Dear Prof. Faeder and Leitner,

We would like to thank you and the reviewers for the constructive feedbacks, and also for coordinating the reviewing process of our work. In our revised manuscript, we tried our best to address them one by one. We hope the new submission answers all the questions you and the reviewers raised in the last round reviewing. In the following, our response is in red. Our responses have also been incorporated to the new submission.

Sincerely,

The authors

# Response to Reviewer 1

R1.1 In “Use of compressed sensing to expedite high-throughput diagnostic testing for COVID- 19 and beyond”, the authors present a compressed sensing-based pooling strategy for accurate identification of positive disease samples at high prevalence rates. The strategy is tested with an MHV-1 model system and with primary human COVID-19 samples. The authors show their method has high accuracy and reduces the required test numbers in the experiments.  
  
This article addresses an issue of importance to public health. The strategy proposed by the authors could be potentially used in other scenarios that require large-scale screening of disease samples.

We would like to thank the reviewer for the constructive feedbacks. In the following, we will provide a point-to-point response for each of the valuable feedbacks, and our responses have been incorporated in our new submission.

R1.2 However, it is not very clear to me how the authors’ approach compares with the current state-of-the-art pooling methods. In page 5, the author state “More sophisticated pooling efforts have been developed and validated during the pandemic through the accuracy and effectiveness of these new approaches falls apart rapidly as the prevalence rate rises”. I think at least two pooling methods work at high prevalence rates (>10%), Abdalhamid et al. and Eberhardt et al. I wonder whether the authors could benchmark their method with existing pooling methods and give more quantitative measurements about the improvement. That doesn’t necessarily mean the authors’ method should be superior to all previous ones. But it could help the readers to better appreciate the rational design of their protocol. To me, it is not very meaningful to claim a reduction in the total number of required tests compared to individualized testing (for example, page 20 last paragraph), since every pooling strategy is expected to be better.

Abdalhamid B, Bilder CR, McCutchen EL, Hinrichs SH, Koepsell SA, Iwen PC. Assessment of specimen pooling to conserve SARS CoV-2 testing resources. Am J Clin Pathol 2020;153(6):715–718  
  
Eberhardt JN, Breuckmann NP, Eberhardt CS. Multi-stage group testing improves efficiency of large-scale COVID-19 screening. J Clin Virol 2020;(e-pub ahead of print). doi: 10.1016/j. jcv.2020.104382

We would like to thank the reviewer for this constructive feedback, and pointing to us the highly relevant references. Kody is taking care of this.

R1.3 I have listed below several other comments which I hope will be helpful for the authors.  
  
1) I am not sure I understand the design of the MHV-1 experiment. What is the pool size? What does a prevalence rate of 0-14.3% mean? Are the 299 samples separated into different subgroups? I would expect the authors to have a better explanation about the experiment details.

We would like to thank the reviewer for this constructive feedback. Kody is taking care of this.

R1.4 2) I would suggest the authors shorten their discussion of the modified RNA extraction protocol in the main text, as they have pointed out in the discussion part “the modified RNA extraction protocol was not required as we were able to accurately and reproducibly identify infected samples in primary human SARS CoV-2 pooling experiments…”

Kody may have the expertise to take care of this comment.

R1.5 3) In figure 2, the mixing matrixes are presented in two ways. I would advise the authors to use either a numerical matrix or pixel matrix.

We would like to thank the reviewer for this constructive feedback. In the new submission, we use pixel matrix to represent all the participation matrices as follows.

A picture containing chart

Description automatically generated

Participation matrix used for MHV-1 experiments:

Qr code

Description automatically generatedParticipation matrix used for MHV-1 experiments:

A picture containing chart

Description automatically generatedParticipation matrix used for MHV-1 experiments:

Qr code

Description automatically generatedParticipation matrix used for COVID-19 experiments:

4) Figure 3B, what are the p values?

We would like to thank the reviewer for this constructive feedback. Kody is taking care of this.

5) I recommend the authors to include a main figure about the compressed sensing decoding.

We would like to thank the reviewer for this constructive feedback. The high level idea of the proposed adaptive decoding approach based on compressed sensing is illustrated in the following figure.

