Overall details

1. You must keep everything relating to the review process confidential:

- Do not talk to anyone else (including other participants of the experiment) about the paper that is assigned to you. Keep all the information about the paper assigned to you (including the title) private.
- Do not use ideas and results from submissions in your own work until they become publicly available (e.g., via a technical report or a published paper).

2. The review is double-blind. Under no circumstances should you attempt to find out the identities of the authors for any of your assigned submissions (e.g., by searching or arXiv).

3. **You should review a paper as if it was submitted to the real ICML conference**. However, for this experiment we assume that all the submissions are in the scope of the ICML conference, are properly anonymized and do not violate formatting guidelines. **Do not consider these criteria as yardsticks for evaluations.**

4. For this experiment you are not given access to any supplementary materials except appendices (where applicable). If a paper makes any claim about other supplementary material (i.e. refers to the attached video or dataset), you should trust this claim.

How to review? Part 1. Theory of a good review

In the review you will need to (i) identify three contributions the paper makes and (ii) write a detailed comment evaluating the paper on four criteria: **originality, quality, clarity,** and **significance**.

0. Keep in mind

Novel and/or interdisciplinary works (e.g., which are not incremental extensions of previously studied problems but instead perhaps formulate a new problem of interest) are often very easy to criticize, because, for example, the assumptions they make and the models they use are not yet widely accepted by the community (due to novelty). However, these works may be of high importance for the successful development of the field in the long run, so please try to be aware of this bias, and avoid dismissive criticism. When reviewing a paper, always think about the impact the work may have on the community.

1. Contributions.

Contributions may be theoretical, methodological, algorithmic, empirical, connecting ideas in disparate fields ("bridge papers"), formulating a novel problem or providing a critical analysis (e.g., principled justifications of why the community is going after the wrong outcome or using the wrong types of approaches.). One measure of the significance of a contribution is (your belief about) the level to which researchers or practitioners will build off of or use the proposed ideas. **Papers that explore new territory or point out new directions for research are preferable to papers that advance the state of the art, but only incrementally.**

Not all good papers will have three contributions. For example, a ground-breaking theoretical paper might simply contribute the key theorem and proof (Significance: High). However, for such a paper, hopefully you could also list “Presented a unified and extended view of several existing results. Significance: Medium.” or “Provided new proof path. Significance: High.”

Please remain polite in this section and avoid writing “This paper did not contribute any new ideas.” Instead write something along the lines of “Authors proposed model that primarily combines models in [cite A] and [cite B]. Significance: Low.”

2. Detailed comments.

Your comments should begin by summarizing the main ideas of the submission and relating these ideas to previous works in the area. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: **originality, quality, clarity, and significance**. Ideally (but not obligatory), you should have a separate paragraph for each criteria.

Your comments should be detailed, specific, and polite. Please avoid vague, subjective complaints. Imagine your paper receiving an unfair, unjustified, short, or dismissive review. Try to not be that reviewer! Always be constructive and help the authors understand your viewpoint, without being dismissive or using inappropriate language. Remember that you are not reviewing your level of interest in the submission, but its scientific contribution to the field!

Importantly, please comment on and take into account the strengths of the submission. It can be tempting to only comment on the weaknesses; however, ACs, SACs, and program chairs need to understand both the strengths and weaknesses in order to make an informed decision. *It is useful for the ACs, SACs, and program chairs if you include a list of arguments for and against acceptance.*

How to review? Part 2. Examples of a good review

1. Contributions:

The following are examples of contributions a paper might make. This list is not exhaustive.

- “The paper provides a thorough experimental validation of the proposed algorithm, demonstrating much faster runtimes without loss in performance compared to strong baselines.”
- “The paper proposes an algorithm for [insert] with computational complexity scaling linearly in the observed dimensions; in contrast, existing algorithms scale cubicly.”
- “The paper presents a method for robustly handling covariate shift in cases where [insert assumptions], and demonstrated the impact on [insert application].”
- “The authors provide a framework that unifies [insert field A] and [insert field B], two previously disparate research areas.”
- “This paper demonstrates how the previously popular approach of [insert] has serious limitations when applied to [insert].”
- “This paper formulates a novel problem [insert description] and clearly shows that it is of interest to the community because [explain why]”

2. Detailed comments:

a) Quality

Try to answer the following questions: Is the submission technically sound? Are claims well supported by theoretical analysis or experimental results? Is this a complete piece of work or work in progress? Are the authors careful and honest about evaluating both the strengths and weaknesses of their work?

