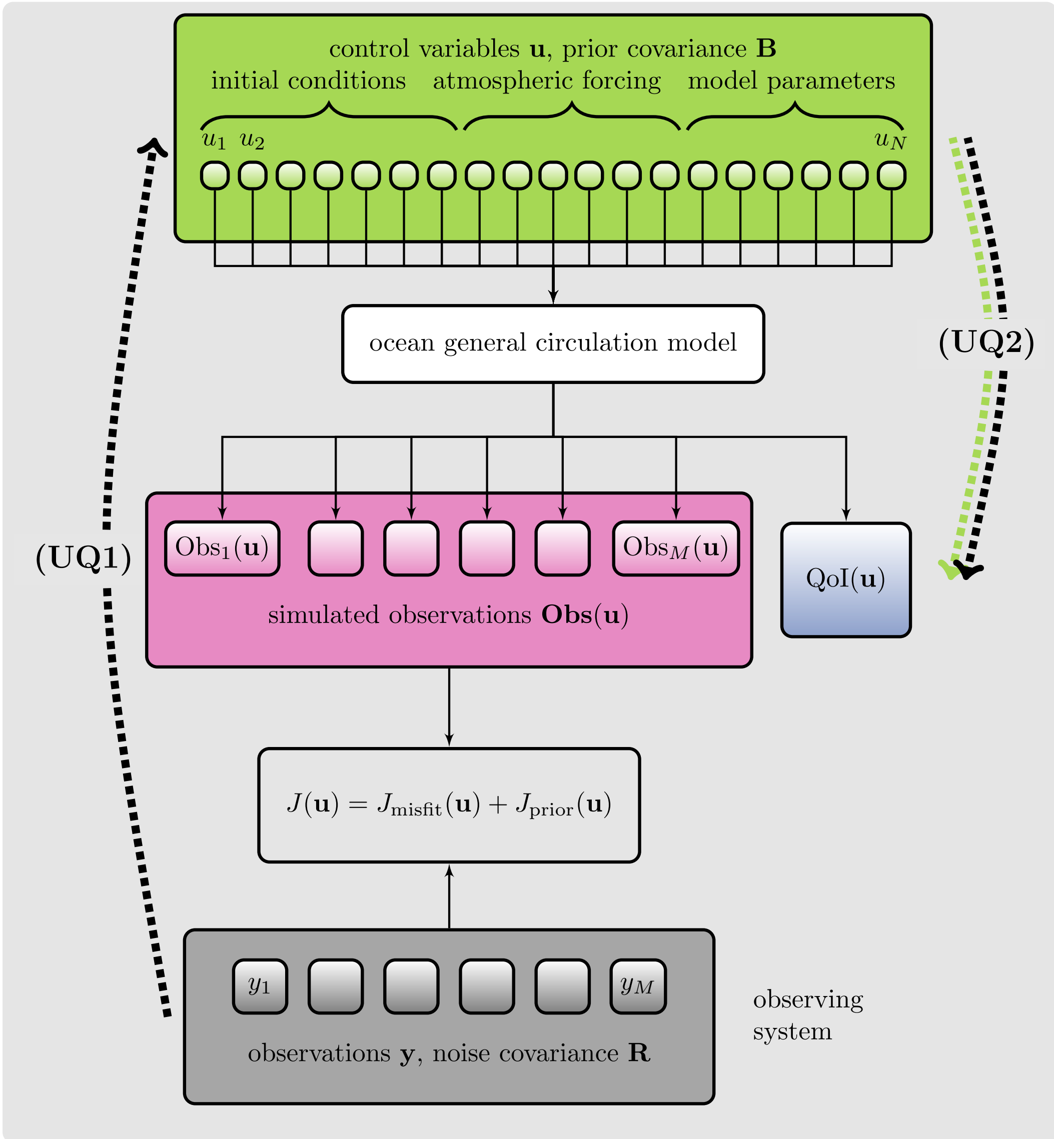

Figure 1. Workflow for Hessian uncertainty quantification (UQ) in ocean state estimation. Starting from an observing system (gray box), inverse uncertainty propagation along path (UQ1) reduces the uncertainty in the control variables (green box), see section 2.2. A subsequent forward uncertainty propagation along path (UQ2) reduces the uncertainty in a chosen quantity of interest (QoI, purple box), see section 2.3. Green and black arrows indicate propagation of prior and posterior uncertainty, respectively. The degree to which the observing system reduces uncertainty in the QoI, via a composite uncertainty propagation along paths (UQ1) and (UQ2), is referred to as the observing system's proxy potential for the QoI (section 2.4).

ances, spelling out the assumption that observational noise and prior uncertainties follow the Gaussian distributions $\mathcal{N}(\mathbf{0}, \mathbf{R})$ and $\mathcal{N}(\mathbf{u}_0, \mathbf{B})$, respectively (Tarantola, 2005).

The solution of the inverse problem is the minimizer of the cost function, $\mathbf{u}_{\min} = \min_{\mathbf{u}} J$; that is, a choice of control variables. The ocean state estimate itself is obtained by running the GCM with inputs $\mathbf{u}_{\min}$.

2.2 Inverse uncertainty propagation

To quantify uncertainties in the solution $\mathbf{u}_{\min}$ of the inverse problem, one propagates observational information and uncertainty along path (UQ1) (Fig. 1). This inverse uncertainty propagation results in the posterior probability distribution of the control variables, given the observations. In practice, it is not feasible to compute the full posterior probability distribution, nor to map this distribution onto the full ocean state space. We therefore need to appeal to approximation methods.

The posterior probability distribution of the control variables can be approximated by the Gaussian $\mathcal{N}(\mathbf{u}_{\min}, \mathbf{P})$, with $N \times N$ covariance matrix $\mathbf{P}$ equal to

$$\mathbf{P} = \mathbf{B} - \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1} \left(\mathbf{B}^{1/2}\,\mathbf{v}_i\right)\left(\mathbf{B}^{1/2}\,\mathbf{v}_i\right)^T. \tag{2}$$

Here, $\{\mathbf{v}_i, \lambda_i\}_{i=1}^{M'}$ is the set of orthonormal eigenvectors $\mathbf{v}_i$ with associated non-zero eigenvalues $\lambda_1 \geq \ldots \geq \lambda_{M'} > 0$ of the misfit Hessian:

$$\mathbf{H}_{\text{misfit}} = \mathbf{B}^{T/2}\,\mathbf{A}^T\,\mathbf{R}^{-1}\,\mathbf{A}\,\mathbf{B}^{1/2} = \sum_{i=1}^{M'} \lambda_i \mathbf{v}_i \mathbf{v}_i^T. \tag{3}$$

In eq. (3), the entries of the $M \times N$ matrix $\mathbf{A}$ are the sensitivities

$$\mathbf{A}_{i,j} = \frac{\partial(\text{Obs}_i)}{\partial u_j}, \tag{4}$$

evaluated at $\mathbf{u}_{\min}$. Furthermore, $\mathbf{B}^{1/2}$ denotes an $N \times N$ matrix which has an inverse, $\mathbf{B}^{-1/2}$, and satisfies $\mathbf{B}^{1/2}\,\mathbf{B}^{T/2} = \mathbf{B}$ (where $\mathbf{B}^{T/2}$ is the transpose of $\mathbf{B}^{1/2}$). The $N \times N$ matrix $\mathbf{H}_{\text{misfit}}$ is the linearized Hessian (or Gauss-Newton Hessian, Chen, 2011) of the rescaled model-data misfit term, $J_{\text{misfit}}(\tilde{\mathbf{u}})$ (eq. (1)). The rescaling is performed through the change of variables $\tilde{\mathbf{u}} = \mathbf{B}^{-1/2}\,\mathbf{u}$, and can be thought of as a nondimensionalization if $\mathbf{B}$ is diagonal. In summary, eq. (2) phrases the posterior uncertainty $\mathbf{P}$ as the prior uncertainty $\mathbf{B}$, reduced by any information $\{\mathbf{v}_i, \lambda_i\}$ obtained from the observations. Expression (2) has been known and used in the NWP and CSE communities for many years (see, e.g., Bui-Thanh et al., 2012; Leutbecher, 2003). A self-contained derivation is relegated to the supporting information (Text S1).

Next, we inspect the set $\{\mathbf{v}_i, \lambda_i\}$ in more detail as it fully characterizes the information obtained from the observations. The eigenvectors $\{\mathbf{v}_i\}_{i=1}^{M'}$ of the misfit Hessian (eq. (3)) are the data-informed directions within the control space. Along a data-informed direction $\mathbf{v}_i$, the function $J_{\text{misfit}}(\tilde{\mathbf{u}})$ has curvature $\lambda_i > 0$ (Fig. 2(a)). The eigenvalue λ_i captures the strength of the data constraint imposed on the control direction $\mathbf{v}_i$, with large λ_i corresponding to a strong observational constraint. The control directions along which $J_{\text{misfit}}(\tilde{\mathbf{u}})$ is not curved are not informed by the observations (Fig. 2(b)). Note that $M' \leq \min(M, N)$; that is, the number of independent data-informed directions, M', is less than or equal to the number of observations, M, and the number of control variables, N.

If an observing system consists of only a single observation ($M = 1$) with simulated counterpart $\text{Obs}_1(\mathbf{u}) = \text{Obs}(\mathbf{u})$ and observational noise variance $\mathbf{R} = \varepsilon^2 > 0$,