Diagram

Description automatically generated

Overall framework of our proposed adaptive decoding approach based on compressed sensing. In each stage we design a mixing matrix which is used to get pooling results . The mixing matrices and the pooling results from all the previous stages are used for support set estimation. If the status all the samples can be determined, we perform virus load estimation to get  . If the support set estimation results and the virus load estimation results are consistent, then we report the final results. Otherwise, we report warning about errors during the pooling stage. If the status all the samples cannot be determined, we design a new mixing matrix , and request pooling results . More specifically, the has only one row which pools all the potentially positive samples after the first stage decoding. If the pool in returns positive, then will be a diagonal matrix with each pool containing only one of the potentially positive sample. The new pooling results will be combined with those from previous stages to go through the decoding process again in each stage.

6) For the human patient samples, 38 tests are required to screen 80 patients. I wonder how many tests are required for traditional pooling. How would pool size affect these results?

We would like to thank the reviewer for this constructive feedback. Kody is taking care of this.

7) I am confused by sample 17 in Table S1, why the viral load is higher than the upper bound?

We would like to thank the reviewer for this constructive feedback. This was because of a typo. In the new submission, we rerun the experiments with another set of parameters to get better decoding results. See the following figure. Kody may need to update the experimental results section since I reran the experiments with different parameters. Here are the data

* C:\Users\jy0821\Dropbox\2020\_CS\_Virus\_Testing\00\_CS\_Virus\_Testing-main\Data

Graphical user interface, application, table, Excel

Description automatically generated

8) Page 10, Wij should be 0 when Pij = 0.

# We would like to thank the reviewer for this constructive feedback. We have corrected the type in the new submission. Response to Reviewer 2

R2.1 In Waldstein and Yi et al, the authors propose to use a compressed-sensing-related algorithm to increase the efficiency of screening for SARA-Cov-2 viral infection. They propose that their method is specifically advantageous in small N, higher-prevalence settings, as compared with other group-testing methods that are ineffective at high prevalence. The use of a modified procedure to control the dilution factor in each sample is interesting, but I have number of concerns, outlined below, that I think ultimately limit the value of the paper.

We would like to thank the reviewer for this constructive feedback. We have tried our best to address these constructive feedbacks, and in the following, we provide a point-to-point response for them. We hope our responses answer all the concerns raised in the previous round of reviewing.  
  
R2.2 Major comments: 1. The authors propose a new algorithm for compressed-sensing-like decoding that they claim is advantageous for small N, high prevalence settings. These are useful settings to consider, but it remains entirely unclear if their approach actually works better than alternative methods. The authors need, at a minimum, to compare to a standard compressed sensing approach, and at least several of the many many published group testing methods. More generally, it would be helpful if they could more rigorously quantify sensitivity, specificity, and efficiency, and evaluate these quantities across their simulations and real tests, and make comparisons with alternative approaches.

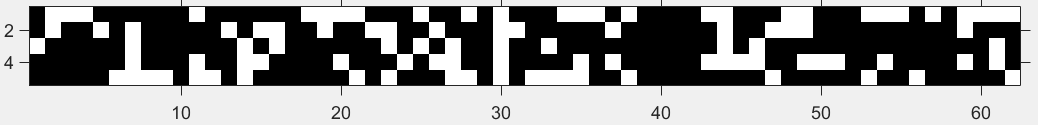
We would like to thank the reviewer for this constructive feedback. We added in the new submission experimental results on numerical experiments for benchmarking the proposed approach with existing ones. Alternative 1 CSL1: standard compressed sensing with minimization. Alternative 2 P3S2: split the population into pools of size 3 with an individual sample appearing in one and only one pool. For a negative pool, all the individual samples are determined as negative. For a positive pool, perform test for each involving individual. Approach 3 Certified (proposed). The individual sample virus load and their status are randomly generated. For both the CSL1 and Certified approaches, we use the same randomly generated mixing matrix for pooling, and we keep the number of pools fixed in the first initial stage for compressed sensing-based approach. In CSL1, we solve the testing problem in one shot without adaptive requests of extra pooling tests, while in Certified, we have potentially two sets of extra adaptive request of pooling results. We define the efficiency as the average number of tests needed to determine the status of an individual sample. We consider different population size and prevalence for the experiments, and use the sensitivity, specificity, and efficiency as the evaluation metrics. The results are presented in the following Table, and we have added it to the revised manuscript.

