

Response to Report 1

We will try our best to address all the points in this report. Since the review points are not clearly identified, we just select them from the report based on our judgment.

Point 1. *This condition is now weakened in this manuscript. The general idea to find the set of parameters $\{\phi_j, \beta_l\}$ then works as before by recovering the corresponding Prony polynomial.*

Response 1. The condition is weakend can be treated as an improvement. However, our method relies on the existence of a linear operator for which both the generating functions are the eigenfunctions. This condition is very strong, and it is very hard to find a linear operator with such a property. We really like to get this condition weakened, and hope we can improve this condition in the future. As the report points out, we basically apply the existing methods in the literature to solve this new problem, but the solution is non-trivial. To weaken the condition, we feel that some new ideas or new techniques should be developed to provide a better solution for this two-generator sparse expansion problem; otherwise, the technique is too restrictive to be applied in solving more general real-world problems. We think that the problem is a good problem, and this is our first try in solving this problem. We wish that our result can inspire more researchers in this field to contribute new ideas and methods to conquer this problem.

Point 2. *The new difficulty is now to extract from the set of zeros $\{\lambda_k, k = 1, \dots, M_1 + M_2\}$ of the Prony polynomial the frequencies, which belong to the u -expansion, and which belong to the v -expansion. The idea to consider a combinatorial problem and to solve the linear system for all possible choices is not feasible.*

Response 2. We feel that this point is the biggest one in this report. So we are willing to spend more time to address this point. The main issue here is that we need to go through a combinatorial number of cases, say $\binom{2n}{n}$ cases, and the referee feels that this approach is not feasible. We would like to view this point in a global way. First, we have not noticed that in solving problems using the Prony method, the researchers in this field have discussed the “*feasibility*” issue. The classical Prony method requires a step that calculates a matrix inverse of an $n \times n$ matrix (Hankel matrix), which is of complexity

$O(n^3)$. It seems that nobody in this field worries about the feasibility of the matrix inverse operation. We agree that there is a big difference between the two numbers: $\binom{2n}{n}$ and n^3 when n is large. Can we simply say that a method of complexity $O(n^3)$ is feasible, while a method of complexity $O(\binom{2n}{n})$ is not? The answer should rely on the individual applications. For some applications involving large data sizes, an algorithm with $O(n^3)$ could be treated as non-feasible. One of the authors of this manuscript happens to work in a field called EMD (*Empirical Mode Decomposition*), where there is a similar situation to the one we are talking about here regarding the feasibility of an algorithm. But that is the difference between $O(n^3)$ and $O(n)$. The EMD implementation requires a step using cubic spline interpolation, and the current method uses a *global* interpolation approach which needs a matrix inverse operation of complexity $O(n^3)$. Due to this reason, many applications require hours for processing, which is not suitable for real-time processing. Since there is no $O(n)$ solution available, people have to take this $O(n^3)$ solution. Can we say that the $O(n^3)$ solution requires hours for processing, thus it has no value in the field? That's why we say the answer relies on individual applications, and it is a decision of individual users who decide if the solution can be used. To evolve from a $O(n^3)$ solution to a $O(n)$ solution, it takes more than 20 years research in the field. The author mentioned above recently developed a $O(n)$ solution for the cubic spline interpolation using a *local* approach, and this work will appear in an ACHA (*Applied Computational Harmonic Analysis*) special issue soon. It took the author more than 5 years research to solve this problem. After a $O(n)$ solution is available, it is highly likely that the researchers would treat a $O(n^3)$ solution non-feasible. As researchers, we know how hard to make significant improvement on a difficult research problem. Here we have a situation that we need to improve a solution from $O(\binom{2n}{n})$ to $O(n^3)$. Before a solution in $O(n^3)$ is available, can we say that a solution in $O(\binom{2n}{n})$ has some value for certain users? If a user needs to solve a two-generator sparse expansion problem for up to 20 parameters, i.e. $n = 20$, since $\binom{20}{10} = 184,756$, our method can handle this type of computation. For these users, our solution is probably the only solution they can use before some super-powerful solution is available.

Definitely we are not satisfied with this solution, and we really want to bring down its efficiency from $O(\binom{2n}{n})$ to $O(n^3)$. But it may take long time with hardworking research to get to that point. This could be a very challenging research problem, and we would like to expose it to the researchers in this field for a solution. In this manuscript, we give a first try on this two-generator sparse expansion problem, and we feel that this problem

is meaningful in the real-world applications. Our solution is just a starting point, and we hope to attract other researchers to provide better solutions. Or if we are lucky enough, we may have a chance to solve this problem in the future.

Point 3. *In particular for the example in Section 3 (equation (37)), wouldn't it make much more sense to consider a simply a larger system of the form*

$$f(x_0 + hn) = \sum_{j=1}^{M_1+M_2} \{c_j \cos(\lambda_j(x_0 + hn)) + d_j \sin(\lambda_j(x_0 + hn))\}$$

for $n = -2(M_1 + M_2) + 1, \dots, 2(M_1 + M_2) - 1$? This system should have at least one solution, since the original model is incorporated, and if there is a unique solution, then you should get some vanishing coefficients c_j and d_j .