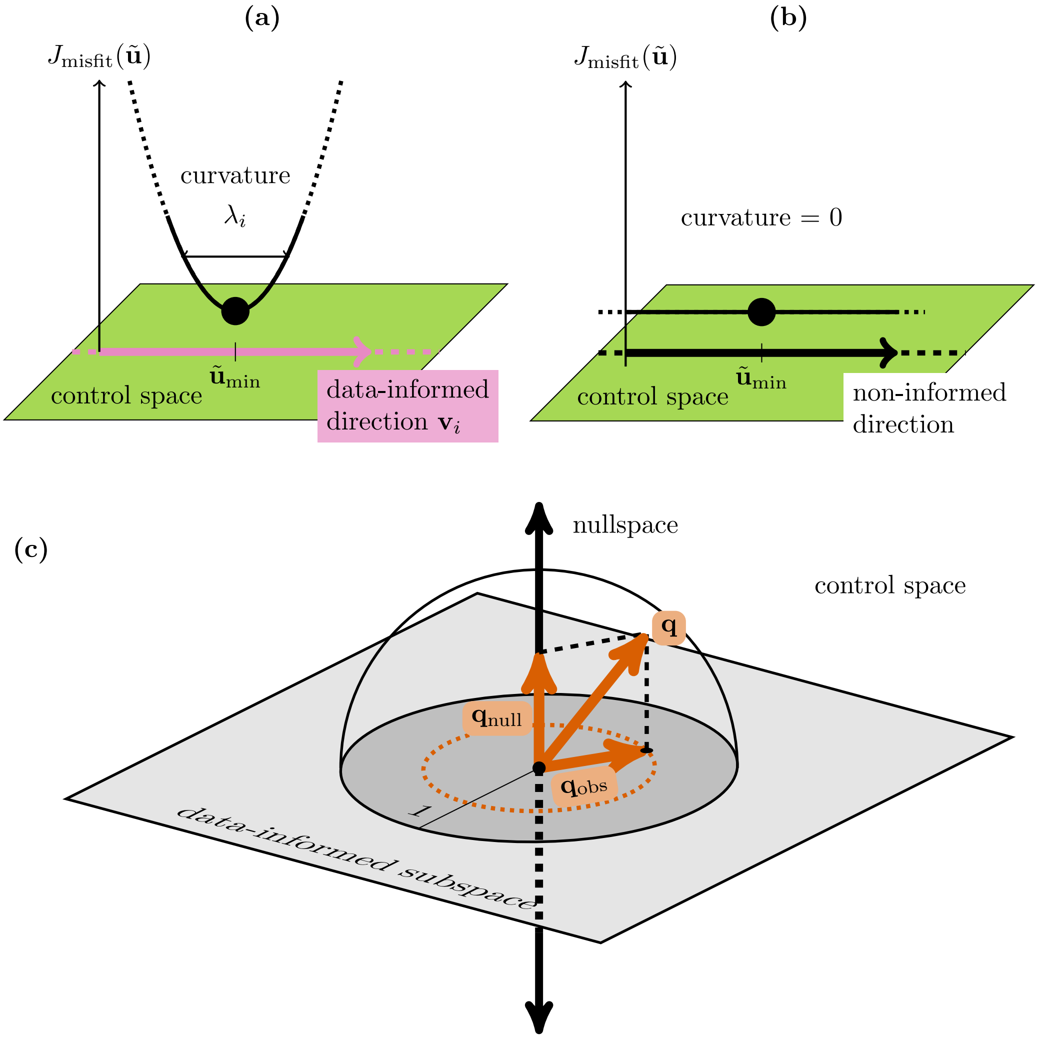

Figure 2. (a),(b) Curvature of the rescaled model-data misfit function, $J_{\text{misfit}}(\tilde{\mathbf{u}})$, at the cost function minimizer $\tilde{\mathbf{u}}_{\text{min}}$, along two directions in the control space: (a) the data-informed direction $\mathbf{v}_i$ (eq. (3)) and (b) a non-informed direction. (c) The direction of interest, $\mathbf{q}$ (eq. (10)), orthogonally decomposed into $\mathbf{q} = \mathbf{q}_{\text{obs}} + \mathbf{q}_{\text{null}}$. The data-informed component, $\mathbf{q}_{\text{obs}}$, is the projection of $\mathbf{q}$ onto the data-informed subspace. The component $\mathbf{q}_{\text{null}}$ lies in the nullspace, i.e., the subspace that is not informed by the data. Parts of the unit sphere of the control space are displayed in black, and $\mathbf{q}$ has unit length. The larger the radius of the orange dashed circle, defined by the length of $\mathbf{q}_{\text{obs}}$, the higher the dynamical proxy potential of the considered observing system for the QoI.

the misfit Hessian (eq. (3)) simplifies to $\mathbf{H}_{\text{misfit}} = \lambda_1 \, \mathbf{v}_1 \mathbf{v}_1^T$, with

$$\mathbf{v}_1 = (\sigma^{\mathbf{B}}_{\text{Obs}})^{-1} \left[\mathbf{B}^{T/2} \, \nabla_{\mathbf{u}} \text{Obs}\right] \in \mathbb{R}^N, \qquad \lambda_1 = \frac{(\sigma^{\mathbf{B}}_{\text{Obs}})^2}{\varepsilon^2} > 0. \tag{5}$$

Here, $\mathbb{R}^{\mathbb{N}}$ refers to the N-dimensional vector space of real numbers, and we denote $\nabla_{\mathbf{u}}\text{Obs} = (\partial(\text{Obs})/\partial u_1 \,, \ldots \,, \partial(\text{Obs})/\partial u_N)^T$, evaluated at $\mathbf{u}_{\min}$, and

$$\sigma^{\mathbf{B}}_{\text{Obs}} = \left\| \mathbf{B}^{T/2} \, \nabla_{\mathbf{u}} \text{Obs} \right\| > 0. \tag{6}$$

Put differently, the only data-informed direction is spanned by the prior-weighted sensitivity vector $\mathbf{B}^{T/2} \, \nabla_{\mathbf{u}}\text{Obs}$ (eq. (5)), where 'prior-weighting' is through multiplication by $\mathbf{B}^{T/2}$. Similarly, for an observing system with more than one observation ($M > 1$), the data-informed subspace of the control space is spanned by the M prior-weighted sensitivity vectors $\mathbf{B}^{T/2}\nabla_{\mathbf{u}}\text{Obs}_1 \,, \ldots \,, \mathbf{B}^{T/2}\nabla_{\mathbf{u}}\text{Obs}_M$. To obtain the eigenvectors of the misfit Hessian, one has to orthonormalize and rotate these M vectors within the data-informed subspace (Appendix A). In particular, the eigenvectors of our misfit Hessian – which contains the *second* derivatives of J_{misfit} – are fully determined by *first* (rather than second) derivatives of the observed quantities, i.e., by $\nabla_{\mathbf{u}}\text{Obs}_i$.

For $M = 1$, the observational noise, ε^2, appears in the denominator of λ_1 (eq. (5)). In particular, for vanishing ε^2, the eigenvalue λ_1 tends to infinity. This fact generalizes to the case $M > 1$: in the limit of vanishing observational noise ($\mathbf{R} \searrow 0$), the eigenvalues λ_i of the misfit Hessian (eq. (3)) tend to infinity,

$$\lambda_i \nearrow \infty. \tag{7}$$

That is, $J_{\text{misfit}}(\tilde{\mathbf{u}})$ becomes increasingly curved (Fig. 2(a)) and the constraint by the observations increasingly strong.

2.3 Forward uncertainty propagation

To assess the observational constraints on a QoI, the inverse uncertainty propagation along path (UQ1) has to be followed by a forward uncertainty propagation along path (UQ2) (Fig. 1). In other words, we quantify how the uncertainty reduction in the controls, due to the new observational information, reduces uncertainty in the QoI, a diagnostic of the model evaluated at $\mathbf{u}_{\min}$. Forward propagation of prior uncertainties ($\mathbf{B}$, dotted green arrow) and posterior uncertainties ($\mathbf{P}$, dotted black arrow) along path (UQ2) results in the prior and posterior QoI variances (see Isaac et al., 2015, or Text S2 in supporting information):

$$(\sigma^{\mathbf{C}}_{\text{QoI}})^2 \; = \; (\nabla_{\mathbf{u}}\text{QoI})^T \, \mathbf{C} \, (\nabla_{\mathbf{u}}\text{QoI}) \; = \; \left\| \mathbf{C}^{1/2} \, \nabla_{\mathbf{u}}\text{QoI} \right\|^2, \qquad \mathbf{C} \in \{\mathbf{B}, \mathbf{P}\}. \tag{8}$$

We infer the prior-to-posterior reduction in QoI uncertainty, relative to the prior uncertainty:

$$\Delta\sigma^2_{\text{QoI}} \; = \; \frac{(\sigma^{\mathbf{B}}_{\text{QoI}})^2 - (\sigma^{\mathbf{P}}_{\text{QoI}})^2}{(\sigma^{\mathbf{B}}_{\text{QoI}})^2} \; = \; \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1} \, (\mathbf{q} \bullet \mathbf{v}_i)^2 \in [0, 1). \tag{9}$$

The second equality in eq. (9) holds by virtue of eqs. (8) and (2). Here, $\{(\mathbf{v}_i, \lambda_i)\}_{i=1}^{M'}$ are the eigenvectors and eigenvalues of the misfit Hessian (eq. (3)), $\bullet$ denotes the 'dot' (or Euclidean inner) product between two vectors in $\mathbb{R}^N$, and

$$\mathbf{q} = (\sigma^{\mathbf{B}}_{\text{QoI}})^{-1} \left[\mathbf{B}^{T/2} \, \nabla_{\mathbf{u}}\text{QoI}\right] \in \mathbb{R}^N. \tag{10}$$

The unit vector $\mathbf{q}$ is of key interest: it is the direction within the control space to be constrained in order to inform the QoI. It can be written as the orthogonal decomposition $\mathbf{q} = \mathbf{q}_{\text{obs}} + \mathbf{q}_{\text{null}}$ (Fig. 2(c)). $\mathbf{q}_{\text{obs}}$ is the component that lies in the data-informed subspace, given by the projection $\mathbf{q}_{\text{obs}} = \sum_{i=1}^{M'} (\mathbf{q} \bullet \mathbf{v}_i) \, \mathbf{v}_i$. The component $\mathbf{q}_{\text{null}}$ lies in the orthogonal complement of the data-informed subspace: the null space, i.e., the subspace that is not informed by the data. Uncertainty is only reduced along the data-informed component, $\mathbf{q}_{\text{obs}}$, not along the nullspace component, $\mathbf{q}_{\text{null}}$.

2.4 Dynamical and effective proxy potential

Relative reduction in QoI uncertainty, $\Delta\sigma^2_{\text{QoI}}$ (eq. (9)), rigorously quantifies the dynamical constraints of an observing system (gray box, Fig. 1) on a QoI (purple box, Fig. 1), as the result of the composite uncertainty propagation along paths (UQ1) and (UQ2). We refer to $\Delta\sigma^2_{\text{QoI}}$ as the *proxy potential* of the observing system for the QoI (Loose et al., 2020). Building on eq. (9), we distinguish between *dynamical* proxy potential

$$\text{DPP}(\text{Obs}_1, \ldots, \text{Obs}_M; \text{QoI}) = \sum_{i=1}^{M'} (\mathbf{q} \bullet \mathbf{v}_i)^2 \in [0, 1] \tag{11}$$

and *effective* proxy potential

$$\text{EPP}(\text{Obs}_1, \ldots, \text{Obs}_M; \text{QoI}) = \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1} (\mathbf{q} \bullet \mathbf{v}_i)^2 \in [0, 1) \tag{12}$$

of the examined observing system for the QoI. Recall that $M' \leq M$ is the number of independent data constraints, characterized by the eigenvectors and eigenvalues $\{\mathbf{v}_i, \lambda_i\}_{i=1}^{M'}$ of the misfit Hessian (eq. (3)). Geometrically, DPP is equal to the squared length of $\mathbf{q}_{\text{obs}}$, the data-informed component of $\mathbf{q}$ in control space (Fig. 2(c)). Note that EPP is smaller than DPP, because all factors $\eta_i = \lambda_i/(\lambda_i + 1)$ are smaller than 1. For vanishing observational noise, EPP approaches DPP, since all eigenvalues λ_i tend to infinity (eq. (7)), and consequently all factors $\eta_i = \lambda_i/(\lambda_i+1)$ tend to 1 (see also Appendix B and Fig. 7). The bounds for DPP and EPP correspond to the cases for which the observing system provides *no* constraint (EPP = DPP = 0), and for which it serves as a *perfect* proxy for the QoI, in the case of noise-free observations (DPP = 1) and noisy observations (EPP $\nearrow$ 1).

