Point-wise-response

Reviewer 1

MAJOR COMMENTS

1) Ground truth dataset
As shown in Method section by the authors (page 10), they used the HIPPIE (Human Integrated protein-protein interaction reference) dataset as the ground truth for single-cell expression. However, it has been well recognized that there might be no strong correlation between protein interactions and gene co-expression pairs (https://academic.oup.com/bioinformatics/article/21/11/2730/294834). Hence, it may be not proper to use the protein interaction dataset to assess the accuracy of predicted gene co-expression pairs.
Also, the authors mentioned that for the 4 datasets (3 real datasets and one simulated dataset) within DREAM5, all the ground truth data are available. It is easily to understand for the availability of simulated datasets, but it remains confusing to me how could you know exactly the "true" gene co-expressions for the other 3 real datasets.

We have used protein-protein interactions (PPI) as reference for evaluation for only one dataset (i.e single-cell expression profile from the pancreatic cells). Even for that dataset, we have also used another method based on the overlap of predicted gene-interactions in old and young cells to show improvement in network inference by our approach of graph-wavelet based filtering.

We partially agree with reviewer-1 that protein-protein interaction may not be an ideal ground truth for evaluating gene-interaction prediction. However, even in the manuscript cited by reviewer-1, it has been written

" For individual genomes, we have found that in *E.coli* there is a strong correlation between the expression profiles for interacting pairs when compared with random pairs, while in other species the correlation is only slightly more significant than random."

Thus protein-protein interaction is still better than random pairs, in terms of correlation with gene co-expression. Even in the study of benchmarking of multiple gene-network inference methods, Chen et al. (cited in our manuscript) has used PPI for evaluation of performance. (https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-018-2217-z)

Nevertheless, we have added a sentence to the corresponding location to make our statement softer

Even though like Chen et al. \cite{Chen-2018} we have used PPI to measure improvement in gene-network inference, it may not be reflective of all gene-interactions. Hence we also used

the criteria of increase in overlap among predicted networks for same cell-types to evaluate our method for scRNA-seq profiles of pancreatic cells.

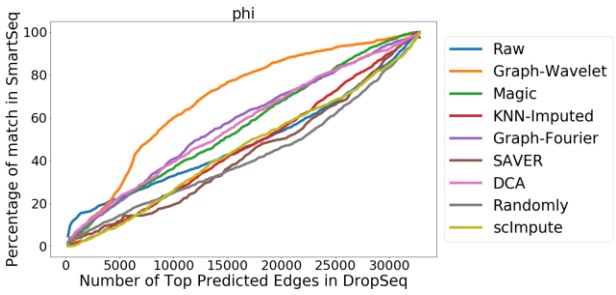
Reviewer-1 has pointed out that the golden sets of interactions provided by DREAM5 consortium are not reliable. There have been many publications (more than 25) where te same DREAM5 datasets and their golden set of interactions reference have been used to evaluate network inference (or gene co-expression) methods. The same 4 datasets and their corresponding golden sets of interactions have been used by Feizi et al. (Nature Biotechnology, 31 : 726–733(2013)) for evaluating their network deconvolution method for correlation and ARACNE based predicted-network (see Figure 2 in Feizi et al. Nature Biotechnology https://www.nature.com/articles/nbt.2635).

For DREAM5 dataset, true positive interactions were based on experimentally validated interactions from the RegulonDB database for *E. coli (Nucleic Acids Res. 39, D98–D105 (2011))*. For *S. cerevisiae,* a high-confidence set of interactions was made using genome-wide transcription-factor binding data (ChIP-chip) and evolutionarily conserved binding motifs (*BMC Bioinformatics* 7, 113 (2006)).

Overall we have used widely used procedure and datasets for evaluating network inference methods as they have been built after a lot of efforts of several scientific groups across the world.

2) Other denoising tools
It remains unclear to me why the authors did not compare the results from their filtering approaches with that from other previous denoising tools. The authors should first collect other denoising tools in the Introduction section, and then compare with results from these tools in the Result part.
What also confuses me is that whether the Graph Fourier is another denoising tool, or just one pre-processing step before Graph Wavelet? The authors should state it clearly in the main text.

In our manuscript, we have focused more on the novelty of our approach and its application in the analysis of single-cell expression profiles from ageing samples for impactful insights. However, on the suggestion of reviewer-1, we have now compared our method to 7 other methods for denoising/imputation of single-cell expression profile.
Here we show one result for mESC dataset, which we update on figure 2C.



Similarly, we have also updated supplementary Figure-S1 and Figure S2A with results of comparison with 7 other methods for denoising/imputation of single-cell expression profile.

