We thank the reviewers for their insightful feedback. We first address a general concern on numerical experiments, before replying to reviewer-specific comments.

Numerical experiments (R2, R3, R4). A shared comment by most referees is the absence of numerical experiments. While the paper is theoretical in nature, and we initially chose to postpone in-depth numerical experiments for future work for space reasons, we agree that simple experiments would be helpful to illustrate the theory. We will therefore include simple synthetic experiments in the final version as an illustration of our theoretical findings. As a preliminary study, Figure 1 illustrates (left) the convergence of a permutation-invariant GCN with random weights on simple random graph models for different sparsity levels, and (center-right) its stability for a particular simple deformation of three-dimensional latent variables. More refined experiments of this type, as well as the code, will be included in the updated version of the paper.

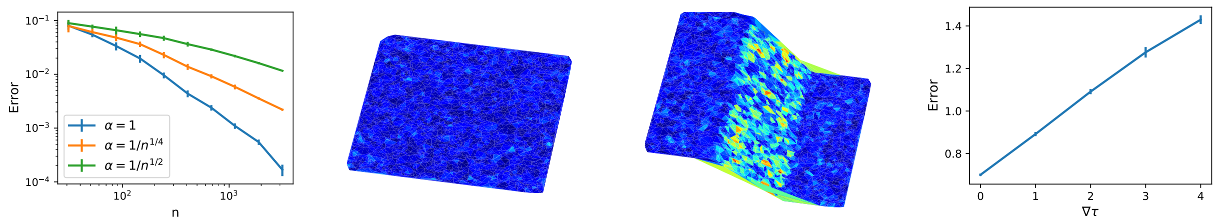

Figure 1: (left) Convergence of permutation-invariant GCNs to its continuous value (approximated using a very large n and $\alpha = 1$) on synthetic data. (middle-right) Illustration of stability to deformations. A Gaussian kernel is used to generate graphs from 3D latent positions on a surface. The deformation adds a ramp with height τ. The middle figures illustrate the difference in output of an equivariant GCN between several draws of the random graph, with the same or deformed latent variables. With a translation-invariant kernel, "flat" parts on both side of the ramp yield the same output, despite being at different heights.

R1. *About "implications [...] to practices" and the "superiority" of continuous GCN*: In this paper, our goal is to provide a novel theoretical analysis which may help explain the behavior of classical GCNs on typical large graphs, rather than to propose a new model: c-GCNs are only theoretical objects that characterize the limit of GCNs. While we indeed hope that our analysis will help derive guidelines for practices (for instance we mention helping to choose the input signal in the absence of node features, but this can go beyond), this is a bit out-of-scope of the present paper and left for future work. Such guidelines will likely come from in-depth analyses of each particular random graph model to which our analysis apply: community graphs, geometric graphs, etc. We will update the discussion in the final version of the paper. Finally, when we say "overcome the difficulties of dealing with [...] isomorphisms", we only meant in the theoretical analysis. We will rephrase this to make this point clearer.
"The overall paper is very dense [...].": The manuscript will be updated to use a unified notation for all assumptions.
Use of the symbol d, $diam(X)$ and grammar comment: We agree and will correct this in the update in order to improve clarity. We will rephrase line 193.
Additional references: We thank the reviewer for pointing out these works. GVAE indeed exploits classical latent-space random graph models (which date back all the way to the Erdös-Rényi model in the 50's) in the context of VAE. In future work, additional properties of c-GCNs, e.g. approximation power, could probably be derived in this case. We will include the suggested references.
R2. *Optimality of the bounds*: As pointed out by the referee, it is still unknown whether the bound in $(n\alpha_n)^{-1/2}$ is optimal, as moreover for the case where α_n is proportional to $\log n/n$, the convergence is extremely slow and difficult to observe in experiments. For the Wasserstein convergence in $n^{-1/d}$, lower bounds do exist however (see Weed and Bach's paper), even if some questions are still open. We will update the discussion.
R3. *Assumptions*: Concerning the main assumption, the random graph modelling, it is indeed a central question to assess how relevant they are to model real data. This however has a very long history in the literature, but we will add some key references. The convergence assumption in line 144 allows us to analyze filters of infinite order, but it is always satisfied for filters of finite order as commonly used in practice (e.g. in ChebNet). For the Euclidean assumption line 261, we think that this is a simple assumption allowing us to discuss deformations through the formalism of diffeomorphism, and it is valid for instance when considering graph models arising from 3D meshes. Both assumptions could be extended in future work to handle different scenarios, and we will update the discussion accordingly.
R4. *About the exposition*: We will do our best to improve the discussions of our results in order to provide more intuition, and hope the experiments presented above will help clarify our message on convergence and stability.
Additional references: Thank you for the additional references. Centrality measures are a key concept for graphons and are definitely related to some considerations in our paper, and we will add the references in the updated paper.