If the observing system under consideration consists of only a single observation ($M = 1$), eq. (11) simplifies to $\text{DPP}(\text{Obs}_1; \text{QoI}) = (\mathbf{q} \bullet \mathbf{v}_1)^2$, which coincides with the definition of dynamical proxy potential in Loose et al. (2020, eq. (4) therein).

3 Application to the North Atlantic

We illustrate the concepts of Hessian UQ and proxy potential in a North Atlantic case study. Section 3.1 describes our experimental setup, including our choice of QoI and observations. We then assess proxy potential of the observations for the QoI, for the cases of noise-free observations (DPP, section 3.2) and noisy observations (EPP, section 3.3).

3.1 Experimental Setup

Our experimental setup coincides with the one described in section 3.1 of Loose et al. (2020) and is embedded in the ECCO version 4, release 2 (ECCOv4r2, Forget et al., 2015) state estimation framework. We use the Massachusetts Institute of Technology general circulation model (Marshall et al., 1997; Adcroft et al., 2018), in a global configuration, at a nominal horizontal resolution of 1°, and with 50 vertical levels. The linear sensitivities of the QoI and observed quantities to all control variables (eqs. (5),(10), and Appendix A) are computed using the respective adjoint models generated through algorithmic differentiation with the commercial tool Transformation of Algorithms in Fortran (TAF, Giering & Kaminski, 2003).

Our QoI is heat transport across the Iceland-Scotland ridge (black line, Fig. 3), denoted by HT_{ISR}. We study four different hypothetical temperature observations in the North Atlantic, located inside the yellow dots in Fig. 3, and labeled by θ^A, θ^B, θ^C, θ^D. Observations θ^A and θ^C are in the Irminger Sea at (40 °W, 60 °N), observation θ^B off the Portuguese coast at (12 °W, 41 °N), and θ^D in Denmark Strait at (28 °W, 66 °N). θ^A,

Figure 3. Overview map of the case study in this work (modi ed from Fig. 2 in Loose et al., 2020). The QoI is heat transport across the Iceland-Scotland ridge (black line), denoted by HT_{ISR}. The temperature observations ^{A}, ^{B}, ^{C}, ^{D} are located inside the yellow dots. ^{A}, ^{B}, ^{D} are subsurface (at 300 m depth), ^{C} at the sea surface. The arrows show approximate pathways of near-surface currents carrying warm Atlantic waters (orange) and cold Arctic waters (purple): NAC = North Atlantic Current; NwAC = Norwegian Atlantic Current; IC = Irminger Current.

Table 2. Observational noise $\varepsilon_\star$, prior uncertainty $\sigma_\star^{\mathbf{B}}$ (eq. (6)), sensitivity-to-noise ratio $\lambda^\star$ (eq. (18)), and effectiveness $\eta^\star = \lambda^\star/(\lambda^\star+1)$, for each observation $\theta^\star$, $\star = A, B, C, D$.

Obs	Location	$\varepsilon_\star$	$\sigma_\star^{\mathbf{B}}$	$\lambda^\star$	$\lambda^\star/(\lambda^\star+1)$
θ^A	Irminger Sea (subsurface)	0.1 °C	0.048 °C	0.23	19%
θ^B	Portuguese Coast (subsurface)	0.1 °C	0.059 °C	0.35	26%
θ^C	Irminger Sea (surface)	0.1 °C	0.230 °C	5.29	84%
		0.2 °C	0.230 °C	1.32	57%
		0.3 °C	0.230 °C	0.59	37%
θ^D	Denmark Strait (subsurface)	0.1 °C	0.071 °C	0.50	33%

surface observation θ^A (Fig. 6(l)). Thus, θ^C does not lead to a gain in DPP when added to θ^A (Fig. 6(d)).

Finally, we are interested in the maximum achievable DPP for $\mathrm{HT_{ISR}}$, obtained by combining all four observations in our case study. Viewed within the three-dimensional subspace that is informed by the observing system $\{\theta^A, \theta^B, \theta^D\}$ (Fig. 5(f)), the $\{\theta^A, \theta^D\}$-informed yellow plane is almost orthogonal to the $\{\theta^A, \theta^B\}$-informed green plane (where the black plane would be exactly orthogonal to the green plane). Hence, when adding θ^D to the observing system $\{\theta^A, \theta^B\}$, the gain in DPP (green-yellow hatched, Fig. 6(e)) almost coincides with $(\mathbf{q}\bullet(\mathbf{v}^D)^\perp)^2$ (yellow hatched, Fig. 6(c)), leading to a total DPP of 35% (Fig. 6(e)). Completing the observing system by θ^C does not increase the DPP any further (Fig. 6(f)).

3.3 Noisy Observations

So far, our analysis has assumed noise-free observations. Next, we study the EPP of our observations $\theta^\star$; this notion does account for observational noise, as encoded in the Gaussian noise matrix $\mathbf{R}$ (eq. (1)). Recall that the EPP of $\theta^\star$ for $\mathrm{HT_{ISR}}$ is equal to the relative uncertainty reduction in $\mathrm{HT_{ISR}}$ that is achieved when adding $\theta^\star$ to the underlying state estimation framework (eq. (12)). Following the common assumption of uncorrelated observation errors (e.g., Forget et al., 2015), we only need to specify the diagonal entries of $\mathbf{R}$, i.e., the error variance $\varepsilon_\star^2$ of each observation $\theta^\star$. We assign $\varepsilon_\star = 0.1\,°\mathrm{C}$ for all observations (Table 2). We also consider the impact of varying ε_C, by testing for $\varepsilon_C = 0.2\,°\mathrm{C}$ and $\varepsilon_C = 0.3\,°\mathrm{C}$. The rationale for this addition is that climatological surface temperature, measured by θ^C, is more variable than climatological subsurface temperature (Locarnini et al., 2013), and can therefore be expected to be more noisy.

3.3.1 Sensitivity-To-Noise Ratio

The strength of the constraint provided by each individual observation $\theta^\star$ is quantified by the eigenvalue $\lambda^\star$ (eq. (5)) corresponding to the $\theta^\star$-informed direction $\mathbf{v}^\star$ (Fig. 2(a)). It is given by

$$\lambda^\star = \frac{(\sigma_\star^{\mathbf{B}})^2}{\varepsilon_\star^2} = \frac{1}{\varepsilon_\star^2} \sum_{m=1}^{4} \sum_{i,j} \left(\frac{\partial \theta^\star}{\partial F_m(i,j)} \Delta F_m \right)^2. \tag{18}$$

$\lambda^\star$ describes the *sensitivity-to-noise ratio* (SensNR): it is large if either $\theta^\star$ has high overall prior-weighted sensitivity, $(\sigma_\star^{\mathbf{B}})^2$, or if the observational noise $\varepsilon_\star^2$ is small. Since surface temperature is much more sensitive to atmospheric forcing than subsurface temperature ($\sigma_C^{\mathbf{B}} \gg \sigma_\star^{\mathbf{B}}$, $\star = A, B, D$, Table 2), the SensNR of θ^C is higher than that of $\theta^A, \theta^B, \theta^D$ (Fig. 7). This remains true if the noise variance for θ^C (i.e., ε_C^2) is assumed four – or even nine – times as large as that of the subsurface observations (Fig. 7).

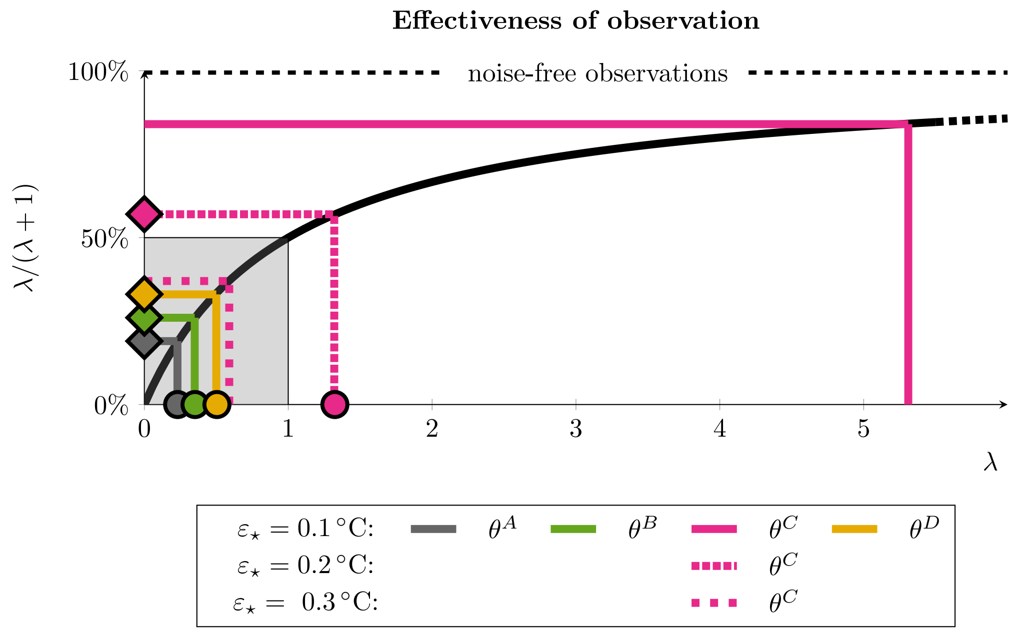

Figure 7. The black curve is the function $\lambda \mapsto \lambda/(\lambda+1)$. The colored lines map the SensNR of $\theta^\star$ ($\lambda^\star$, eq. (18), circles) to the effectiveness of $\theta^\star$ ($\eta^\star = \lambda^\star/(\lambda^\star+1)$, diamonds), cf. the values in Table 2. An observation that falls into the light gray rectangle has SensNR smaller than 1.

Note that $(\sigma_\star^{\mathbf{B}})^2$ is equal to the prior uncertainty in the observed quantity $\theta^\star$ (cf. eq (8)), i.e. the uncertainty given the prior knowledge in the ocean state estimate, before taking the actual measurement. Thus, observations with SensNR smaller than 1 (here: $\theta^A, \theta^B, \theta^D$, gray rectangle in Fig. 7) are characterized by a prior uncertainty, $(\sigma_\star^{\mathbf{B}})^2$, that is smaller than their assumed observational uncertainty, $\varepsilon_\star^2$.