Graph Fourier is another denoising approach where Chebyshev filter is used for low-pass filtering for denoising expression profiles. It was already mentioned and cited in the previous version of the manuscript. Graph Fourier based denoising is independent of our method. Now we have clearly mentioned it. Now we have clearly mentioned it in the results section. The text is as such

We compared Graph Fourier based low pass-filtering with graph-wavelet based denoising using three different approaches to threshold the wavelet-coefficients

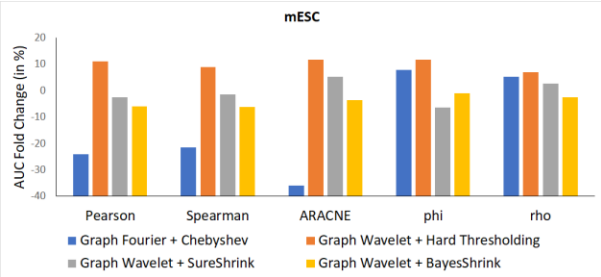
MINOR POINTS:

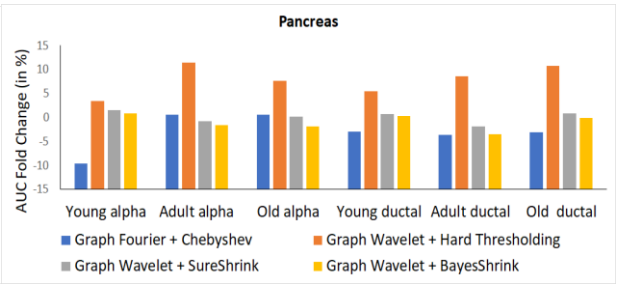
1) As the main focus of this study is to present a denoising tool for gene co-expression network analysis, I would suggest the authors should shorten the description on the application of this tool (line 16- line 42 in page 2), or move it to the discussion section.

We are repeating here that the main focus of our study is denoising gene-expression as well showing its utility in study regulatory changes due to ageing with single-cell expression profile. Such gene-network based study of single-cell expression profiles has rarely been done in the context of ageing. We also showed the relevance of our study by adding some results from the analysis of expression from SARS-COV-2 infected lungs.
Nevertheless, now we have moved those sentences to corresponding results section to maintain the flow.

2) It is not clear why only the Hard Thresholding result for Graph Wavelet is presented in Fig 3, as in total three thresholding approaches were applied?

We applied hard thresholding as proof of principle, and it also gave consistently good results. The soft-thresholding results were satisfactory but were comparable to hard-thresholding results for bulk expression profile. However, sometimes soft thresholding did not perform well.





We have now put a sentence in thresholding sub-section of Method Section, which is as such

Here, we have used hard-thresholding for most the datasets as proper soft-thresholding of Graph-wavelet coefficient is itself a topic of intensive research and may need further fine-tuning.

3) In the first paragraph of the Discussion section, the authors mentioned that they also compared with the results from imputation and smoothing, but these two approaches have not been documented in the Method section.

Now we have made a comparison against more methods and added a new sub-section in the Method section, which is as such

\subsection*{Comparison with other methods}
We compared the results of our approach of graph-wavelet based denoising with other methods meant for imputation or reducing noise in scRNA-seq profiles. For comparison we used Graph-Fourier based filtering \cite{Burkhardt_2019}, MAGIC \cite{Zhou-2007}, scImpute \cite{LiWV_2018}, DCA \cite{Eraslan_2019} , SAVER \cite{Huang_2018}, Randomly \cite{Aparicio_2020}, KNN-impute \cite{troyanskaya_2001}. Brief descriptions and corresponding parameters used for other methods are written in supplementary Method.

In addition, we have described other imputation and denoising methods briefly in supplementary Method file.

4) In Fig 4, the authors should state clearly what the raw data is and what the filtered data is (analysis method, which thresholding).

We have now added the following text to the caption of figure-4. Here is the text we added.

The label "Raw" here means that both networks (for old and young) were inferred using unfiltered scRNA-seq profiles. Whereas, the same result from denoised scRNA-seq profile is shown as filtered. Networks were inferred using correlation-based co-expression.

5) What is the value of K for the K-nearest neighbors (KNN) method?

We have now written is clearly

Here we decide the value of K in the range of 10-50, based on the number of samples(cells) in the expression datasets.

6) Some grammatical errors need to be corrected, for instance:

We have now made improvements wherever correction was needed.

a. Paragraph 2 in page 1, “few methods” should be “a few methods”

We have fixed this typo

b. First paragraph in page 2, “Thus the major problem of handling noise and dropout in scRNA-seq profile is an open problem.” seems weird, and need to be adapted.

We have changed this sentence as given below

Thus the major challenge of handling noise and dropout in scRNA-seq profile is an open problem.

c. line 48 in page 2, “network inference method” should be “network inference methods”.