From the results, we can see that the proposed approach can outperform the CSL1 and P3S2 in various settings. In all the setup, CSL1 is the most efficient due to the single round of test, but its testing performance such as sensitivity and specificity are outperformed by the proposed approach. For example, the CSL1 needs only 0.081 tests to determine the status of an individual sample, but the sensitivity and the specificity are only 0.526 and 0.746, respectively. In all the settings we tested, the P3S2 achieves the highest sensitivity and specificity, but its efficiency is outperformed by the proposed approach. For example, when the population size is 22 and the prevalence is 0.1, the P3S2 needs 0.667 tests for each sample while the proposed approach only needs 0.5 tests per sample. However, when the prevalence is high enough, e.g., reaching 0.3 or higher, neither P3S2 and Certified can provide any efficiency gain, e.g., 1.145 tests/sample and 1.097 tests/sample for P3S2 and Certified, respectively, and individual tests without pooling should be preferred. This is also consistent with [1].

[1] Eberhardt JN, Breuckmann NP, Eberhardt CS. Multi-stage group testing improves efficiency of large-scale COVID-19 screening. J Clin Virol 2020;(e-pub ahead of print). doi: 10.1016/j. jcv.2020.104382

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Decoding Methods | Population Size 62, Prevalence 0.3 | | | Population Size 12, Prevalence 0.2 | | | Population Size 22, Prevalence 0.1 | | |
| Sensitivity | Specificity | Efficiency | Sensitivity | Specificity | Efficiency | Sensitivity | Specificity | Efficiency |
| P3S2 | 1.000 | 1.000 | 1.145 | 1.000 | 1.000 | 0.667 | 1.000 | 1.000 | 0.667 |
| CSL1 (n=5) | 0.526 | 0.746 | 0.081 | 0.500 | 1.000 | 0.417 | 1.000 | 0.900 | 0.227 |
| Certified (n=5) | 1.000 | 1.000 | 1.097 | 1.000 | 1.000 | 0.528 | 1.000 | 1.000 | 0.500 |

Table 1: comparison between proposed approach and existing approach.

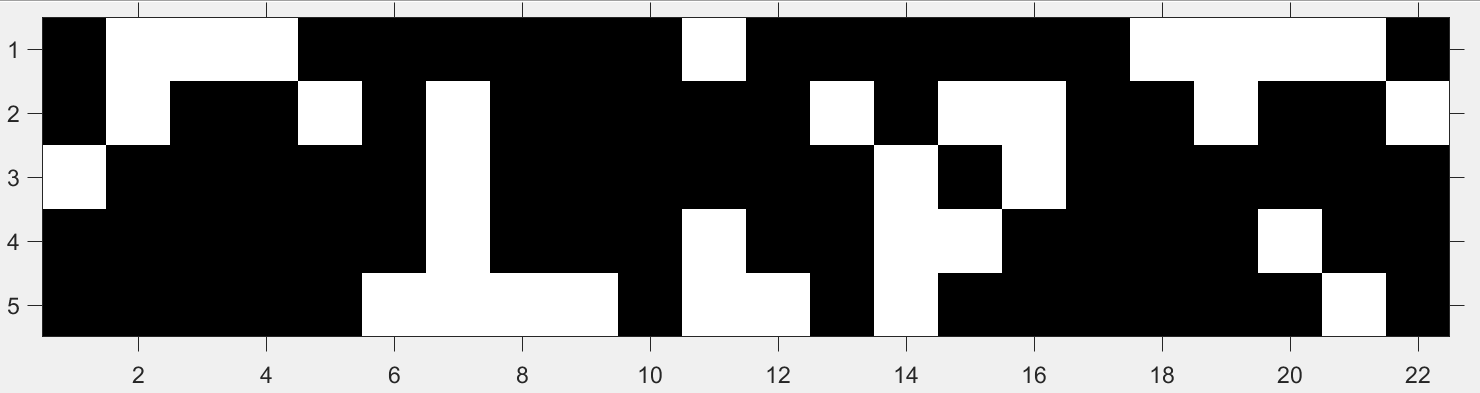


Participation matrix used in CSL1 and Certified with pool number n=5, population size N=62, and prevalence 0.3.

Logo

Description automatically generated with medium confidence

Participation matrix used in CSL1 and Certified with pool number n=5, population size N=12, and prevalence 0.2.



Participation matrix used in CSL1 and Certified with pool number n=5, population size N=22, and prevalence 0.1.

R2.3 2. The authors discuss both simulation and testing on real samples, and potentially also simulation to determine parameters for tests on real samples. I am uncertain. In general, when it came to presentation of results, I was extremely confused as to what was actually done. If there were simulated tests performed, what was the simulation procedure? Where is this described? For example, the authors refer on pg 18 to “1325 computer simulated samples…with prevalence rates ranging from 0-14%.” Why were there 1325 simulated samples? How many batches were simulated, and what were the values for N and n in each batch? What was the actual distribution of positivity rates (not just the overall range) within each batch? What were the viral loads in positive samples, and how were these selected? Much more detail is needed here.