The EPP of $\theta^\star$ for HT_{ISR} is given by $\text{EPP}^\star = \eta^\star \cdot \text{DPP}^\star$ (eq. (12)), with factor

$$\eta^\star = \frac{\lambda^\star}{\lambda^\star + 1} < 1.$$

The factor $\eta^\star$ indicates what fraction of $\text{DPP}^\star$ can be retrieved and will therefore be referred to as the 'effectiveness' of the observation $\theta^\star$. Note that, in contrast to $\text{DPP}^\star$, the observation's effectiveness, $\eta^\star$, is independent of the QoI under consideration. Instead, it is solely determined by the observation's SensNR $\lambda^\star$. Since the function $\lambda \mapsto \lambda/(\lambda+1)$ increases monotonically with λ (Fig. 7), observations with higher SensNRs are more effective. Therefore, the effectiveness of the surface observation θ^C is higher than that of the subsurface observations $\theta^A, \theta^B, \theta^D$ (Fig. 7). In fact, the effectiveness of $\theta^A, \theta^B, \theta^D$ is less than 50%, due to their SensNR being less than 1 (gray rectangle, Fig. 7).

In this section, we studied the SensNR, $\lambda^\star$, and associated effectiveness, $\eta^\star$, of each *individual* observation $\theta^\star$. In the next section, we will establish a connection between the $\lambda^\star$ and the set of eigenvalues $\{\lambda_i\}_{i=1}^4$, where the latter set characterizes the observing system that is *jointly* formed by $\{\theta^A, \theta^B, \theta^C, \theta^D\}$.

3.3.2 Combining Noisy Observations

We now combine all four temperature observations of our case study, while taking into account their observational noise. In eq. (1), the resulting observing system is

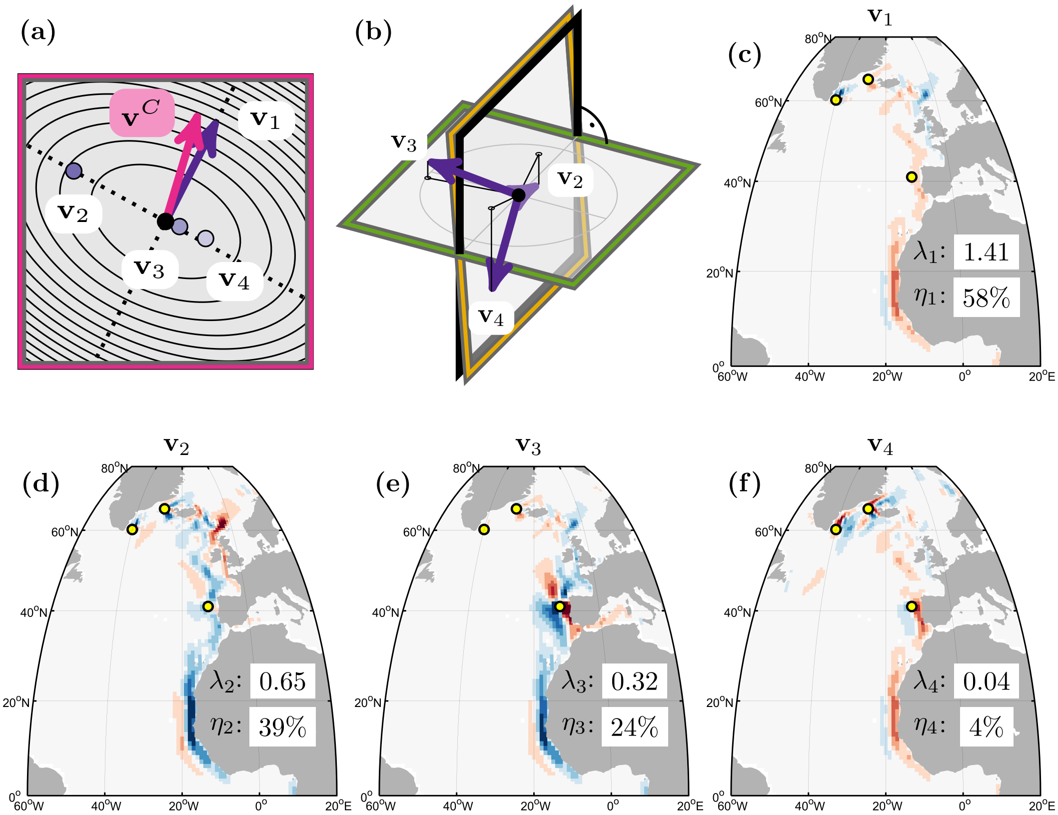

Figure 8. Eigen-decomposition $\{\mathbf{v}_i, \lambda_i\}_{i=1}^4$ of $\mathbf{H}_{\text{misfit}}$ (eq. (3)) for the observing system in eq. (19), with $\varepsilon_\star = 0.1\,^\circ\mathrm{C}$ for $\star = A, B, D$ and $\varepsilon_C = 0.2\,^\circ\mathrm{C}$. (a),(b): Orientation of the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ (purple vectors/dots) within the (a) $\{\theta^A, \theta^C\}$-informed, (b) $\{\theta^A, \theta^B, \theta^D\}$-informed subspace of the control space (cf. Fig. 6(l), Fig. 5(f)). The ellipses in (a) show the contour lines of $J_{\text{misfit}}(\tilde{\mathbf{u}})$. (c)-(f): τ_y component of the four eigenvectors. Each inset reports the eigenvalue λ_i, and the associated effectiveness $\eta_i = \lambda_i/(\lambda_i + 1)$.

represented by

$$\mathbf{Obs} = \left(\theta^A, \theta^B, \theta^C, \theta^D\right)^T, \qquad \mathbf{R} = \mathrm{diag}(\varepsilon_A^2, \varepsilon_B^2, \varepsilon_C^2, \varepsilon_D^2), \tag{19}$$

where the latter denotes a diagonal 4×4 matrix, with diagonal entries equal to the noise variances $\varepsilon_\star^2$, chosen as in Table 2. For the sake of brevity, we focus on the case $\varepsilon_C = 0.2\,^\circ\mathrm{C}$ (cases with alternative choices for ε_C are presented in the supporting information, Fig. S.1). We compute the eigenvectors and eigenvalues, $\{\mathbf{v}_i, \lambda_i\}_{i=1}^4$, of the misfit Hessian $\mathbf{H}_{\mathrm{misfit}}$ (eq. (3)) as described in Appendix A.

By definition, the first eigenvector $\mathbf{v}_1$ points in the direction of maximal curvature of $J_{\mathrm{misfit}}(\tilde{\mathbf{u}})$. This direction is almost aligned with the θ^C-informed direction, spanned by $\mathbf{v}^C$ (Fig. 8(a)), because the surface observation θ^C has a much higher SensNR than the remaining observations (Fig. 7). The remaining eigenvectors, $\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$, have little contribution from θ^C (purple dots, Fig. 8(a)), and are instead a linear combination of $\mathbf{v}^\star$, $\star = A, B, D$ (Fig. 8(b)). The τ_y component of $\mathbf{v}_2$ (Fig. 8(d)) extracts the dominant sensitivity patterns along the eastern boundary of the North Atlantic (region (I)), shared by θ^A, θ^B, and θ^D, and in the northeast Atlantic and the Nordic Seas (region (II)), shared by θ^A and θ^D (Figs. 5(a)-(c)). The τ_y component of $\mathbf{v}_3$ (Fig. 8(e)) is governed by sensitivities in region (I), as set by θ^B. Meanwhile, the τ_y component of $\mathbf{v}_4$ (Fig. 8(f)) is dominated by sensitivity dipoles local to the three observing sites (yellow dots), which emerge due to the effect of Ekman pumping.

The eigenvalues λ_i (inserted in Figs. 8(c)-(f)) are closely linked to the SensNRs $\lambda^\star$ (Table 2) of all observations $\theta^\star$ involved, via the following relations (Bunch et al., 1978):

$$\sum_{i=1}^{4} \lambda_i = \lambda^A + \lambda^B + \lambda^C + \lambda^D \qquad \text{and} \qquad \lambda_1 \geq \max\{\lambda^A, \lambda^B, \lambda^C, \lambda^D\}. \tag{20}$$

Each eigenvalue λ_i determines the effectiveness, $\eta_i = \lambda_i/(\lambda_i + 1)$, of the eigenvector $\mathbf{v}_i$ (insets in Figs. 8(c)-(f) and diamonds in Fig. 9(b)). We consider how the effectiveness of the observing system components changes with observational noise, and inflate the observational noise covariance by a factor α, to $\alpha\mathbf{R}$. As α varies from 0 (no noise) to 1 (full noise), effectiveness decays as $\lambda/(\lambda+\alpha)$, from 100% to $\lambda/(\lambda+1)$ (eq. (B1)). Here, λ is a placeholder for either a SensNR $\lambda^\star$ (Fig. 9(a)) or an eigenvalue λ_i (Fig. 9(b)). The decay in effectiveness of the surface observation θ^C (Fig. 9(a)) as well as $\mathbf{v}_1$ (Fig. 9(b)) is slower than that of the remaining observations and eigenvectors.

Decay in effectiveness causes decay in proxy potential. When considering individual observations $\theta^\star$, inflating observational noise leads to a decay in proxy potential according to $\lambda^\star/(\lambda^\star + \alpha) \cdot (\mathbf{q} \bullet \mathbf{v}^\star)^2$ (Fig. 9(c)). If instead, the full observing system is considered jointly, the decay in proxy potential is given by

$$\sum_{i=1}^{4} \frac{\lambda_i}{\lambda_i + \alpha} \, (\mathbf{q} \bullet \mathbf{v}_i)^2, \tag{21}$$

see Fig. 9(d) (and Appendix B). The expression in equation (21) involves the eigenvectors and eigenvalues of the misfit Hessian. For noise-free observations, proxy potential is equal to DPP ($\alpha = 0$, pentagons in Figs. 9(c),(d), eq. (11)). It decays to EPP for fully inflated noise ($\alpha = 1$, squares in Figs. 9(c),(d), eq. (12)). Even though the surface observation θ^C has highest SensNR and, thus, slowest decay in effectiveness (Fig. 9(a)), its proxy potential for $\mathrm{HT}_{\mathrm{ISR}}$ is lower than that of the subsurface observations θ^A and θ^D, due to its almost negligible DPP at the very outset $\alpha = 0$ (Fig. 9(c)). Through a similar argument, $\mathbf{v}_1$ contributes less to proxy potential than $\mathbf{v}_2$ and $\mathbf{v}_3$ (Fig. 9(d)), despite its relatively highest effectiveness (Fig. 9(b)). The insignificance of θ^C implies that proxy potential of the observing system $\{\theta^A, \theta^B, \theta^C, \theta^D\}$ for $\mathrm{HT}_{\mathrm{ISR}}$ is essentially insensitive to the choice of the observation error ε_C (Figs. S1(g)-(i)).