It is fixed now

d. line 22 in the right part of page 2, “can also exist” should be “can both exist”.

It is fixed now

e. last line in the left part of page 2, should be “choose the thresholds of …..”

We have now changed it to ..

we choose the threshold for wavelet coefficients using sureShrink and BayesShrink or ….

f. last paragraph of the right part in page 7, “more closer” should be “closer”.

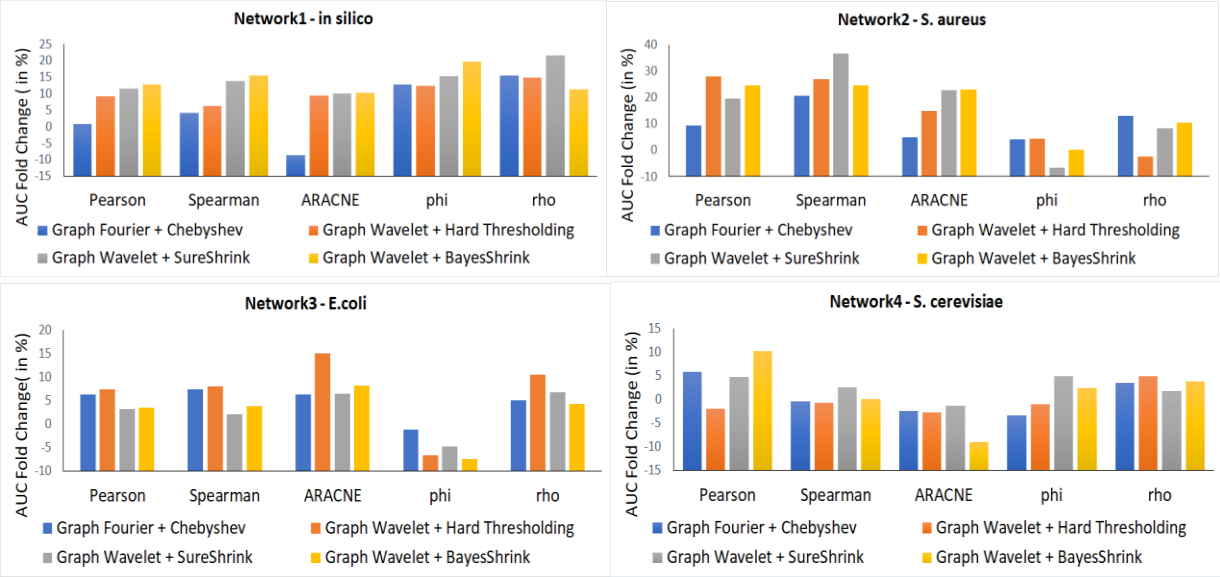
It is fixed now in the new version

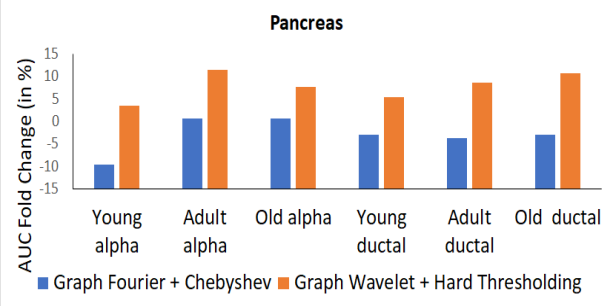
Reviewer: 2

Comments to the Author

1. For example, the term fold change is used in understanding the changes in gene expression measurements. The authors appear to have extended it to network overlap with changes in area under curve. This caught me quite confused and it took some time for me to comprehend the presentation. I am quite sure many readers out there would also undergo similar experience thereby not being able to appreciate the work.

We thank reviewer-2 for highlighting this point. We have now added more information to Y-axis labels in figure. We also provided more detail in the related figure-caption





In the caption of figure 2 we have explained with following sentences

The Y-axis shows fold change in area under curve(AUC) for receiver operating characteristic curve (ROC) for overlap of predicted network with golden-set of interactions.

2. There is practically no detail given so that I can reproduce the work with another dataset. for example the KNN protocol is rather described more in an 'idea' fashion than in a way that one could implement for their datasets.

We had already provided code on weblinks (http://reggen.iiitd.edu.in:1207/GraphWavelet/index.html) and https://github.com/reggenlab/GWNet/
One can use the code to reproduce the work on other data-set.