We would like to thank the reviewer for this constructive feedback. Kody is taking care of this.  
  
R2.4 3. I am similarly confused about what was actually done in testing real samples. Apparently 299 MHV-1 samples were tested, but it is unclear why 299 samples, how many were positive, what their viral loads were, how many positive samples per batch, how many batches and what values of N and n per batch. I am similarly confused about what was actually done with testing human samples. More detail is needed, including in the description of results on pg 21, which I could not follow.

Kody may have the expertise to take care of this comment.  
  
R2.5 4. In presenting results the authors refer to Algorithms 1, 2, and 3, but it was never stated what these algorithms correspond to – presumably something related to earlier-described methods, but it was never clearly defined. This made it difficult to understand the results presented on pages 18 and onward.

We would like to thank the reviewer for this constructive feedback. In the new submission, we have corrected typos of the references to algorithms.  
  
R2.6 5. It is unclear why the authors use parity check matrices for MHV-1 samples and bipartite-graph construction for human samples. As the authors note, the parity check matrices are not practically useful for even moderate values of N. However, these matrices are also not particularly useful if there is uncertainty in the positivity rate. If the positivity rate is fixed at 0 or 1 positive samples in a given set, these matrices are fine, but this is not a useful scenario to evaluate since we never know how many positive samples there will be in a given batch.

We would like to thank the reviewer for this constructive feedback, and we totally agree with the reviewer about this aspect, i.e., the parity check matrices can only perform well when there is only single positive sample in a population. The major advantage of them is the high efficiency in populations with one single positive, i.e., we only need pools to accurately identify the positive from a population of size . This implies tests/sample, e.g., 0.0098 tests/sample when .

R2.7 6. The presentation of the bipartite-graph design is not sufficient. The authors should state at least the basic principles of how samples are assigned to pools (i.e., how the subset of edges is selected), and not just describe what a bipartite graph is. Moreover, they should describe any constraints imposed by their method. Can they handle any number of individuals and any number of pools? Usually the answer to this is “no,” unless one allows the designs to be poorly-behaved. Finally, why was the specific design shown in Fig 2D chosen? (i.e., why 16 pools, why 4 pools per sample, and how were the specific pools for each individual sample selected?)

We would like to thank the reviewer for this constructive feedback. In the new submission, we give more details about how the mixing matrices are constructed, and we also give discussions about the advantages and disadvantages of the mixing matrices constructed from Bipartite graphs when compared with other mixing matrices. Let Michael take care of this.  
  
R2.8 7. The authors state that dilution “has required pools to remain small, usually less that 5 patients per pool.” This is not accurate. There are many published examples by now of effective pooling strategies with many more than 5 samples per pool, and it is well-understood that in many infected individuals the viral load is many orders of magnitude above the limit of detection (thus, most samples can be diluted far more than 5-fold and still be identified). This doesn’t mean that limiting dilution is not helpful, but doing the extra work to fix dilution at 4-fold (what the authors do) vs say 16-fold may not have much of an impact and might not be worth it. The additional samples one would expect to identify with the modified procedure, vs those missed due to “normal” dilution, would be those with a viral load falling between 4- and 16-fold of the limit of detection. It is unclear how beneficial it would be to identify the small fraction of positive samples falling in this range, but this benefit must somehow be weighed against the cost of the additional work.

Kody may have the expertise to take care of this comment.  
  
R2.9 8. The authors suggest their approach, which recovers an estimate of the original signal (i.e., viral loads), is advantageous over group testing methods that output the positive or negative status of each sample, but they do not provide any explanation for why it is useful to know the viral load in addition to positive/negative status. Moreover, in many 2-stage pooling methods, putatively positive samples are individually tested in a second stage, and estimates of the viral load can be easily determined from the Ct values of individual tests.

We would like to thank the reviewer for this constructive feedback. Kody is taking care of this.  
  
R2.10 9. The authors state that they reduce computational complexity by finding the sparsest solution with error smaller than a given tolerance. The computational motivation for this is clear, but from a diagnostic perspective this bias towards a sparse solution seems undesirable, and is in contrast to many other group testing methods that are premised on the presumption of sparsity, but which do not incorporate sparse priors in their decoding (in other words the decoding is not biased towards a sparse solution, and for example returns all definite or putative positives without using any sparse prior or sparsity-inducing penalty). If there are other possible (denser) solutions that fit equally-well or even better than the sparsest fit within error tolerance, why should we not consider these?