Figure 9. (a),(b) Decay in e ectiveness of (a) each individual observation $^{?}$ and (b) the eigenvectors v_i of the combined observing system (eq. (19)), as a function of . Incresasing the parameter in ates the observational noise (R) from no noise (= 0) to full noise (=1). Without noise, all observations have an equal e ectiveness of 100%. The colored diamonds repeat the values for /(+ 1) from Fig. 7 and Figs. 8(c)-(f). (c),(d) Decay in proxy potential for the QoI, HT_{ISR}, again as a function of . Without noise, proxy potential is equal to DPP (pentagons, cf. Figs. 4(b)-(e), Fig. 6(f), eq. (11)); but decays to EPP (squares, eq. (12)) for fully in ated noise. The black dashed curve in (c) coincides with the one in (d), and shows proxy potential for all observations combined.

The success of Hessian UQ relies on approaches that are more computationally efficient, two of which we consider: first, an a-priori reduction, and second, a data-informed reduction of the control space dimension. In this paper we have pursued the second approach, as further discussed in the next paragraph. In contrast, Kaminski et al. (2015, 2018) follow the first approach, by aggregating and adjusting their control variables uniformly over large regions (e.g., Fig. 2 in Kaminski et al., 2015), rather than on a model grid point basis. This 'large region approach' reduces their control space to a total of about 150 control variables, and it is then feasible to explicitly compute the full Hessian (150^2 entries). In practice, the large region approach requires to spatially accumulate sensitivities of QoIs and observed quantities over the pre-defined large regions, as exemplified in Fig. 10. The eight regions defined in Fig. 10(c) reduce the dimension of our control space from $\mathcal{O}(10^6)$ (Table 1) to $8 \cdot 4 = 32$. However, the spatial accumulation of sensitivities implies that proxy origins and adjustment mechanisms, e.g., along the basin boundaries, are no longer resolved (Figs. 10(d),(e)) and proxy potential is artificially lost (right yellow label). Note that for other QoIs, this approach could overestimate (rather than underestimate) proxy potential and uncertainty reduction.

Because of the ad-hoc nature of a-priori control space reductions, and the difficulties it incurs (Fig. 10), we advocate the approach of data-informed reductions of the control space for the following reason. Even though the Hessian in our North Atlantic case study consists of $\mathcal{O}(10^{12})$ entries (section 3.1), the misfit Hessian is only of rank 4, equal to the number of observations involved. The four Hessian misfit eigenvectors with non-zero eigenvalues capture the Hessian's full information. They were extracted efficiently while preserving the physical mechanism that led to uncertainty reduction. The concept of data-informed control space reduction generalizes to large, complex observing systems, e.g., mixed mooring arrays and autonomous instruments, which include thousands to millions of observations in time and space. While it becomes intractable to compute all (thousands to millions of) misfit Hessian eigenvectors, randomized numerical linear algebra for low-rank approximations can be used to extract the *leading* eigenvectors with highest eigenvalues ($M' \ll M$ in eq. (3), Bui-Thanh et al., 2012; Kalmikov & Heimbach, 2014; Liberty et al., 2007).

Moore et al. (2017) used a related technique in a regional ocean setting. They derived data-informed reduced-rank approximations of the Hessian, but with reductions sought in the observation space, rather than the control space. The two approaches are equivalent (or 'dual' to each other), and the implementation of the underlying variational data assimilation scheme may determine which of the two approaches is more convenient to employ. We argue that an eigen-decomposition in the control space, as suggested here, has the appeal of a straightforward dynamical attribution of proxy origins.

4.4 Limitations

Some shortcomings of the method presented should be acknowledged. Hessian UQ relies on an accurate specification of the prior and noise covariance matrices, $\mathbf{B}$ and $\mathbf{R}$ (eq. (1)). This is emphasized, for instance, by the fact that the relative weight of surface vs. subsurface observational noise determines the observations' relative effectiveness, and thus the patterns that dominate the leading eigenvectors of the misfit Hessian (Fig. S.1). A second shortcoming is that the results may suffer from model dependency, a problem common to all methods for model-informed observing system design. A third limitation is that Hessian UQ makes a Gaussian approximation of the posterior probability function for the uncertain control space and the estimated ocean state space. This approximation is accurate if the linearized model provides a good representation of the ocean dynamics on the time scales investigated. The results by Loose et al. (2020) indicate that on the five-year time scale considered, nonlinearity is not a major obstacle, at least not in the non-eddy resolving model under consideration. In situations where strong non-linearities are barriers to Gaussian approximations, Hessian UQ in combination with non-

Gaussian sampling methods have shown promise (Petra et al., 2014), but are yet to be explored in ocean and climate modeling. For instance, Stochastic Newton MCMC employs a local Gaussian approximation (given by the local inverse Hessian), which is then used as a proposal distribution for the posterior probability distribution (Petra et al., 2014).

4.5 Outlook

In our case study, we made simplifying assumptions regarding the control variables and the prior error covariance matrix (Table 1) to enable a clearer understanding of the methodology. These simplifications are readily relaxed in future work. Based on the insights gained here, we aim to compute reduced-rank approximations of the Hessian for large observing systems within the ECCO framework. Our case study highlights that the stopping criterion for truncating the eigenvalue spectrum has to be chosen carefully, because the leading Hessian eigenvectors are not always the most important ones for informing a given QoI. Indeed, eigenvectors lower down in the spectrum captured important dynamical teleconnections originating from the sensitivity of subsurface (rather than surface) observations. Future work should address the interesting question whether the abundance of surface observations (available from satellite altimetry) and their mutual complementarity (due to their local sensitivity) may be able to cover for the large-scale sensitivities of subsurface observations.

The technique presented in this paper is complementary to the more widely used OSSEs. Hessian UQ elucidates dynamical teleconnections that communicate observational constraints – via ocean currents, wave dynamics, Ekman dynamics, and geostrophy – over basin-scale distances and on monthly to interannual time scales. It provides an approach for guiding the design of observing systems that (1) maximize the information about (possibly remote) QoIs that are difficult or impossible to observe directly, and (2) are complementary to the existing observational database. We hope that Hessian UQ, in combination with OSSEs and other tools, will be more widely used for tackling the grand community challenge of co-designing a cost-effective and long-term Atlantic observing system in the coming years.

Appendix A Eigen-Decomposition of the Misfit Hessian

For an observing system with M observations, the eigen-decomposition of the misfit Hessian (eq. (3)) can be computed from the prior-weighted sensitivity vectors $\mathbf{c}_i = \mathbf{B}^{T/2}\,\nabla_{\mathbf{u}}\mathrm{Obs}_i$ via the following two steps: (M.1) a QR decomposition of $\mathbf{B}^{T/2}\,\mathbf{A}^T$ in $\mathbb{R}^N$ and (M.2) and an eigen-decomposition in $\mathbb{R}^M$.

In step (M.1), the QR decomposition of $\mathbf{B}^{T/2}\,\mathbf{A}^T = \left(\ \mathbf{c}_1 \,\middle|\, \cdots \,\middle|\, \mathbf{c}_M\ \right)$ is computed via the Gram-Schmidt process:

- $\tilde{\mathbf{w}}_1 := \mathbf{c}_1, \qquad \mathbf{w}_1 = \|\tilde{\mathbf{w}}_1\|^{-1}\,\tilde{\mathbf{w}}_1$
- For $j = 2, \ldots, M$: $\qquad \tilde{\mathbf{w}}_j = \mathbf{c}_j - \sum_{i=1}^{j-1}(\mathbf{c}_j \bullet \mathbf{w}_i)\,\mathbf{w}_i, \qquad \mathbf{w}_j = \|\tilde{\mathbf{w}}_j\|^{-1}\,\tilde{\mathbf{w}}_j$

Then the $N \times M$ matrix $\mathbf{Q} := \left(\ \mathbf{w}_1 \,\middle|\, \ldots \,\middle|\, \mathbf{w}_M\ \right)$ and the $M \times M$ matrix

$$\tilde{\mathbf{R}} = \begin{pmatrix} \|\tilde{\mathbf{w}}_1\| & \mathbf{w}_1 \bullet \mathbf{c}_2 & \mathbf{w}_1 \bullet \mathbf{c}_3 & \cdots \\ 0 & \|\tilde{\mathbf{w}}_2\| & \mathbf{w}_2 \bullet \mathbf{c}_3 & \cdots \\ 0 & 0 & \|\tilde{\mathbf{w}}_3\| & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

provide the desired QR decomposition, i.e., they satisfy $\mathbf{B}^{T/2}\,\mathbf{A}^T = \mathbf{Q}\tilde{\mathbf{R}}$. In step (M.2), one finds an orthogonal $M \times M$ matrix $\mathbf{O}$ and $\lambda_1 \geq \ldots \geq \lambda_M \geq 0$ such that

$$\tilde{\mathbf{R}}\,\mathbf{R}^{-1}\,\tilde{\mathbf{R}}^T = \mathbf{O}\,\mathrm{diag}(\lambda_1, \ldots, \lambda_M)\,\mathbf{O}^T,$$

by means of dense matrix algebra. Combining steps (M.1) and (M.2) gives

$$\mathbf{H}_{\text{misfit}} = \mathbf{B}^{T/2}\,\mathbf{A}^T\,\mathbf{R}^{-1}\,\mathbf{A}\,\mathbf{B}^{1/2} = \mathbf{Q}\,\tilde{\mathbf{R}}\,\mathbf{R}^{-1}\,\tilde{\mathbf{R}}^T\,\mathbf{Q}^T = \mathbf{QO}\;\text{diag}(\lambda_1,\ldots,\lambda_M)\;\mathbf{O}^T\mathbf{Q}^T,$$

and the i^{th} column of $\mathbf{QO}$ contains the i^{th} eigenvector of $\mathbf{H}_{\text{misfit}}$, with corresponding eigenvalue $\lambda_i \geq 0$. The eigenvectors corresponding to non-zero eigenvalues are the data-informed directions $\mathbf{v}_1,\ldots,\mathbf{v}_{M'}$.

Step (M.2) is feasible, as long as the number of observations, M, is small enough to allow for dense matrix algebra in $\mathbb{R}^M$. For large M, one has to resort to randomized numerical linear algebra for low-rank approximations of the misfit Hessian. Such randomized algorithms continue to follow the outlined steps (M.1) and (M.2), except that the decomposition in (M.1) is substituted by an approximate, low-rank QR factorization (Halko et al., 2011; Liberty et al., 2007).

