Reviewer #3

- *This paper investigates the key criterion for ocean state estimate, which is commonly called data assimilation (DA) in the oceanography and meteorological communities. There have been a lot of theoretical research and development over the past 30 years, the number of literature is just too many to list. The Lagrange multiplier method is just another way to express the minimization problem presented in the traditional 4D- Var. This can be found out Andrew Bennetts 1992 book: Inverse Methods in Physi- cal Oceanography. Cambridge Monographs on Mechanics and Applied Mathematics. Cambridge University Press.*

  The reviewer is clearly correct in their assessment of the relationship between 4D-VAR and the Lagrange multiplier method. We prefer the notation, "Lagrange multiplier method," because it will hopefully be understood by scientists in the greater mathemathics and physics communities. We include a sentence in the Introduction that shows that these terms are interchangeable.

  > The Lagrange multiplier method (e.g., *Thacker and Long*, 1988; *Wunsch*, 2010), sometimes called the adjoint method (e.g., *Hall et al.*, 1982; *Tziperman and Thacker*, 1989), "4D-VAR" (e.g., *Courtier et al.*, 1994; *Ferron and Marotzke*, 2003), or variational data assimilation (e.g., *LeDimet and Talagrand*, 1986; *Bonekamp et al.*, 2001; *Bennett*, 2002), is a method that satisfies both of these criteria, unlike the Kalman filter (e.g., *Fukumori and Malanotte-Rizzoli*, 1995) or nudging techniques (e.g., *Malanotte-Rizzoli and Tziperman*, 1995).

  The reviewer's point that the Lagrange multiplier method is just another way to express a minimization problem is in line with the point of our Section 4.1. In that section, we suggest that the criteria for successful use of the Kalman Filter/Smoother, which also minimizes the same cost function, is the same as that for the Lagrange multiplier method.

- *The reviewer does not agree the statement: the dimensionality of many million state variables is not a a fundamental problem. I think both high dimensionality and nonlinearity of ocean models are challenging issues for the ocean prediction and data assimilation. The controllability in this article is very vaguely defined. In fact the importance of boundary condition in ocean state estimation has long been recognized.*

  Some of the sentences in the Introduction may have given the impression that we are the first to consider the ocean state estimation problem as a time-variable boundary control problem. Given our long list of references on the topic, this is obviously not true! We hope to strike a better balance by revising the final paragraph of the Introduction as follows.

  > As has been documented in detail by many authors including the textbook of *Bennett* (1992), the ocean state estimation problem is better described as a time-variable boundary value problem because synoptic atmospheric variability acts as an external forcing on the ocean (Section 2).

  The Estimating the Climate and Circulation of the Ocean (ECCO) Consortirum has specifically focused on large-scale ocean state estimation with ocean models of coarse-enough resolution that they are essentially linear (e.g., *Stammer*, 1997; *Stammer et al.*, 2002). In this context, *Wunsch and Heimbach* (2007) stated, "The main issue for the oceanographic problem is one of dimension," in accordance with the reviewer's comment. The success of the ECCO Consortium, insofar as a solution can be found that fits the ocean data, indicates that the dimensionality of the problem can be overcome and therefore is not a fundamental obstacle. Here we revise the sentence.

  > Research conducted by the ECCO (Estimating the Circulation and Climate of the Ocean) Consortium (*Stammer et al.*, 2002, 2004) has demonstrated that (1), the dimensionality of many million state variables presents a challenge, but, insofar as a solution can be found that fits the ocean data, it can be overcome and it is does not pose a fundamental obstacle.

  Observability and controllability conditions for nonlinear state estimation are difficult problems. Results for certain nonlinear systems are found in the book by *Casti* (1985) (which evolves from the review paper of *Casti* (1982)). For general nonlinear estimation problems, the down-gradient-based iterative optimization is likely one of the best methods, and we have shown the relevance of controllability to the iterative process in our example. We include the following new paragraph in the Discussion that also expands upon observability and controllability.

> To recover the true trajectory of a system, observability is also important, as the estimation problem is the dual of the control problem (*Fukumori et al.*, 1993; *Marchal*, 2014). For the linear problem, *Cohn and Dee* (1988) showed that completely observability implies asymptotic stability of the Kalman filter/smoother. Defining observability and controllability conditions for nonlinear state estimation problems is difficult *Casti* (1985). In practice, the important criterion is ability to solve equation (10). Strictly speaking, the solution criteria will therefore depend upon both the controllability matrix, **C**, and the observational matrix, **E**, which combines the issues of observability and controllability. Here, we suggest the operational definition that a system is effectively controllable when the solution to (10), generalized to multiple observations, exists.

