**General Comments:**

In this paper, we present a nonlinear mixture autoregressive model (MixAR) that attempts to directly model nonlinearities in the trajectories of the speech features. Performance benefits are demonstrated across several databases including TIMIT, NTIMIT, and NIST-2001. We demonstrate that MixAR, using only half the number of parameters and only static features, can achieve a lower equal error rate when compared to GMMs, particularly in the presence of previously unseen noise. This is the first demonstration of improved performance on a significant speech recognition task for this approach, and demonstration of its efficacy across several databases makes the paper unique.

Below we address the specific concerns of the reviewers:

**Reviewer #2:**

**FIGURE ONE IS INCOMPLETE (variables wi and gi)**

The relationship of wi and gi to the weights on the Gaussian mixtures is indirect. We have included an equation explaining this relationship, and redrawn the figure to explicitly show these dependencies. Hopefully it is clearer now.

**Reviewer #3:**

**The authors present a nonlinear mixture autoregressive model and its application to the problem of speaker verification. The paper is well organized and the theoretical explanation is concise and technically sound. The authors conducted extensive experiments to illustrate the advantages of the presented model over the commonly used GMM model. However, in experiments the evaluation of the GMM model was not conducted on the state-of-the-art GMM system, what authors also mention in the summary. The authors should mention this fact already in the sections III and IV and should discuss the results emphasizing the performance limitations of the implemented GMM, which was not the state-of-the art system. Therefore, further elaboration of the section III and especially IV is needed to adequately include this aspect in the discussion.**

We were careful not to use the term “state of the art” when describing GMMs. State of the art, of course, can mean many different things. The baseline system our work is based on uses the same number of mixture components per state. We understand not everyone does this today, but many practical systems are still constrained this way. We added some discussion of this at the beginning of Section III, and we reinforce this notion in the summary.

**The curves in the Figures 4, 5 and 6 are not readable (the curves cannot be distinguish) if the article is BW printed. To provide readable figures also in black/white layout of the article the authors may consider to change line types of the curves in Figures 4 - 6.**

We no longer have the ability to reproduce these figures. They are more legible in color than B&W. We will do our best to provide improved figures when the final color figures are requested.

**On page 3 last line the Section VVI should be changed to Section VI.**

Fixed.

**Summary Comments:**

We would like to thank the reviewers for their thoughtful and insightful feedback. We have done our best to address each of their concerns. We hope our responses have adequately addressed your concerns and we can proceed with publication.