The author would like to thank the associate editor as well as the reviewer for their valuable feedback and their interest in the paper. The following is a summary of the answers to the questions raised by the reviewer (the four questions raised by the reviewer are in **bold red**):

1. **The authors term their approach to be ultrafast, but no results are presented to show, how the approach will lead to ultrafast relaying, since wavelet and ANN are still mostly offline.**

The question has two parts to it, the first one is about the speed of the proposed algorithm and the second one on whether it is useful for relaying since ANN training is done offline. I will deal with both in succession.

The term ultrafast is used to mean that the time needed to reach a decision about whether there is a fault or not in the system is much less than the time needed by conventional phasor estimation methods and this consistent with usage of the word in literature. The approach given in this paper uses only 1/8th of a cycle (that is 0.002 seconds on a 60 Hz system) to judge whether the fault has occurred on the protected line or another type of transient has happened on the protected line or its adjacent lines. Compared to other methods such as the phasor based methods -which need at least a cycle for secure phasor estimation- or wavelets based methods -which need at least one quarter of a cycle- or even single ended traveling wave methods, the method proposed in the paper is faster. For example in [1], the authors uses the term ultrafast to mean that they need at least 1.5ms to operate. This is very comparable to the speed reported in this paper. It should be pointed out that in this analysis the author has not counted the time used by the microprocessor for pre-processing and post-processing, i.e., the word "time" in this context, is used to mean the time window needed for reaching a secure decision without taking into consideration any hardware or software delays which is consistent with the usage in [1] and [2]. It is interesting to see that the time used by the proposed algorithm is very comparable to the time taken by the latest SEL traveling wave relay [2] which takes at least 1 ms and can take up to 4 ms to reach a decision even with high speed fiber communications transferring massive amount of data between the two ends. The approach in this paper is single ended and is guaranteed to give a decision using 2 ms only of the post-event signal (post-event means from the moment of the event initiation)

Secondly, the reviewer has concerns on using offline simulations for relaying when the claim is doing relaying which is online. The author would like to turn the reviewer's attention that the training has been done using hundreds of thousands of cases. For example, 8066 fault cases per line and a total of 54336 transient cases have been used to train ANN as given in section V-B. These cases encompass all variations of any transient that can occur on a transmission line. EMTP is mature enough that it can capture the actual system response. A real event oscillography would look like the simulated case in EMTP if the EMTP model is an accurate representation of the system, which the author thinks is the case as stated in section V-A. It is not possible to exhaust all foreseen transient parameters, for example it is not computationally feasible to include all values of fault resistance into consideration when creating all transient cases for training. However, it is not hard that one uses his engineering judgment to see that the oscillations present in the transient signal corresponding to a fault case at 50% of line length shall be in between the oscillations of the same fault case evaluated at 5% and 95% of line length. This notion of “in-betweenness” is illustrated in Fig. 1 and is further explored in the answer to question 3 below. Similarly, a fault case with a fault resistance of 20 Ω should have oscillations that fall in between

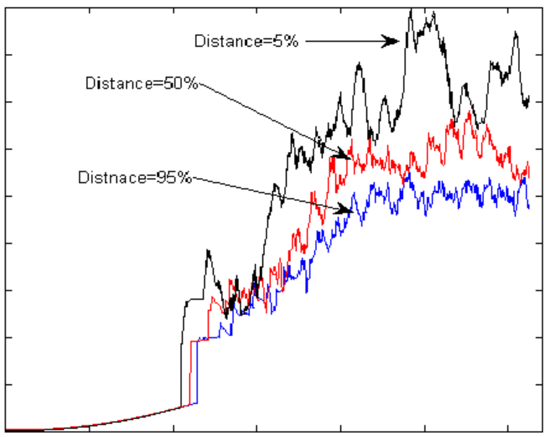

Fig. 1. Figure shows the aerial mode current of a certain fault case at different fault location. The same fault type, fault resistance and incipient angle is evaluated but at different fault location.

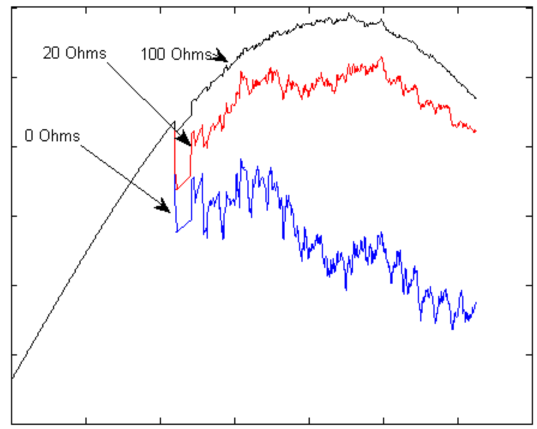

Fig. 2. Figure shows the aerial mode current of a certain fault case with different fault resistance. The same fault type, fault location and incipient angle is evaluated but at different fault resistance.