- *My view is that the system the authors have employed for the study and the observations are too simplified to draw valuable conclusions for the DA research and development communities. The reviewer recommends a major revision to include more realistic models.*

Our experimental setup may not be as simple as the reviewer understood. We note that stochastic noise is used to generate the synthetic observations, mimicking the imperfect nature of ocean observations. This is a common approach that has appeared in ocean state estimation studies such as Tziperman et al. (1992).

Regarding the simplified model of this study, it has been explicitly chosen with the motivation and goals of the manuscript in mind. Much time and effort has been spent to develop the Lagrange multiplier method in real-world scenarios, yet it is unclear whether this method should be applied to eddy-resolving models and how long the time window should be. For the Lagrange multiplier method to be successful in state-of-the-art ocean models, two major issues need to be addressed: (1) the high dimensionality of the forward model and estimation problem, and (2) the nonlinearity of ocean models at increasingly fine resolution. Issue (1) has been overcome by groups such as the ECCO Consortium, leading us to focus on (2).

With this problem in mind, it is logical to find a numerical model that can be thoroughly understood and one that is highly nonlinear. It is not the goal of this manuscript to use a state-of-the-art numerical model. We believe that these expectations should be set at the outset, so we include the following in the Introduction.

> Because the effect of nonlinearity is seen as the major roadblock for application of the Lagrange multipler method, we isolate this effect by choosing a model that is highly nonlinear but low-dimensional: the forced, chaotic pendulum (Section 2). Toy models are worth revisiting because the dynamics are comparatively simple to understand, and they have strongly influenced when the Lagrange multiplier method has been deployed to realistic ocean problems. We will show that previous toy models have sometimes been misinterpreted.

We now also emphasize upfront that the development of a new state-of-the-art data assimilation technique is not the goal of this work, either. Instead, we wish to evaluate the current use of the Lagrange multiplier method. Now, the Introduction makes this explicit.

> Rather than developing a new state-of-the-art data assimilation technique, we proceed by taking the existing Lagrange multipler method and developing diagnostics regarding when

> and why it succeeds or fails, as evaluated by the ability to fit observations. Relative to the initialization problem, the prospects for a successful state estimate are shown to be improved in the boundary control problem, even if one uses a highly nonlinear model such as the forced, chaotic pendulum (Section 3).

The typical criterion for successful state estimation has been the stability to initial perturbation. In the manuscript, we provide a counterexample showing that state estimation can be successful for an unstable system when it is controllable. This is one novel result we are reporting, and the previous works suggested by the reviewers have not already made this point, nor do they appear to contradict it.

One aspect of the analysis that we have improved is the first-guess of the forcing field. In a case study where the first-guess of the forcing is zero, the results are similar to the original case. This is reported in a new Section 3.5 and a new Figure 8.

Without doubt, it is a worthy goal to demonstrate the importance of controllability for the Lagrange multiplier method with a more complex geophysical model. Based on the results of this manuscript, a thorough test with that type of model is a logical next step, as suggested in Sec. 5.

- *The presentation also needs improvement, there are some sentences that are either not correct or clear to the reader. e.g. Page 2, line 34-35 is not correct.*

  We believe lines 34-35 to be grammatically correct as originally formulated.

- *Page 3, equation (1). It is stated that $\omega_d = 2/3$, so the right hand side forcing $f(t)$ is just a cosine function of time, I cannot see two independent variables (ω, θ) in this equation.*

  Certainly the angular velocity and displacement are related by a time derivative and are not independent. Those two variables describe the state of the model, and their time tendencies depend on one another, constituting an algebraic system of equations (discrete version of a system of differential equations). We find no mention of independence in the text and we do not believe that the dependence impacts the relevance of this model to state estimation.

- *Page 5, lines 26-27: Kalman filter equation is normally solved by in lower space with the covariance represented by ensembles, there is no need for explicit representation.*

  The reviewer makes a nice point that we have now included in the manuscript:

  > One remedy is to solve the Kalman filter equation in reduced space with the covariance represented by ensembles rather than being explicitly represented. Instead, we design a whole-domain method that is computationally efficient and provides a good first guess for the boundary control problem.

- *Page 14, equations (A1), (A2). Omega is the time derivative of theta. They are not completely independent.*

  We agree with the reviewer's statement, but admit that we are unclear as to the larger relevance of this statement.

- *Page 25, the 2nd to the last sentence in Figure 7 caption is clear not a correct sentence.*

  We have revised the sentence to the following.

  > The local minimum in the first two dimensions is no longer an extremum in the other two dimensions or the combined three-dimensional space.