Appendix B Inflating Noise and Prior Covariances

Modifying the noise covariance matrix via $\mathbf{R} \to \alpha\mathbf{R}$ reflects a uniform deflation ($0 < \alpha < 1$) or inflation ($\alpha > 1$) of observational noise. This modification results in a reciprocal scaling of the misfit Hessian, $\mathbf{H}_{\text{misfit}} \to \mathbf{H}_{\text{misfit}}/\alpha$. Here, we substituted $\alpha\mathbf{R}$ for $\mathbf{R}$ in eq. (3), and assume the sensitivity matrix $\mathbf{A}$ unchanged (even though its evaluation point may change). The scaled misfit Hessian, $\mathbf{H}_{\text{misfit}}/\alpha$, has unchanged eigenvectors $\mathbf{v}_i$, and new eigenvalues λ_i/α. Therefore, effectiveness scales as

$$\frac{\lambda}{\lambda+1} \to \frac{(\lambda/\alpha)}{\lambda/\alpha+1} = \frac{\lambda}{\lambda+\alpha} \tag{B1}$$

and effective proxy potential (eq. (12)) as in eq. (21).

We note that the same scaling of the misfit Hessian, $\mathbf{H}_{\text{misfit}} \to \mathbf{H}_{\text{misfit}}/\alpha$, can be achieved by modifying the prior covariance matrix via $\mathbf{B} \to \mathbf{B}/\alpha$, while keeping the noise covariance matrix unchanged. The value $\alpha = 0$ in Fig. 9 corresponds therefore either to the limit of vanishing observational noise or inifinite prior uncertainty. Similarly, $\alpha = 1$ represents not only the case of unchanged $\mathbf{R}$ and $\mathbf{B}$, but also the case of $\gamma\mathbf{R}$ and $\mathbf{B}/\gamma$, for any $\gamma > 0$.

Acknowledgments

We gratefully acknowledge comments by two anonymous reviewers, which helped improve the quality of this paper. We thank Gaël Forget for developing the flexible adjoint modeling tools within ECCOv4, which this work substantially benefitted from. We also thank Timothy Smith, Helen R. Pillar, Daniel N. Goldberg, Carl Wunsch, and Kerim H. Nisancioglu for useful discussions and suggestions. This study is supported by the grants NSF-OCE-1924546, NSF-OPP-1751120, NSF-OPP-1603903, and by the ECCO project via a JPL/Caltech subcontract. The ECCOv4r2 model setup used in this work can be accessed from public repositories and permanent archives (Campin et al., 2019; Forget, 2018, 2016a, 2016b). Adjoint code was generated using the TAF software tool, created and maintained by FastOpt GmbH (http://www.fastopt.com/).

References

Adcroft, A., Campin, J., Dutkiewicz, S., Evangelinos, C., Ferreira, D., Forget, G., ... Molod, A. (2018). Mitgcm User Manual. *Zenodo*. doi: 1721.1/117188

Balmaseda, M., Anderson, D., & Vidard, A. (2007). Impact of Argo on analyses of the global ocean. *Geophysical Research Letters*, *34*(16). doi: 10.1029/2007GL030452

Bellman, R. (1957). *Dynamic Programming.* Princeton University Press.

Bennett, A. F. (1985). Array design by inverse methods. *Progress in Oceanography*, *15*(2), 129–156. doi: 10.1016/0079-6611(85)90033-3

Bui-Thanh, T., Burstedde, C., Ghattas, O., Martin, J., Stadler, G., & Wilcox, L. C. (2012). Extreme-scale UQ for Bayesian Inverse Problems Governed by PDEs. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis* (pp. 1–11).

Bunch, J. R., Nielsen, C. P., & Sorensen, D. C. (1978, March). Rank-one modification of the symmetric eigenproblem. *Numerische Mathematik*, *31*(1), 31–48. doi: 10.1007/BF01396012

Campin, J.-M., Heimbach, P., Losch, M., Forget, G., Hill, E., Adcroft, A., ... McRae, A. T. T. (2019). *MITgcm/MITgcm: checkpoint67m.* Zenodo. doi: 10.5281/zenodo.3492298

Chen, P. (2011). Hessian Matrix vs. Gauss-Newton Hessian Matrix. *SIAM Journal on Numerical Analysis*, *49*(4), 1417–1435. doi: 10.1137/100799988

Forget, G. (2016a). *ECCO version 4 release 2 inputs: model initialization.* Harvard Dataverse. doi: 10.7910/DVN/7XYXSF

Forget, G. (2016b). *ECCO version 4 release 2 inputs: surface forcing fields.* Harvard Dataverse. doi: 10.7910/DVN/9WYSZF

Forget, G. (2018). *gaelforget/eccov4: Documentation updates.* Zenodo. doi: 10.5281/zenodo.1211363

Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C. (2015). ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geosci. Model Dev.*, *8*(10), 3071–3104. doi: 10.5194/gmd-8-3071-2015

Fujii, Y., Rémy, E., Zuo, H., Oke, P., Halliwell, G., Gasparin, F., ... Usui, N. (2019). Observing System Evaluation Based on Ocean Data Assimilation and Prediction Systems: On-Going Challenges and a Future Vision for Designing and Supporting Ocean Observational Networks. *Frontiers in Marine Science*, *6*. doi: 10.3389/fmars.2019.00417

Fujii, Y., Tsujino, H., Usui, N., Nakano, H., & Kamachi, M. (2008). Application of singular vector analysis to the Kuroshio large meander. *Journal of Geophysical Research: Oceans*, *113*(C7). doi: 10.1029/2007JC004476

Gasparin, F., Guinehut, S., Mao, C., Mirouze, I., Rémy, E., King, R. R., ... Masina, S. (2019). Requirements for an Integrated in situ Atlantic Ocean Observing System From Coordinated Observing System Simulation Experiments. *Frontiers in Marine Science*, *6*. doi: 10.3389/fmars.2019.00083

Giering, R., & Kaminski, T. (2003). Applying TAF to generate efficient derivative code of Fortran 77-95 programs. *Proc. Appl. Math. Mech.*, *2*(1), 54–57. doi: 10.1145/293686.293695

Griffa, A., Molcard, A., Raicich, F., & Rupolo, V. (2006). Assessment of the impact of TS assimilation from ARGO floats in the Mediterranean Sea. *Ocean Science*, *2*(2), 237–248. doi: https://doi.org/10.5194/os-2-237-2006

Halko, N., Martinsson, P. G., & Tropp, J. A. (2011). Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions. *SIAM Review*, *53*(2), 217–288. (Publisher: Society for Industrial and Applied Mathematics) doi: 10.1137/090771806

Halliwell, G. R., Mehari, M. F., Hénaff, M. L., Kourafalou, V. H., Androulidakis, I. S., Kang, H. S., & Atlas, R. (2017). North Atlantic Ocean OSSE system: Evaluation of operational ocean observing system components and supplemental seasonal observations for potentially improving tropical cyclone prediction in coupled systems. *Journal of Operational Oceanography*, *10*(2), 154–175. doi: 10.1080/1755876X.2017.1322770

Heimbach, P., Fukumori, I., Hill, C. N., Ponte, R. M., Stammer, D., Wunsch, C., ... Zhang, H. (2019). Putting It All Together: Adding Value to the Global Ocean and Climate Observing Systems With Complete Self-Consistent

Ocean State and Parameter Estimates. *Frontiers in Marine Science*, *6*. doi: 10.3389/fmars.2019.00055

Heimbach, P., Wunsch, C., Ponte, R. M., Forget, G., Hill, C., & Utke, J. (2011). Timescales and regions of the sensitivity of Atlantic meridional volume and heat transport: Toward observing system design. *Deep Sea Research Part II: Topical Studies in Oceanography*, *58*(17-18), 1858–1879. doi: 10.1016/j.dsr2.2010.10.065

Hirschi, J., Baehr, J., Marotzke, J., Stark, J., Cunningham, S., & Beismann, J.-O. (2003). A monitoring design for the Atlantic meridional overturning circulation. *Geophysical Research Letters*, *30*(7). doi: 10.1029/2002GL016776

Isaac, T., Petra, N., Stadler, G., & Ghattas, O. (2015). Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet. *Journal of Computational Physics*, *296*, 348–368. doi: 10.1016/j.jcp.2015.04.047

Jones, D., Forget, G., Sinha, B., Josey, S., Boland, E., Meijers, A., & Shuckburgh, E. (2018). Local and Remote Influences on the Heat Content of the Labrador Sea: An Adjoint Sensitivity Study. *Journal of Geophysical Research: Oceans*, *0*(0). doi: 10.1002/2018JC013774

Kalmikov, A., & Heimbach, P. (2014). A Hessian-Based Method for Uncertainty Quantification in Global Ocean State Estimation. *SIAM Journal on Scientific Computing*, *36*(5), S267–S295. doi: 10.1137/130925311

Kaminski, T., Kauker, F., Eicken, H., & Karcher, M. (2015). Exploring the utility of quantitative network design in evaluating Arctic sea ice thickness sampling strategies. *The Cryosphere*, *9*(4), 1721–1733. doi: 10.5194/tc-9-1721-2015

Kaminski, T., Kauker, F., Toudal Pedersen, L., Voßbeck, M., Haak, H., Niederdrenk, L., … Gråbak, O. (2018). Arctic Mission Benefit Analysis: impact of sea ice thickness, freeboard, and snow depth products on sea ice forecast performance. *The Cryosphere*, *12*(8), 2569–2594. doi: https://doi.org/10.5194/tc-12-2569-2018

Köhl, A. (2005). Anomalies of Meridional Overturning: Mechanisms in the North Atlantic. *Journal of Physical Oceanography*, *35*(8), 1455–1472. doi: 10.1175/JPO2767.1

Köhl, A., & Stammer, D. (2004). Optimal Observations for Variational Data Assimilation. *Journal of Physical Oceanography*, *34*(3), 529–542. doi: 10.1175/2513.1

Leutbecher, M. (2003). A Reduced Rank Estimate of Forecast Error Variance Changes due to Intermittent Modifications of the Observing Network. *Journal of the Atmospheric Sciences*, *60*(5), 729–742. doi: 10.1175/1520-0469(2003)060⟨0729:ARREOF⟩2.0.CO;2

Li, F., Lozier, M. S., & Johns, W. E. (2017). Calculating the Meridional Volume, Heat, and Freshwater Transports from an Observing System in the Subpolar North Atlantic: Observing System Simulation Experiment. *Journal of Atmospheric and Oceanic Technology*, *34*(7), 1483–1500. doi: 10.1175/JTECH-D-16-0247.1

Liberty, E., Woolfe, F., Martinsson, P.-G., Rokhlin, V., & Tygert, M. (2007). Randomized algorithms for the low-rank approximation of matrices. *Proceedings of the National Academy of Sciences*, *104*(51), 20167–20172. doi: 10.1073/pnas.0709640104

Locarnini, R., Mishonov, A., Antonov, J., Boyer, T., Garcia, H., Baranova, O., … Johnson, D. (2013, January). World Ocean Atlas 2009, Volume 1: Temperature. *S. Levitus, Ed. NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 pp.*.