the same fault case but with a fault resistance of 0 Ω and 100 Ω. This is illustrated in Fig. 2 and is also further explored in the answer to question 3 below. This idea of “in-betweenness” makes this approach applicable to online applications as long as the extreme cases are included in the offline training. As explained in the paper, the transient signal is eventually translated into a vector that belongs to the n-dimensional space. This n-dimensional vector cannot be visualized as n is much larger than 3 and ANN does region separation in this higher dimensional space. Thus, if the fault cases corresponding to 5%, 20%, 80% and 95% of line length can be linearly separated from all other transients, then the fault cases corresponding to 50% of line length should fall in the region of those fault cases. There is no rigorous mathematical proof for this but engineering judgment tells us that this should be the case. This idea shall be very clear by referring to Fig. 3, 4, 5 and 6 which illustrate the effect of varying one parameter of the transient case on the DWT coefficients. For Fig. 6, please note that the oscillations of the three curves do in fact lie “in between” each other if the black curve is multiplied by -1, i.e., the magnitude of oscillations do change around a certain magnitude but sometimes they have different signs. All figures are also attached separately as vectorized pdf files so that the reviewer can zoom in infinitely as needed to see the difference between the curves. Also, please note that Fig. 3, 4, 5 and 6 are for one fourth (1/4) of a cycle of post information data and the author is using only one eighths (1/8) of cycle.

I think a more descriptive title of the paper would be “Ultrafast Transmission Line Fault Detection using a DWT based ANN”

2. **The authors need to compare the performance of the proposed approach with the accuracy of the existing approaches**

This is done in section IX. Additionally, the author has inserted a section on the background material for DWT and ANN. The author intends to write a complete paper on the literature survey regarding this point and submit to the next IAS meeting.

3. **The oscillatory nature and signal analysis is not presented**

As the author has explained in response to question 1 above, the essence of the method is converting the transient signal to an n-dimensional vector. When a transient occurs, the oscillations imposed on the fundamental span a wide range of the frequency spectrum. Thus, it becomes unpractical to identify all oscillatory modes within this wide spectrum and take a decision based on each oscillatory mode. What is done in the paper instead is defining an n-dimensional vector that captures all oscillatory modes present in the signal. Since these oscillatory modes cannot be captured using an FFT since they change over time, DWT is used for building this n-dimensional vector. The oscillatory modes are captured in this vector. The length of this vector is considerably less than the range of spectrum under study. For example, if the n-dimensional vector is built using level 5 coefficients, then the dimension of the vector is 30, i.e., n equals to 30. For example, level 5 spans the frequency range from 6.25 kHz to 3.125 kHz - per the sampling rate used in the paper as pointed in section IV-which is much larger than the dimension of the vector space and it would be unpractical to study each frequency in this large frequency band. If one uses level 1 instead, this level spans the frequency band from 100 kHz to 50 kHz and the dimension of the vector space is 480- much less than level 1 frequency band. As opposed to FFT, DWT gives out coefficients that do not represent the actual cosine or sine coefficients of the oscillatory modes but rather those coefficients are to be combined with the mother wavelet chosen to reconstruct the transient signal. For this reason, the author has refrained from providing those coefficients in the paper as they do not measure an intuitive physical quantity as one would expect with the FFT coefficients. However, since the reviewer has asked for the oscillatory nature, the author has inserted Fig. 3, 4, 5 and 6 in this document to illustrate this oscillatory nature and provided one plot only of the figures in the body of the paper for the sake of the length requirement of the paper. The author does not object to including all the plots in this document to the paper if the reviewer and the associate editor sees this feasible from the publication's a point of view.

4. **The paper requires support of more results and analysis related to the wavelet and signal processing to justify the claims**

This questions supplements question 3 above. The reviewer sees that more analysis is needed as well as there are some claims that are not supported in the paper. The author will address both points.

a. First Part:

For the first part, the author will explain why the n-dimensional vector built the way provided in the paper is capturing all information needed for classification.

When a fault (or any transient event) occurs, it launches a traveling wave as well as high frequency oscillatory components. Traveling waves can be easily quantified as they arrive at the line terminals but the high frequency oscillatory components cannot. The reason for the oscillatory components not being easily quantified before in literature will be explained below.

At this point the author needs to introduce new terminology. Speaking about a certain transmission line, the author calls a fault on that line a *fault case*. A *fault case* has its parameters. Those parameters are: incipient angle, fault resistance, fault type and fault location.

Thus a certain fault case on a specific transmission line causes voltages and currents to oscillate in a manner in accordance with the parameters of that fault case. The formal solution of the currents and voltages of a single phase line is given by the following two equations (called telegrapher equations) [3]:

$$\frac{dV}{dx} = R\times I + L\times\frac{dI}{dt} \qquad (1)$$

$$\frac{dI}{dx} = R\times V + C\times\frac{dV}{dt} \qquad (2)$$

Where:

- I, V are the current and the voltage at the line terminals
- R is the resistance per unit length of the line
- L and C are inductance and capacitance of the line per unit length
- x is the distance from a zero reference frame generally taken to be at either end of the line.