Response 3. This suggestion provides a smart way to get around the frequency separation problem. Actually, it makes the two expansion generators share the same set of parameters, and the frequency separation issue is handled by certain vanishing coefficients c_j and d_j in a hidden way. If every step in the (modified) Prony method can work through, the solution is very simple, and there is no need to consider the multi-generator sparse expansion problem. However, any vanishing coefficients c_j and d_j would cause some fundamental problem in the (modified) Prony method, because the critical Hankel matrix would be singular. The Hankel matrix has a matrix factorization that is the product of two Vandermonde matrices and a diagonal matrix with the given coefficients at the diagonal entries. If any one of the coefficients becomes zero, then the whole matrix becomes singular. All the existing modified Prony methods inherit this property, and in our method, we have to assume that all the coefficients c_j and d_j are nonzero. That means the two-generator sparse expansion problem cannot be solved in this special way.

Point 4. *On page 4, line 2, it is claimed that these results are from [8], but this is wrong. The given results are from [9].*

Response 4. We feel sorry that we made this kind of mistakes. We fixed it, and thank the referee to help us reduce the number of errors in the manuscript.

Point 5. *The paper needs a major revision, where in each section it is exactly made clear, which problem is solved, which assumptions are needed etc. It would be best to clearly formulate a theorem in section 3 (and similarly in section 4), which is then proved*

under the given assumptions.

Response 5. We significantly changed the structure of our section 3 and section 4 with each one containing one theorem for the main result in that section. In this way, the readers can view our main results much better. We appreciate the theorem suggestion of the referee, because in this way, we can make the assumptions/conditions clearer for the readers.

Point 6. *For example, I miss clear assumptions for the range of ϕ_j and β_l as well as on the choice of h . Are the coefficients c_j and d_l complex or real? Anyway they should be nonzero! How to choose h and x_0 ? It is important to relate h to the range of ϕ_j and β_l ! What happens if just $x_0 = 0$? Do we need then less function values?*

Response 6. This point is very important to make our results easy to follow. We addressed the questions above in 3 conditions for each of the two theorems. Specifically, the range of ϕ_j and β_l and the choice of h are specified in Condition 2° of **Theorem 3.1**. The requirement for c_j and d_l is specified in Condition 1° of **Theorem 3.1**. The restriction on x_0 is specified in 3° of **Theorem 3.1**. In order to avoid listing too many cases, we just say that $\cos(\phi_j x_0)$'s and $\sin(\beta_l x_0)$'s are nonzero, which eliminates the case that $x_0 = 0$. Since $x_0 = 0$ is not allowed, we thus do not discuss the number of function values needed.

Point 7. *Is the a priori knowledge needed about the values M_1 and M_2 or maybe only $M = M_1 + M_2$? Or can the numerical method be generalized to find also M_1 and M_2 , if one knows an upper bound L of M and has sufficiently many function values?*

Response 7. These questions are raised in our section 6, the conclusion section. In this paper, we cannot answer these questions, which require extensive research to provide deep results. The combinatorial number of cases problem exists in these questions. To answer these questions, we may need to split it into two steps: 1) Do not worry about the combinatorial number of cases. Only focus on the numerical method to determine M_1 and M_2 . 2) Find some special property, so that we do not have to go through all the possible cases, only a relatively small number of cases. But the discovery of that special critical property would not be easy.

Point 8. *In particular it is important to find a feasible method to get ϕ_j and β_l out of*

$\{\lambda_k\}$ without combinatorial try out.

Response 8. This point is seriously addressed in our **Response 2**, and we said that it is the biggest point in this report. The success of a solution that overcomes this problem relies on the discovery of some critical property that can help us eliminate a large number of cases. In that way, we only need to try out a relatively small number of cases. Before that key property is available, the current solution is only treated as a temporary solution, and we wish to inspire more experienced researchers to conquer this important research problem. The method developed solving this problem could be used to solve many other related problems.

Point 9. *Similar improvements are needed for section 4.*

Response 9. We basically do the same thing in section 4 as we did in section 3. Since there is no need to separate the parameters, the method is much simpler. But we feel that the section 3 case is more valuable for the research field, because it allows us to go deeper in research.

Point 10. *Summarizing, the paper needs a major revision, before it can be considered for publication.*

Response 10. We work the best we can to revise the paper. Specifically, section 3 and section 4 are formulated with two major theorems containing clear conditions. In section 5, we replace the first example by a more complex example that is suggested by the second referee. More importantly, to solve the problem, we developed a *two-stage least-square detection* method which contains some interesting property. Although there is a long way to go for a satisfying solution, we feel that our solution is not a bad starting point towards solving this two-generator sparse expansion problem completely.