We would like to thank the reviewer for this constructive feedback. TBD Test all negative pools, before individual use pooled negatives  
  
R2.11 10. The code for the work needs to be made publicly available, not just available upon request. "Available upon request" severely limits the actual adoption of the work in my experience, especially when the description of methods is not particularly clear.

We would like to thank the reviewer for this constructive feedback, and we agree that our codes and raw data should be publicly released to facilitate further research in related topics. In our new submission, we have added links for accessing the our codes and data for reproduction of our experimental results.  
  
  
R2.12 Minor comments: In the introduction the authors motivate the need for pooled testing in high-prevalence scenarios, stating that “prevalence rates continue to be >10% with a worldwide estimate of 30.” The reference cited is a preprint from 2020 (not up-to-date), and, more importantly, appears to provide estimates of cumulative rates of infection (i.e., including past infections), whereas only current positivity rates within a pool are relevant for this study (these will be much, much lower). There is still value in developing methods that are efficient under high positivity rates, but this is not an appropriate way to motivate it.

Kody may have the expertise to take care of this comment.  
  
R2.13 The final sentence of the results section could be read as implying that the modified RNA extraction process was tested on human samples, but this process was only tested on the MHV-1 samples (as far as I can tell). This should be clarified.

Kody may have the expertise to take care of this comment.  
  
  
R2.14 The authors focus on cases where N is “small,” i.e., N= 7,15,31,40. These numbers are chosen because they are convenient for fitting within constraints of the mixing matrix designs, but are not particularly convenient for typical plate and array sizes in laboratory environments (typically with rows/columns as multiples of 8 and 12).  
  
We would like to thank the reviewer for this constructive feedback. Kody is taking care of this.  
  
R2.15 The authors describe a very general framework for representing mixing matrices and results across multiple rounds, but appear to only propose and use one specific set of designs (parity or bipartite-graph method in round 1, testing all undetermined in and single pool, and individual testing). If this is only procedure the authors propose, it might be more straightforward to present the specific thing you do, and skip the general framework that leads to all sorts of questions about how the designs are chosen at each stage.

We would like to thank the reviewer for this constructive feedback. TBD

# Response to General Requirements

**Have the authors made all data and (if applicable) computational code underlying the findings in their manuscript fully available?**  
The [PLOS Data policy](https://journals.plos.org/ploscompbiol/s/materials-and-software-sharing) requires authors to make all data and code underlying the findings described in their manuscript fully available without restriction, with rare exception (please refer to the Data Availability Statement in the manuscript PDF file). The data and code should be provided as part of the manuscript or its supporting information, or deposited to a public repository. For example, in addition to summary statistics, the data points behind means, medians and variance measures should be available. If there are restrictions on publicly sharing data or code —e.g. participant privacy or use of data from a third party—those must be specified.

Reviewer #1: None

Reviewer #2: **No:**The code is neither provided nor deposited in a public repository.

 We would like to thank the reviewer for this constructive feedback. The raw data and the implementations are available at: <https://github.com/JAMES-YI/00_CS_Virus_Testing>.

Figure Files:

While revising your submission, please upload your figure files to the Preflight Analysis and Conversion Engine (PACE) digital diagnostic tool, [https://pacev2.apexcovantage.com](https://pacev2.apexcovantage.com/). PACE helps ensure that figures meet PLOS requirements. To use PACE, you must first register as a user. Then, login and navigate to the UPLOAD tab, where you will find detailed instructions on how to use the tool. If you encounter any issues or have any questions when using PACE, please email us at [figures@plos.org](mailto:figures@plos.org).

Data Requirements:

Please note that, as a condition of publication, PLOS' data policy requires that you make available all data used to draw the conclusions outlined in your manuscript. Data must be deposited in an appropriate repository, included within the body of the manuscript, or uploaded as supporting information. This includes all numerical values that were used to generate graphs, histograms etc.. For an example in PLOS Biology see here: http://www.plosbiology.org/article/info%3Adoi%2F10.1371%2Fjournal.pbio.1001908#s5.

Reproducibility:

To enhance the reproducibility of your results, we recommend that you deposit your laboratory protocols in protocols.io, where a protocol can be assigned its own identifier (DOI) such that it can be cited independently in the future. Additionally, PLOS ONE offers an option to publish peer-reviewed clinical study protocols. Read more information on sharing protocols at <https://plos.org/protocols?utm_medium=editorial-email&utm_source=authorletters&utm_campaign=protocols>