Loose, N., Heimbach, P., Pillar, H. R., & Nisancioglu, K. H. (2020). Quantifying Dynamical Proxy Potential Through Shared Adjustment Physics in the North

Atlantic. *Journal of Geophysical Research: Oceans*, *125*(9), e2020JC016112. doi: 10.1029/2020JC016112

Marshall, J., Adcroft, A., Hill, C., Perelman, L., & Heisey, C. (1997). A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research: Oceans*, *102*(C3), 5753–5766. doi: 10.1029/96JC02775

Masuda, S., Awaji, T., Sugiura, N., Matthews, J. P., Toyoda, T., Kawai, Y., ... Fukasawa, M. (2010). Simulated Rapid Warming of Abyssal North Pacific Waters. *Science*, *329*(5989), 319–322. doi: 10.1126/science.1188703

Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B., ... Robinson, P. (2011). The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part III - Observation impact and observation sensitivity in the California Current System. *Progress in Oceanography*, *91*(1), 74–94. doi: 10.1016/j.pocean.2011.05.005

Moore, A. M., Arango, H. G., & Edwards, C. A. (2017). Reduced-Rank Array Modes of the California Current Observing System. *Journal of Geophysical Research: Oceans*, *123*(1), 452–465. doi: 10.1002/2017JC013172

National Academies of Sciences, Engineering, and Medicine. (2017). *Sustaining Ocean Observations to Understand Future Changes in Earth's Climate.* Washington, DC: The National Academies Press. doi: 10.17226/24919

Perez, R. C., Garzoli, S. L., Meinen, C. S., & Matano, R. P. (2011). Geostrophic Velocity Measurement Techniques for the Meridional Overturning Circulation and Meridional Heat Transport in the South Atlantic. *Journal of Atmospheric and Oceanic Technology*, *28*(11), 1504–1521. doi: 10.1175/JTECH-D-11-00058.1

Petra, N., Martin, J., Stadler, G., & Ghattas, O. (2014). A Computational Framework for Infinite-Dimensional Bayesian Inverse Problems, Part II: Stochastic Newton MCMC with Application to Ice Sheet Flow Inverse Problems. *SIAM Journal on Scientific Computing*, *36*(4), A1525–A1555. doi: 10.1137/130934805

Pillar, H. R., Heimbach, P., Johnson, H. L., & Marshall, D. P. (2016). Dynamical Attribution of Recent Variability in Atlantic Overturning. *Journal of Climate*, *29*(9), 3339–3352. doi: 10.1175/JCLI-D-15-0727.1

Smith, T., & Heimbach, P. (2019). Atmospheric origins of variability in the South Atlantic meridional overturning circulation. *Journal of Climate*, *32*(5), 1483-1500. doi: 10.1175/JCLI-D-18-0311.1

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation.* Society for Industrial and Applied Mathematics.

Thacker, W. C. (1989). The role of the Hessian matrix in fitting models to measurements. *Journal of Geophysical Research: Oceans*, *94*(C5), 6177–6196. doi: 10.1029/JC094iC05p06177

Zanna, L., Heimbach, P., Moore, A. M., & Tziperman, E. (2012). Upper-ocean singular vectors of the North Atlantic climate with implications for linear predictability and variability. *Quarterly Journal of the Royal Meteorological Society*, *138*(663), 500–513. doi: 10.1002/qj.937

Zhang, W. G., Wilkin, J. L., & Levin, J. C. (2010). Towards an integrated observation and modeling system in the New York Bight using variational methods. Part II: Representer-based observing strategy evaluation. *Ocean Modelling*, *35*(3), 134–145. doi: 10.1016/j.ocemod.2010.06.006

Supporting Information for "Leveraging Uncertainty Quantification to Design Ocean Climate Observing Systems"

Nora Loose[1], Patrick Heimbach[1,2,3]

[1]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX, USA

[2]Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA

[3]Institute for Geophysics, The University of Texas at Austin, Austin, TX, USA

Contents of this file

1. Text S1: Inverse Uncertainty Propagation.
2. Text S2: Forward Uncertainty Propagation.
3. Figure S1.

Text S1. Inverse Uncertainty Propagation. The Bayesian approach states the deterministic inverse problem (eq. (1)) as one of Bayesian inference over the space of unknown control variables, which are to be inferred from the observations and the ocean GCM dynamics. The solution is given by the posterior probability density function $\pi_{\text{post}}(\mathbf{u}|\mathbf{y}) \propto e^{-J(\mathbf{u})}$. Hence, the deterministic and Bayesian formulation of the inverse problem are interconnected by the fact that the deterministic least squares cost function J is the negative log-posterior in the Bayesian interpretation. Furthermore, the deterministic solution $\mathbf{u}_{\min}$ is the Maximum a Posteriori (MAP) point, i.e., the most likely solution, in the Bayesian framework. For more details, the reader is referred to the books by Tarantola (2005); Law, Stuart, and Zygalakis (2015).

To make the computation of the posterior probability density function $\pi_{\text{post}}(\mathbf{u}|\mathbf{y})$ computationally tractable, a linearization of the observation operator about the MAP Point, $\mathbf{u}_{\text{min}}$, is necessary (e.g., Bui-Thanh et al., 2012). This yields

$$\mathbf{Obs}(\mathbf{u}) \approx \mathbf{Obs}(\mathbf{u}_{\text{min}}) + \mathbf{A}(\mathbf{u} - \mathbf{u}_{\text{min}}), \tag{S.1}$$

where $\mathbf{A} = \frac{\partial(\mathbf{Obs})}{\partial \mathbf{u}}_{|\mathbf{u}_{\text{min}}}$ is the Jacobian matrix of the observation operator $\mathbf{u} \mapsto \mathbf{Obs}(\mathbf{u})$, evaluated at $\mathbf{u}_{\text{min}}$. The posterior distribution $\pi_{\text{post}}(\mathbf{u}|\mathbf{y}) \propto e^{-J(\mathbf{u})}$ then becomes

$$\pi_{\text{post}}(\mathbf{u}|\mathbf{y}) \propto C \cdot \exp\left(-\frac{1}{2}\left[(\mathbf{u} - \mathbf{u}_{\text{min}})^T (\mathbf{A}^T \mathbf{R}^{-1} \mathbf{A} + \mathbf{B}^{-1})^{-1} (\mathbf{u} - \mathbf{u}_{\text{min}})\right]\right), \tag{S.2}$$

where C is a constant factor, given by

$$C = \exp\left(-\frac{1}{2}\left[(\mathbf{y} - \mathbf{Obs}(\mathbf{u}_{\text{min}}))^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{Obs}(\mathbf{u}_{\text{min}})) + (\mathbf{u}_{\text{min}} - \mathbf{u}_0)^T \mathbf{B}^{-1} (\mathbf{u}_{\text{min}} - \mathbf{u}_0)\right]\right),$$

and can therefore be absorbed by the proportionality $\propto$. The right hand side of eq. (S.2) describes a Gaussian $\mathcal{N}(\mathbf{u}_{\text{min}}, \mathbf{P})$ with mean $\mathbf{u}_{\text{min}}$ and covariance matrix

$$\mathbf{P} = (\mathbf{A}^T \mathbf{R}^{-1} \mathbf{A} + \mathbf{B}^{-1})^{-1}. \tag{S.3}$$

The covariance matrix $\mathbf{P}$ (eq. (S.3)) is equal to $\mathbf{H}_J^{-1}$, the inverse of the linearized Hessian matrix of the cost function J (eq. (1)) at $\mathbf{u}_{\text{min}}$ (see also Thacker, 1989).

The linearized Hessian $\mathbf{H}_J$, in turn, is the sum of two matrices: $\mathbf{A}^T \mathbf{R}^{-1} \mathbf{A}$, which is the linearized Hessian of the model-data misfit term J_{misfit} (eq. (1)), and $\mathbf{B}^{-1}$, which is the Hessian of the regularization term J_{prior} (eq. (1)). It is the first matrix, $\mathbf{A}^T \mathbf{R}^{-1} \mathbf{A}$, that characterizes the observational constraints on the control variables. Therefore, as a next step, we perform an eigen-decomposition of the misfit Hessian. This will give further insights into the model input components that are best determined by the observing system under consideration.

As a preparatory step, we rescale the model-data misfit term, J_{misfit}, through a change of variables $\tilde{\mathbf{u}} = \mathbf{B}^{-1/2}\,\mathbf{u}$. Here, $\mathbf{B}^{1/2}$ is an invertible square root of $\mathbf{B}$, i.e., satisfies $\mathbf{B}^{1/2}\,\mathbf{B}^{T/2} = \mathbf{B}$. The rescaling can be thought of as nondimensionalization if $\mathbf{B}$ was diagonal. The rescaling is necessary in order to treat all control variables equally, since they represent different physical variables, characteristic of different orders of magnitudes. In the new coordinates, the Hessian of $J_{\mathrm{misfit}}(\tilde{\mathbf{u}})$ is given by the rescaled misfit Hessian (also referred to as the *prior-preconditioned* misfit Hessian, e.g., Bui-Thanh et al., 2012), equal to the $N \times N$ matrix

$$\mathbf{H}_{\mathrm{misfit}} = \mathbf{B}^{T/2}\,\mathbf{A}^{T}\,\mathbf{R}^{-1}\,\mathbf{A}\,\mathbf{B}^{1/2}. \tag{S.4}$$

For simpicity, we will hereafter drop the term 'rescaled', and refer to $\mathbf{H}_{\mathrm{misfit}}$ solely as the misfit Hessian. The misfit Hessian can be rewritten in terms of its eigen-decomposition:

$$\mathbf{H}_{\mathrm{misfit}} = \sum_{i=1}^{M'} \lambda_i \mathbf{v}_i \mathbf{v}_i^T, \tag{S.5}$$

with an orthonormal set of eigenvectors $\{\mathbf{v}_i\}_{i=1}^{M'}$ and corresponsing eigenvalues $\lambda_i > 0$. M' is defined as the number of strictly positive eigenvalues, while all remaining eigenvalues λ_i, $i > M'$, are equal to zero.

