Review of Excursion Probabilities for Informative Sampling

Constructive comments:

- I know very little about the AUVs you discuss, and I suspect many of your readers will be in a similar position. Additional contextual information would be helpful:
    - What sort of onboard computing do they have? It may not be something like an Intel 2.1 GHz, 4-thread processor, but a more informative description of something most readers relate to would be helpful. Similar information related to RAM and hard drive space would be helpful. What are the computational limits of these machines?
    - How long do the AUVs typically explore in one go? Is there a minimum/maximum time?
    - What kind of movements can the AUVs make? Chess-like moves (horizontal, vertical, diagonal)?
    - Is there a minimum/maximum number of points you want to sample?
    - What is the size of the grid that is typically considered? i.e., what is a typical value of n in Eq. (9)?
    - Does the AUV use a selection algorithm from the very beginning or is there a minimum number of initial measurements that are made before the optimization algorithm begins?
- p. 9, Eq. (8): the dependency on $\boldsymbol{x}$ and $\boldsymbol{x}$' has been omitted for some of the parameters without mention. Or are the parameters constant across locations?
- p. 9, How reasonable is the assumption of stationary and isotropic random fields? You may be limited by the computational abilities of the AUVs so that more complex models are computationally infeasible. That is a valid reason to choose a simpler, less accurate model. I'm just wondering if that is the reason here.
- p. 9, Eq. (10). Are the temperature and salinity processes at a specific location effectively independent of one another? (I truly don't know). That is what is implied by Eq. (10).
- p. 9, Eq. (11). The form of the equation is reminiscent of the Sherman-Morrison-Woodbury solution for fixed-rank kriging. Σ may not be small here, but I wonder if there may be a computational advantage to considering an alternative computational solution. See e.g., Cressie, N. and Johannesson, G. (2008), Fixed rank kriging for very large spatial data sets. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 70: 209-226. doi:10.1111/j.1467-9868.2007.00633.x.
- p. 9, Eq. (13). Why does $\boldsymbol{m}$ have a distribution before seeing the data? I can see how $\boldsymbol{m}(\boldsymbol{y})$ would have a distribution since it depends on $\boldsymbol{y}$. It almost feels Bayesian, but nothing indicates that this analysis is being pursued from a Bayesian perspective. Why is it valid for $\boldsymbol{m}$ to have a distribution prior to seeing data?
- p. 11, Eq. (17). I'm not understanding the advantage to conditioning on $\boldsymbol{m}$ instead of $\boldsymbol{y}$? $\boldsymbol{m}$ is a function of $\boldsymbol{y}$. The probability is the same two-dimensional integral either way. It seems to be related to the standardization that results in $\boldsymbol{Z}$? Please clarify.
- p. 11, Is $\boldsymbol{\pi}(\boldsymbol{m})$ the density associated with the distribution in Eq. (14)?
- p. 11, Is it purely for stylistic reasons that $P(\xi_T \leq \boldsymbol{t}_T, \xi_S \leq \boldsymbol{t}_s|\boldsymbol{m})P(\xi_T \leq \boldsymbol{t}_T, \xi_S \leq \boldsymbol{t}_s|\boldsymbol{m})$ isn't written as $P(\xi_T \leq \boldsymbol{t}_T, \xi_S \leq \boldsymbol{t}_s|\boldsymbol{m})^2$? It seems to be a more natural way of expanding Eq. (18).
- p. 11, It is stated that $\boldsymbol{Z}$ is chosen to be independent of $\boldsymbol{m}$, so presumably the joint CDF $F_{z,m} = F_Z F_m$, which would certainly be convenient, but is not obvious to me why this would be true.

Instead, it appears that $\boldsymbol{Z}$ is chosen to be a pivotal quantity to get the simple CDF in Eq. (20). Please clarify.

- p. 11, Eq. (21). The trailing $p(\boldsymbol{m})$ should be $\pi(\boldsymbol{m})$. Should the right side of the equality be squared?
- This method relies on computing the marginal probabilities in each grid cell. This is computationally convenient and is likely required because of the limited computing power of the AUVs. However, the optimization should technically be performed with respect to the joint distribution across all grid cells. While the context may not allow this in practice, it would be helpful to get an idea of how much error is introduced by this approximation. I recommend creating a small but realistic-as-feasible simulation study comparing the accuracy of the probabilities, a comparison of the computational expense, and examining if/how navigation decisions change based on whether one uses the marginal/joint computations.
- Fig. 7 is difficult to understand. The goal is to show the various paths taken by the AUV in the simulation example. But it is very difficult to determine which paths are very common and which are uncommon. Some ideas that may make this graphic more useful:
    - Don't worry about the paths but focus on the nodes visited. Represent each node using a circle or hexagon.
    - Make the node shading darker depending on whether that node is visited in a simulation. Nodes with no visits would be white. Nodes with 100 visits would be black.
    - The true excursion set is a random variable, but it would be informative to overlay the average position of the excursion set over each of the panels so the reader can see how the simulated AUV moves back and forth across the excursion set.
- While the context doesn't allow full reproducibility of the paper's results, no implementation details are provided. What software was used to perform the bulk of analysis? Are there scripts available that can be adapted by others for similar purposes? Please provide resources for the practical implementation of the methodology discussed in the paper.

Minor comments:

- "closed-form" is sometimes written as "closed form" (e.g., p. 7, p. 18)
- p. 9, \`1s\` should be '1s'. Similar for 0s.
- p. 15, my manuscript has -on where to sample-, which is an unusual use of dashes.