We have now added the following text to explain KNN based graph construction in more detail

…..Hence we first made a base-graph (networks) where supposedly identical cells are connected by edges. For every gene we use this base-graph and apply graph-wavelet transform to get an estimate of variation of its expression in every sample (cells) with respect to other connected samples at different levels of graph-spectral resolution. For this purpose, we first calculated distances among samples (cells). To get a better estimate of distances among samples (cells) one can perform dimension reduction of the expression matrix using tSNE \cite{maaten2008visualizing} or principal component analysis. We considered every sample (cell) as a node in the graph and connected two nodes with an edge only when one of them was among K-nearest neighbors of the other. Here we decide the value of K in the range of 10-50, based on the number of samples(cells) in the expression datasets. Thus we calculated the preliminary adjacency matrix using K-nearest neighbours (KNN) based on euclidean distance metric between samples of the expression matrix. We used this adjacency matrix to build a base-graph. Thus each vertex in the base-graph corresponds to each sample and edge weights to the euclidean distance between them.

3. When the authors say denoising what exactly they refer to as noise? We know some terms like white noise. Does noise mean the variances in the measuerement of gene expression values among the different cells? If so, then how does the graph-wavelet filter dampen these variances? Due to this lack of basic mathematical tretament of noise I express my inability to fully comprehend the data presented by authors.

Now we have described the noise at two different locations.

In the introduction, we have written the following sentences

The noise in single-cell expression profiles could be due to biological and technical reasons. {\color{red}The biological source of noise could include thermal fluctuations and a few stochastic processes involved in transcription and translation such as allele specific expression and irregular binding of transcription factors to DNA. Whereas technical noise could be due to amplification bias and stochastic detection due to low amount of RNA. Raser and O'Shea \cite{Raser_2005} used the term noise in gene expression as measured level of its variation among cells supposed to be identical. Raser and O'Shea categorised potential sources of variation in gene-expression in four types : (i) the inherent stochasticity of biochemical processes due to small numbers of molecules; (ii) heterogeneity among cells due to cell-cycle progression or a random process such as partitioning of mitochondria (iii) subtle micro-environmental differences within a tissue (iv) genetic mutation}. Overall noise in gene-expression profiles hinders in achieving reliable inference about regulation of gene activity in a cell-type. Thus, }

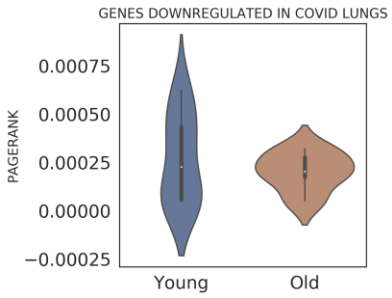
In the beginning of Methods we have written the intuition behind the graph-wavelet based denoising

We used the term noise in gene-expression, according to its definition by Raser and O'Shea \cite{Raser_2005}; that is, measured level of variation in gene-expression among cells supposed to be identical. Hence we first made a graph (networks) where supposedly identical cells are connected by edges. For every gene we use this graph and apply graph-wavelet transform to get an estimate of variation of its expression in every sample (cells) with respect to other connected samples at different levels of graph-spectral resolution. For this purpose, we first calculated distances among samples (cells). To get a better estimate of distances among samples(cells).....

4. In Figure 5 only positively regulated genes with PAGERANK is presented. The system has both positively and negatively regulated genes and the physiological effects is likely a

summation of both. It also looks like the FC values in Fig. 5E are log values but they are simply mentioned as FCs..

We thank reviewer-2 for this point. We have now added figures showing pageRank of genes down-regulated in COVID-infected lungs. The figure below has been added to Figure 5A.



For Figure 5E, now we have mentioned that Y axis show log(Fold change in expression)

5. The authors have done extensive work but I feel that the manuscript has to be presented more focussed in order to appreciate the graph-wavelet filters and their useful properties.

We thank reviewer-2 for appreciating our analysis. The title of the manuscript reveals that it is also about getting some insight from single-cell expression profiles from ageing samples, we have kept this section. However, as the reviewer had requested to focus more on filtering, we have added text about the intuition behind graph-wavelet filters and its benefit, as shown in the text above.

Reviewer: 3

Comments to the Author

This is an interesting powerful method. Some parts are too difficult to bioinformatics and biologists readers, particularly related to the graph wavelets methodology. I think the authors should make more efforts to explain the reasoning of using this approach and discuss more insightful its limitation or potentialities.

We thank reviewer-3 for his appreciation, Now in the beginning of Method section we have added the reasoning. The text is like this

We used the term noise in gene-expression according to its definition by several researchers such as Raser and O'Shea \cite{Raser_2005}; as the measured level of variation in gene-expression among cells supposed to be identical. Hence we first made a base-graph (networks) where supposedly identical cells are connected by edges. For every gene we use this base-graph and apply graph-wavelet transform to get an estimate of variation of its expression in every sample (cells) with respect to other connected samples at different levels of …..