The Woodbury Formula (e.g., Section 2.7.3 in Press et al., 2007) states that, given any $N \times N$ matrix $\mathbf{M}$ and any $N \times M'$ matrix $\mathbf{V}$, the following identity holds true:

$$(\mathbf{M} + \mathbf{V}\mathbf{V}^T)^{-1} = \mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{V}\left(\mathbf{1}_{M'\times M'} + \mathbf{V}^T\mathbf{M}^{-1}\mathbf{V}\right)^{-1}\mathbf{V}^T\mathbf{M}^{-1}. \tag{S.6}$$

Here, $\mathbf{1}_{M'\times M'}$ is the $M' \times M'$ identity matrix. Assuming $\mathbf{M}^{-1}$ is already known, the formula provides an efficient way to compute the inverse of the sum of $\mathbf{M}$ and a low-rank matrix $\mathbf{V}\mathbf{V}^T$ (if $M' \ll N$). Using eqs. (S.3), (S.4), and (S.5), we have

$$\begin{aligned} \mathbf{P} &= \mathbf{B}^{1/2}\left(\mathbf{H}_{\text{misfit}} + \mathbf{1}_{N\times N}\right)^{-1}\mathbf{B}^{T/2} \\ &= \mathbf{B}^{1/2}\left(\sum_{i=1}^{M'} \lambda_i \mathbf{v}_i \mathbf{v}_i^T + \mathbf{1}_{N\times N}\right)^{-1}\mathbf{B}^{T/2}. \end{aligned} \tag{S.7}$$

The Woodbury Formula (eq. (S.6)) can now be applied to the inner piece in eq. (S.7), with $\mathbf{M} = \mathbf{1}_{N\times N}$ and $\mathbf{V}$ defined as the matrix formed by columns of $\sqrt{\lambda_i} \cdot \mathbf{v}_i$:

$$\mathbf{V} = \left[\sqrt{\lambda_1} \cdot \mathbf{v}_1 \middle| \cdots \middle| \sqrt{\lambda_{M'}} \cdot \mathbf{v}_{M'}\right].$$

This yields

$$\mathbf{P} = \mathbf{B}^{1/2}\left(\mathbf{1}_{N\times N} - \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1}\mathbf{v}_i\mathbf{v}_i^T\right)\mathbf{B}^{T/2} = \mathbf{B} - \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1}\left(\mathbf{B}^{1/2}\mathbf{v}_i\right)\left(\mathbf{B}^{1/2}\mathbf{v}_i\right)^T,$$

using the fact that $\{\mathbf{v}_i\}_{i=1}^{M'}$ is a set of orthonormal vectors.

Text S2. Forward Uncertainty Propagation. Consistent with the linearization of the observation operator (eq. (S.1)), the function $\mathbf{u} \mapsto \text{QoI}(\mathbf{u})$ is linearized about $\mathbf{u}_{\min}$:

$$\text{QoI}(\mathbf{u}) \approx \text{QoI}(\mathbf{u}_{\min}) + \frac{\partial(\text{QoI})}{\partial \mathbf{u}}_{|\mathbf{u}_{\min}}(\mathbf{u} - \mathbf{u}_{\min}). \tag{S.8}$$

The posterior distribution of the Bayesian solution of the inverse problem, $\pi_{\text{post}}(\mathbf{u}|\mathbf{y})$, is approximately Gaussian, given by $\mathcal{N}(\mathbf{u}_{\min}, \mathbf{P})$, with $\mathbf{P}$ given by eq. (2). A forward propagation of the posterior uncertainty (dotted black arrow, (UQ2), Fig. 1) leads to a posterior Gaussian distribution for QoI($\mathbf{u}$), since eq. (S.8) describes an affine transformation. The distribution is given by $\mathcal{N}(\mathrm{QoI}(\mathbf{u}_{\min}), (\sigma_{\mathrm{QoI}}^{\mathbf{P}})^2)$, where $(\sigma_{\mathrm{QoI}}^{\mathbf{P}})^2$ is the (scalar) variance, given by the projection

$$(\sigma_{\mathrm{QoI}}^{\mathbf{P}})^2 = (\nabla_{\mathbf{u}}\mathrm{QoI})^T \, \mathbf{P} \, (\nabla_{\mathbf{u}}\mathrm{QoI}) \, .$$

Similarly, the prior distribution of QoI($\mathbf{u}$) is obtained by a forward uncertainty propagation of the Gaussian prior $\mathcal{N}(\mathbf{u}_0, \mathbf{B})$ (dotted green arrow, (UQ2), Fig. 1). This leads to a prior Gaussian distribution for QoI($\mathbf{u}$), given by $\mathcal{N}(\mathrm{QoI}(\mathbf{u}_0), (\sigma_{\mathrm{QoI}}^{\mathbf{B}})^2)$ with

$$(\sigma_{\mathrm{QoI}}^{\mathbf{B}})^2 = (\nabla_{\mathbf{u}}\mathrm{QoI})^T \, \mathbf{B} \, (\nabla_{\mathbf{u}}\mathrm{QoI}) \, .$$

References

Bui-Thanh, T., Burstedde, C., Ghattas, O., Martin, J., Stadler, G., & Wilcox, L. C. (2012). Extreme-scale UQ for Bayesian Inverse Problems Governed by PDEs. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis* (pp. 1–11).

Law, K., Stuart, A., & Zygalakis, K. (2015). Data Assimilation - A Mathematical Introduction. *Springer*.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing.* Cambridge University Press.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation.* Society for Industrial and Applied Mathematics.

Thacker, W. C. (1989). The role of the Hessian matrix in fitting models to measurements. *Journal of Geophysical Research: Oceans*, *94*(C5), 6177–6196. doi: 10.1029/JC094iC05p06177

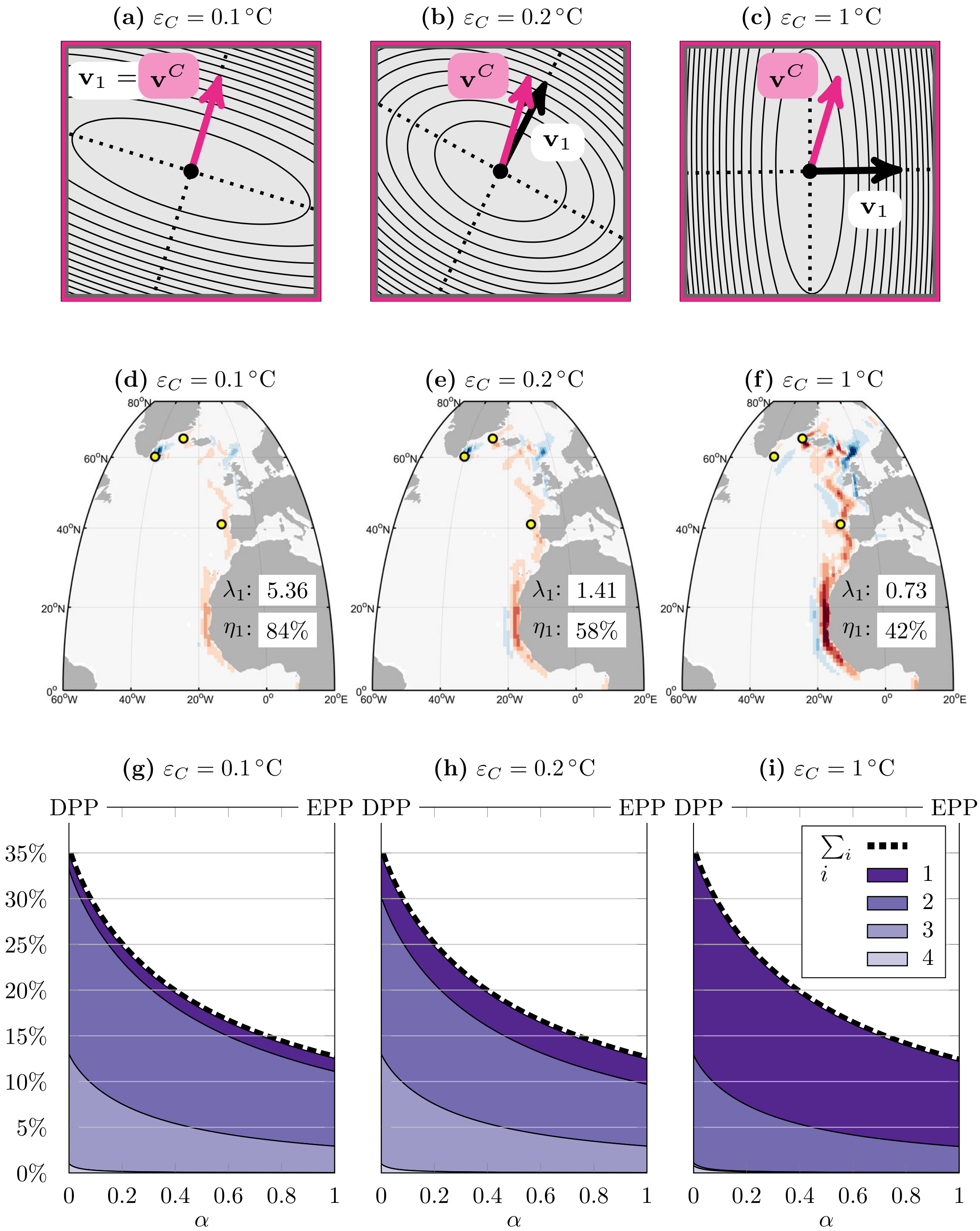

Figure S1. Figure caption on following page.

Figure S1. The panels in the middle column, i.e., (b),(e),(h), coincide with Figs. 8(a),(c), and Fig. 9(d). The left and the right column are as the middle column, but for different choices of ε_C. (a)-(c): Orientation of the first eigenvector, $\mathbf{v}_1$ (black vector) within the $\{\theta^A, \theta^C\}$-informed subspace of the control space. The ellipses show the contour lines of $J_{\mathrm{misfit}}(\tilde{\mathbf{u}})$. The larger ε_C, the more $\mathbf{v}_1$ deviates from $\mathbf{v}^C$, from $\mathbf{v}_1 = \mathbf{v}^C$ in (a) to $\mathbf{v}_1$ being almost orthogonal to $\mathbf{v}^C$ in (c). (d)-(f): τ_y component of $\mathbf{v}_1$. The inlets show the corresponding eigenvalue λ_1, as well as the associated effectiveness $\eta_1 = \lambda_1/(\lambda_1 + 1)$. The larger ε_C, the more $\mathbf{v}_1$ reflects the characteristic sensitivity patterns of the subsurface observations, concentrated along the eastern and northern boundary of the North Atlantic (cf. Fig. 4), but the lower λ_1 and η_1. (g)-(i): Decay in proxy potential for the QoI, $\mathrm{HT}_{\mathrm{ISR}}$, as a function of α. For all three cases, the DPP ($\alpha = 0$) is equal to 35.0%. From left to right, the EPP ($\alpha = 1$) decreases slightly (although almost negligibly) from 12.6% in (g) to 12.2% in (i). From left to right, the main contribution to proxy potential shifts from the second to the first eigenvector.