Response to Reviewer's Comments

We thank the editor and the two referees for thoroughly reading the manuscript and for their helpful comments. We are very pleased to see many positive remarks. After we have addressed his/her concerns, the first referee recommended accepting this manuscript as it is. Also, we are happy to see that the new referee said "I see a lot of potential in this approach". In light of the comments, we have made a thorough revision addressing all major concerns, resulting in a significantly improved paper version.

# 1 Main concerns of the third referee

1. *It is said in the abstract that machine learning approaches are used in the current manuscript. I do not see any place where such tools are used, unless I am mistaken. This is misleading for the reader. This should be revisited in the whole manuscript.*

   Thanks for pointing out this problem. We have added the explanation in the revised manuscript. Followed by the first sentence in the abstract (with the superscript label 1 (**line 3**)), we add the explanation for machine learning approaches as

   - Generally, the statistical machine learning techniques refer to the marriage of traditional optimization methods and statistical methods, or says, stochastic optimization methods, where the iterative behavior is governed by the distribution instead of the point due to the attention of noise. Here, the sampling algorithm used in this paper is to numerically implement the stochastic gradient descent method, which takes the sample average to obtain the inaccurate gradient

2. *In Section 3, the method is outlined based on some demonstration of theorems, but there is no information on how the approach is implemented and compared with traditional methods. Please revisit completely this aspect in order to show the way to follow in order to develop the method by others.*

   Thanks for the good suggestion. We have added the content on how the approach is implemented and compared with traditional methods in **Section 3**. (**Line 130**)

   - In the numerical computation, we obtain the approximate gradient, $\nabla \hat{J}(u_0)$, via the sampling as
   $$\frac{d}{\epsilon} \cdot \mathbb{E}_{v_0 \in \mathbb{S}^{d-1}} \left[ J(u_0 + \epsilon v_0) v_0 \right] \approx \frac{d}{n\epsilon} \sum_{i=1}^{n} J(u_0 + \epsilon v_{0,i}) v_{0,i}, \tag{1.1}$$
   where $v_{0,i} \sim \mathrm{Unif}(\mathbb{S}^{d-1}),\ (i = 1, \ldots, n)$ are the independent random variables following the identical uniform distribution on $\mathbb{S}^{d-1}$. Since the expectation of the random variable $v_0$ on the unit sphere $\mathbb{S}^{d-1}$, we generally take the following way with better performance

in practice as

$$\begin{aligned}
\frac{d}{\epsilon} \cdot \mathbb{E}_{v_0 \in \mathbb{S}^{d-1}} \left[ J(u_0 + \epsilon v_0) v_0 \right] &= \frac{d}{\epsilon} \cdot \mathbb{E}_{v_0 \in \mathbb{S}^{d-1}} \left[ \left( J(u_0 + \epsilon v_0) - J(u_0) \right) v_0 \right] \\
&\approx \frac{d}{n\epsilon} \sum_{i=1}^{n} \left( J(u_0 + \epsilon v_{0,i}) - J(u_0) \right) v_{0,i}, \qquad (1.2)
\end{aligned}$$

where $v_{0,i} \sim \mathrm{Unif}(\mathbb{S}^{d-1}),\ (i = 1, \ldots, n)$ are independent. From (1.2), $n$ is the number of samples and $d$ is the dimension. Generally in practice, the number of samples is far less than the dimension, $n \ll d$. Hence, the times to run the numerical model is $n + 1 \ll d + 1$, which is the times to run the numerical model via the definition of the numerical method as

$$\frac{\partial J(u_0)}{\partial u_{0,i}} \approx \frac{J(u_0 + \epsilon e_i) - J(u_0)}{\epsilon},$$

where $i = 1, \ldots, d$. For the adjoint method, the gradient is numerically computed as

$$\nabla J(u_0) \approx M^\top M u_0 \approx M^\top g^T (U_0 + u_0),$$

where $M$ is a product of some tangent linear models. Practically, the adjoint model, $M^T$, is hard to develop. In addition. we cannot obtain the tangent linear model for the coupled ocean-atmosphere models.

3. *I have really hard time to read and understand the English. It should considerably be improved to make the results understandable.*

Thanks. We have improved considerable places for the English expression in the revised version, with bold text as a marker for editor and reviewer to track the changes.