We thank all reviewers for their time, effort and constructive feedback on this paper.

Experimental validation **(R1,R2,R4).** We appreciate the reviewers' concerns w.r.t. the experiments. We believe that the contributions of this paper are primarily theoretical, namely, (i) showing that GNNs are transferable between deterministic graphs obtained from a graphon and (ii) non-asymptotically quantifying this transfer error. Due to space constraints, we only included limited numerical experiments to briefly illustrate these results in a more practical context. Still, to address the reviewers' points, we will use the extra space in the camera-ready to include an additional experiment using the citation network setting of (Kipf and Welling, 2017). We will also clarify in the manuscript that the user network in the MovieLens example is built from training data alone as in (Ruiz et al., 2019b). The RMSE was chosen for being a standard performance metric in collaborative filtering.

Feasibility of assumptions **(R1,R4).** To address the concerns of reviewer 4, we will add a discussion on the assumptions of Thms. 1–2 to the camera-ready distinguishing between technical (e.g., normalized filters) and binding assumptions (e.g., Lipschitz graphons) and highlighting their consequences and limitations. As for the point raised by reviewer 1, note that the assumption on the filters h is not contradicted by (NT and Maehara, 2019) and (Oono and Suzuki, 2020), since they consider the spectrum of the graph Laplacian rather than the adjacency matrix. Hence, $|\lambda| \approx 1$ corresponds to low frequencies and $|\lambda| \approx 0$ to high frequencies rather than the converse. Moreover, our results show that when the relevant features have low frequency components ($|\lambda| \approx 1$), a simple graph convolutional filter is transferable (Thm. 3 in the appendices). This is not the case for high frequency features ($|\lambda| \approx 0$). In contrast, a GNN is likely to remain transferable as it leverages nonlinearities to scatter high frequency components to lower frequencies.

Relevance and weaknesses of graphon framework **(R3, R4).** Reviewer 3 raises a good point in noting that the deterministic generative model we consider–which appears and is used throughout the seminal book on graphons (Lovász, 2012, Chapter 10)–cannot produce certain types of sparse graphs (e.g., with fixed degree). However, it doesn't necessarily yield complete graphs since $\mathbf{W}(u, v)$ can be 0 for some (u, v). While we agree with the reviewer the model has limitations, we believe that it still provides important insights into when and how transferability is possible. While it may be intuitive that GNNs converge on sequences of deterministic graphs, this is not necessarily the case. For instance, Thm. 3 shows that graph filters are only really transferable for large eigenvalues, i.e., "low frequency" features. This is evidenced by the hypothesis that $h(\lambda)$ is constant for $|\lambda| < c$, which addresses the fact that the graph eigenvalues accumulate near 0 and thus become harder to distinguish. GNNs, on the other hand, are expected to remain transferable even for high frequency features ($|\lambda| \approx 0$) because the nonlinearities scatter high frequency components to lower frequencies. These conclusions are not trivially obtained from the model. Naturally, as noted by reviewer 4, our analyses focus on dense graphs for which graphons are models. Convergence of sparse graphs is a topic of active research that requires considerably different entities (*graphings*) which are beyond the scope of this paper. Still, dense graphs have practical applications and provide insights as to the transferability of the convolution operators in GNNs. We will make these discussions more clear in the camera-ready version.

Transferability–discriminability trade-off **(R1).** This trade-off arises from n_c and δ_c, which are both related to the threshold c. In the interval $|\lambda| \in [c, 1]$, the filters $h(\lambda)$ do not have to be constant and can thus distinguish between eigenvalues. Hence, their discriminative power is larger for smaller c. The value of n_c counts the eigenvalues $|\lambda| \in [c, 1]$. When c is small, n_c is large, which worsens transfer error. The quantity δ_c is close to the minimum graphon eigengap $min_{c\in[\lambda_i,\lambda_j]}|\lambda_i - \lambda_j|$. Since the graphon eigenvalues accumulate near 0, δ_c approaches 0 for small c. This also increases the transfer error. Therefore, small values of c provide better discriminability but deteriorate transferability, and vice-versa. Note that for GNNs (Thm. 2) this trade-off is less rigid than for graph filters (Thm. 3 in the appendices) because GNNs leverage nonlinearities to scatter low frequency components ($|\lambda| \approx 0$) to larger frequencies ($|\lambda| > c$). These points will be added to the discussion of Thm. 2 in the paper. Finally, note that this trade-off is hard to verify experimentally, since in practice c can't be controlled as the filters h are learned by training the GNN. How to promote more high-pass filters during learning is an interesting avenue of research, but is beyond the scope of this paper.

Positivity of δ_c **(R1).** The reviewer raises an important point. The value of δ_c measures the minimum distance between the *eigenspaces* corresponding to the closest graph eigenvalue λ_i^n and graphon eigenvalue λ_j on either side of c, i.e. $\delta_c = min_{c\in[\lambda_i^n,\lambda_j]}|\lambda_i^n - \lambda_j|$ [cf. Prop. 2 in the appendices]. Since, $\lambda_i^n \to \lambda_i$, the convergent point of this quantity is the difference between two consecutive distinct graphon eigenvalues, which is always positive. As defined in the paper, however, δ_c only reflects this if $\lambda_i \neq \lambda_{i+\mathrm{sgn}(i)}$ for all i, i.e., if all nonzero eigenvalues have multiplicity 1 and thus each correspond to a distinct eigenspace. We will correct this definition to eliminate this requirement and clarify that δ_c converges to the smallest graphon eigengap $|\lambda_i - \lambda_j|$ such that $c \in [\lambda_i, \lambda_j]$, which is always positive for $c > 0$.

Related work on transferability **(R1).** We thank the reviewer for pointing us to (Keriven et al., 2020). In this paper, the convergence of GNNs to continuous GNNs is studied based on concentration inequalities relating the spectrum of a random graph Laplacian with that of the sampled graphs. Hence, their result depends on the graph sparsity. In contrast, we focus on dense graphs associated with graphons and study GNN transferability by analyzing the spectral behavior of graph convolutions. We will use the extra space in the camera-ready to expand on these distinctions.