- Example:
  "The technical content of the paper appears to be correct albeit some small careless mistakes that I believe are typos instead of technical flaw (see #4 below).
  …
  4. The equation in line 125 appears to be wrong. Shouldn't there be a line break before the last equal sign, and shouldn't the last expression be equal to E_q[(\frac{p(z,x)}{q(z)})^2]?"
  "The idea of having a sandwich bound for the log-marginal likelihood is certainly good. While the authors did demonstrate that the bound does indeed contain the log-marginal likelihood as expected, it is not entirely clear that the sandwich bound will be useful for model selection. This is not demonstrated in the experiment despite being one of the selling point of the paper. It's important to back up this claim using simulated data in experiment."
- Example:
  "The paper is generally well-written and structured clearly. The notation could be improved in a couple of places. In the inference model (equations between ll. 82-83), I would suggest adding a frame superscript to clarify that inference is occurring within each frame, e.g. q_{\phi}(z_2^{(n)} | x^{(n)}) and q_{\phi}(z_1^{(n)} | x^{(n)}, z_2^{(n)}). In addition, in Section 3 it was not immediately clear that a frame is defined to itself be a sub-sequence."

b) Clarity

Try to answer the following questions: Is the submission clearly written? Is it well organized? (If not, please make constructive suggestions for improvement.) Does it adequately inform the reader? (A superbly written paper provides enough information for an expert reader to reproduce its results.)

- Example:
  "While the paper is pretty readable, there is certainly room for improvements in the clarity of the paper. I find paragraphs in section 1 and 2 to be repetitive. It is clear enough from the Introduction that the key advantages of CHIVI are the zero avoiding approximations and the sandwich bound. I don't find it necessary to be stressing that much more in section 2. Other than that, many equations in the paper do not have numbers. The references to the appendices are also wrong (There is no Appendix D). There is an extra period in line 188."
- Example:
  "The paper is generally well-written and structured clearly. The notation could be improved in a couple of places. In the inference model (equations between ll. 82-83), I would suggest adding a frame superscript to clarify that inference is occurring within each frame, e.g. q_{\phi}(z_2^{(n)} | x^{(n)}) and q_{\phi}(z_1^{(n)} | x^{(n)}, z_2^{(n)}). In addition, in Section 3 it was not immediately clear that a frame is defined to itself be a sub-sequence."

c) Originality

Try to answer the following questions: Are the tasks or methods new? Is the work a novel combination of well-known techniques? Is it clear how this work differs from previous contributions? Is related work adequately cited? Does the work formulate a novel problem?

- Example:

  "The main contribution of this paper is to offer a convergence proof for minimizing sum fi(x) + g(x) where fi(x) is smooth, and g is nonsmooth, in an asynchronous setting. The problem is well-motivated; there is indeed no known proof for this, in my knowledge.

  …

  There are two main theoretical results. Theorem 1 gives a convergence rate for proxSAGA, which is incrementally better than a previous result. Theorem 2 gives the rate for an asynchronous setting, which is more groundbreaking."

- Example:

  "The paper is missing a related work section and also does not cite several related works, particularly regarding RNN variants with latent variables (Fraccaro et al. 2016; Chung et al. 2017), hierarchical probabilistic generative models (Johnson et al. 2016; Edwards & Storkey 2017) and disentanglement in generative models (Higgins et al. 2017). The proposed graphical model is similar to that of Edwards & Storkey (2017), though the frame-level Seq2Seq makes the proposed method sufficiently original. The study of disentanglement for sequential data is also fairly novel."

d) Significance

Try to answer the following questions: Are the results important? Are others (researchers or practitioners) likely to use the ideas or build on them? Does the submission address a difficult task in a better way than previous work? Does it advance the state of the art in a demonstrable way? Does it provide unique data, unique conclusions about existing data, or a unique theoretical or experimental approach?

- Example:

  "I liked this article very much. It answers a very natural question: gradient descent is an extremely classical, and very simple algorithm. Although it is known not to be the fastest one in many situations, it is widely used in practice; we need to understand its convergence rate. The proof is also conceptually simple and elegant, and I found its presentation very clear."

- Example:

  "There are several things to like about this paper:

  - The problem of safe RL is very important, of great interest to the community and without too much in the way of high quality solutions.

  - The authors make good use of the developed tools in model-based control and provide some bridge between developments across sub-fields.

  - The simulations support the insight from the main theoretical analysis, and the algorithm seems to outperform its baseline.

However, I found that there were several shortcomings:

- I found the paper as a whole a little hard to follow and even poorly written as a whole. For a specific example of this see the paragraph beginning 197.

- The treatment of prior work and especially the "exploration/exploitation" problem is inadequate and seems to be treated as an afterthought: but of course it is totally central to the problem! Prior work such as [34] deserve a much more detailed discussion and comparison so that the reader can understand how/why this method is different.

- Something is confusing (or perhaps even wrong) about the way that Figure 1 is presented. In an RL problem you cannot just "sample" state-actions, but instead you may need to plan ahead over multiple timesteps for efficient exploration.

- The main theorems are hard to really internalize in any practical way, would something like a "regret bound" be possible instead? I'm not sure that these types of guarantees are that useful.