- For ANA, the diffusion matrix $\boldsymbol{S}$ can be **precomputed before training**. The ultimate objective of ANA in Eq. (6) only involves simple matrix multiplication ($O(N^2d)$), which can be **efficiently parallelized for execution with GPUs**.
- For IC with **random projection**, its computation mainly involves **a projection operation (matrix multiplication)** and **computing variances along sample dimension**. For QR factorization, the computation is also efficient, whose performance can surpass projection based on eigen-decomposition. The results in Table 2 and can serve as evidence of efficiency of our method. In terms of practical application, ANA-IC (QR) is the most suitable choice.

Table 2

Algorithm	Cora		Citeseer		Pubmed		Computers	
	Time	Memory	Time	Memory	Time	Memory	Time	Memory
DGI	6.8s	3.8GB	9.4s	7.8GB	44.9s	11.2GB	71.2s	11.3GB
GRACE	5.1s	1.2GB	7.4s	1.5GB	1,169s	12.2GB	362.8s	7.4GB
MVGRL	23.7s	3.8GB	48.4s	7.9GB	2,010s	9.1GB	78.8s	16.6GB
ANA-IC (Random).	2.7s	2.5GB	2.0s	2.6GB	11.3s	4.8GB	5.3s	4.4GB
ANA-IC (Eigen)	56.7s	2.5GB	9.7s	2.6GB	110.4s	4.8GB	51.3s	4.3GB
ANA-IC (QR)	3.8s	2.5GB	2.5s	2.6GB	13.1s	4.8GB	6.2s	4.3GB

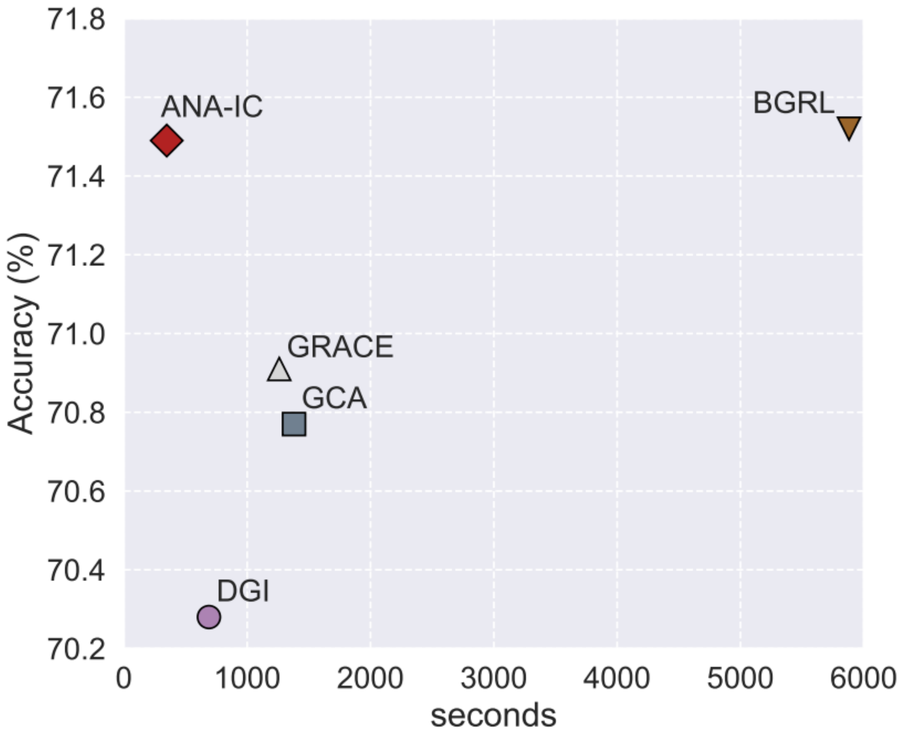

Figure 12: Test accuracy and training time on Ogbn-Arxiv.