A main concern shared by reviewers #1 and #3 is the lack of mathematical rigour. We address this in two points:

- The proposed parametrization of H_r should not be seen as a simple tuning of a hyper-parameter. It is on the contrary the only consistent possible choice, according to three arguments: i) ζ is the eigenvalue resulting from the linearization of BP around its trivial fixed point (see Eq. (11) of [9]) and the corresponding eigenvector g, once processed as $g_i^{\rm in} = \sum_{i\in\partial j} g_{ij}$ satisfies $H_\zeta g^{\rm in} = 0$, that is precisely the eigenvector studied in the present article; ii) the mapping to the Ising Hamiltonian, from which H_r was derived in the first place ([8]), is consistent only at $r = \zeta$ as explained in Section A of the supplementary material; iii) choosing $r = \zeta$ enables a resilience to degree heterogeneity in the DC-SBM, as developed in Section 2.2. These three arguments are independent and lead to the same parameter, creating a deep connection between B and H_ζ other than the graph Laplacian to which H_ζ tends in easy problems.
- We agree that most of our derivations are heuristic. This has nevertheless been the case in the last few years in this research domain: a first heuristic derivations of the results via different techniques – mainly based on tools from statistical physics – is a first necessary step before the mathematical formalization. The study of B in the sparse regime is a hot topic in research, and technically very challenging, see e.g. [10], [15], [16]. As suggested by reviewer #2, even fewer results are available on H_r, and this work, to our knowledge, represents a first attempt to characterize its eigenvectors. Finally, we also point out that our work already triggered new mathematical research (Coste, Zhu 2019: arXiv:1907.05603) in which some of our claims were formally proved.

We now address more specific concerns raised by each reviewer.

- Reviewer#1. Thank you for your detailed review. We hope that your concerns are fully addressed.

  Comparison with regularized Laplacian techniques of, e.g., (Qin 2013). Based on simulations, these methods have comparable performance to ours (ours never being worse). The regularized Laplacian, however, relies heavily on the normalization of the rows of the matrix containing the eigenvectors. Such normalization appears very powerful and in practice effective for many matrices. We believe that such step is not completely justified because i) the study of $||L_\tau - \mathscr{L}_\tau||$ is relevant when such quantity is small compared to the eigengap (Joseph, Yu 2014), that is, for simple classification problems. No guarantee is given for harder scenarios. ii) the detectability threshold is not mentioned in (Qin 2013) and it represents instead a fundamental aspect of our algorithm. We are currently working on a precise description of the connection between these two spectral techniques.

  In Eq. (5) the result is the expectation of the sum of Bernoulli random variables. This has been clarified in the new version. The independence comes from the tree like approximation: neighbours of a same node belong to conditionally independent branches. This is a standard technique used in BP (Mezard 2009), also formalized in e.g. (Salez 2011).

  In l120 we are taking a conditional expectation. The notation has been clarified in the new version. ν_p is the p-th smallest eigenvalue of H_r, as defined in l177. In l214-215, the vector $u^{(p)}$ corresponds to the p-th largest eigenvalue of $C\Pi$, denoted with τ_p , hence to the p-th smallest $\zeta_p = c/\tau_p$. Decreasing the value of r, starting from $r = \sqrt{c\Phi}$, the informative eigenvalues go from negative to positive, hitting zero at ζ_p. The k-th smallest will be the first, so it will correspond to ζ_k then the others follow, so the the p-th corresponds to ζ_p. This has been clarified in the new version.

  The statement $\tilde{\beta} = O(\beta_i)$ means that the random variable β_i has the same scaling (with respect to the average degree) than its expectation. The argument of Gaussianity is an assumption made on reasonable intuitions to conclude the calculus and is to be tested on the expression of the overlap compared to the simulations. Given the very good agreement, we understand that the approximations made to that point are reasonable and justify the description we made on the shape of the informative eigenvector.

  Figure 2 supp mat: $\hat{k}$ represents the number of classes estimated from our algorithm, while k_d is the number of classes that are theoretically detectable. The color scale plots the quantity $2(\hat{k} - k_d)/(\hat{k} + k_d)$ as a function of the actual number classes ($k \geq k_d$) and the hardness of the problem. When this quantity is zero (white), the algorithm has detected all the detectable classes. In the caption $\mathcal{U}(\cdot)^4$ stands for uniform distribution raised to power 4.
- Reviewer #2. Thank you for your very positive review.

  We agree that the term $(r^2 - 1)I_n$ doesn't affect the spectral properties, but it is necessary to make the connections with B and H_r. In order not to introduce a further matrix $(D - rA)$, we chose to write everything in terms of H_r for the sake of clarity. Your interpretation about the kernel is correct: in the new version this has been pointed out.

  To estimate ζ_p we need to recompute each time the first p eigenvalues (not the whole spectrum). This method is self contained in terms of H_r, but using the eigenvalues of B' ([9]) can be an alternative way to estimate ζ_p. We are currently working on an alternative and faster solution, based on a polynomial approximation.
- Reviewer #3. Thank you for your review. We hope that your concerns are fully addressed.

  The comparison of the performances for different values of r is certainly interesting and has been added to the new version. Note however that the spectral algorithm on the matrix B corresponds to $H_{(c_{\rm in} - c_{\rm out})\Phi/2}$ and has been seen in the literature (see e.g. [8]) to underperform the $H_{\sqrt{c\Phi}}$, consistently with the fact that $(c_{\rm in} - c_{\rm out})\Phi/2$ is farther away from ζ then $\sqrt{c\Phi}$.