A general closed form solution of equations (1) and (2) is hard and generally impossible as it depends on the boundary and initial conditions of the case involved. Perhaps the most straightforward solution method is to obtain the solution in terms of an infinite time series expansion. Using infinite time series expansion would mystify the solution, because the author's intention is to analyze the frequency content of the signal as this content changes over time. Application of Laplace and Fourier transforms to those equations gives out integrals that yet to be solved formally. Moreover, when a three phase line is studied, the above equations become six equations (two equations for each phase: a voltage and current equation). And those equations are interdependent which makes decoupling them hard undertaking, if not impossible in most cases. If we have mutual coupling between two lines sharing the same tower, then we have 3 more equations making them a total of nine equations. And this is only for one line. Those equations then need to be solved simultaneously along with all other equations in the system. For today power systems which consist of thousands of lines, solving all equations analytically in a closed form is behind human ingenuity.

The main aim is to detect the fault once it is initiated. It is very apparent that high frequency fault generated transients are different for each fault case on a specific line. This is because each fault case corresponding to a specific line has its set of equations as given in (1) and (2) above with its unique boundary (lines connected to it) and initial conditions. Moreover, the same fault case on different lines will cause different fault generated transients because each case has its own boundary and initial conditions, which is different for all neighboring lines. For this reason, no attempt is made to solve the equations formally but solve them numerically by running EMTP simulations, extracting the high frequency content then analyzing them.

Since the author argues that those frequencies -which change over time- are different for each line, the author only needs to see the aggregate of all those frequency components for each line, and if intuition is true then the aggregate of those fault generated high frequencies will be different for each line. What remains is to investigate how the feature space can be defined such that this difference can be attained which is done in section VI. If it is possible to map those frequency components to an n-dimensional space, then faults corresponding to a certain line should be separable from faults on adjacent lines. The author takes this idea further and investigates whether the aggregate of all faults from each line is actually ***"linearly"*** separable [4] from each other. This is the main argument behind the use of pattern recognition approach. Since the author is trying to define a feature space that may be ***"linearly"*** separable for each line, then ANN is the method of choice.

For any fault occurring on a transmission line, the transient generated frequency components depend on the initial and the boundary conditions (the lines connected the faulted line). Since by definition all adjacent lines are different then the frequency generated transients for faults on transmission lines will be unique for each line. The question is whether those frequency components can be combined in a certain way to form a pattern to be recognized.

The same argument not only applies to faults but also apply to switching events and lightning strikes on transmission lines. The author calls faults, lighting strikes, and switching events the *transient event types*. Evidently, the frequency content of faults –on all lines- would be different from switching which in turn different from lighting because the equations governing the solutions are always different because of the different transient event type and the different boundary and initial conditions. Those events can then be first recognized and then the event can be traced back to the line originating it as given in section VIII-D. This basically means that there is a mapping that can divide a certain subspace to linearly separable regions that correspond to those event types. If we zoom in that event subspace we can see which line is causing that event type.

b. Second part:

The only claim the author is aware that he stated in the paper and has not shown its validity, is the trials of with different tower configurations. However, it is beyond the length of the paper to discuss all tower configurations. As a matter of fact, the use of the two aerial modes only is the reason for the success of this method for all tower configurations. The ground mode is known to cause inaccuracies in EMTP [5] and for this reason it has removed from the training.

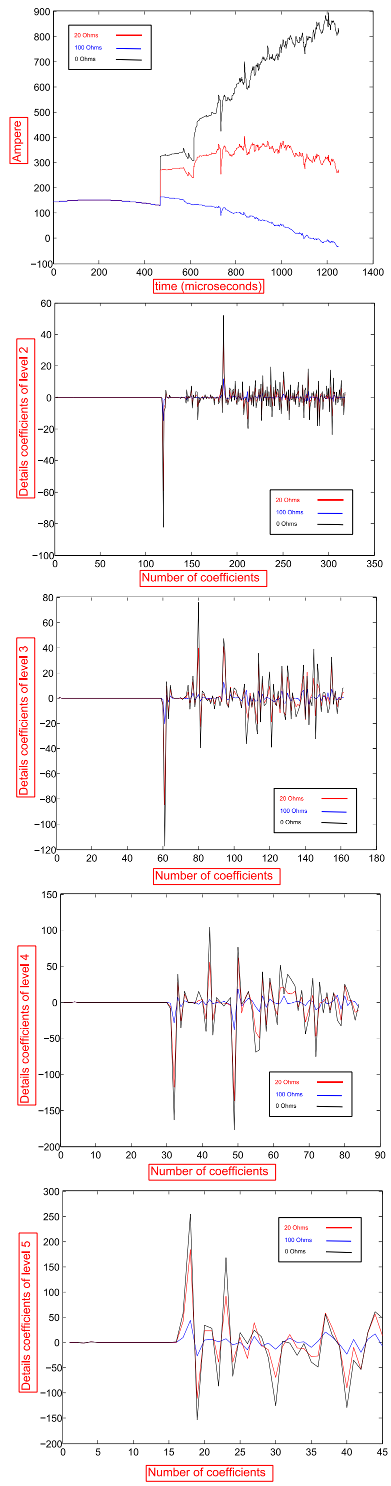

Fig. 3. Figure shows the aerial mode current of a certain fault case at different fault resistances. The same fault type, fault location and incipient angle is evaluated but at different fault resistances.

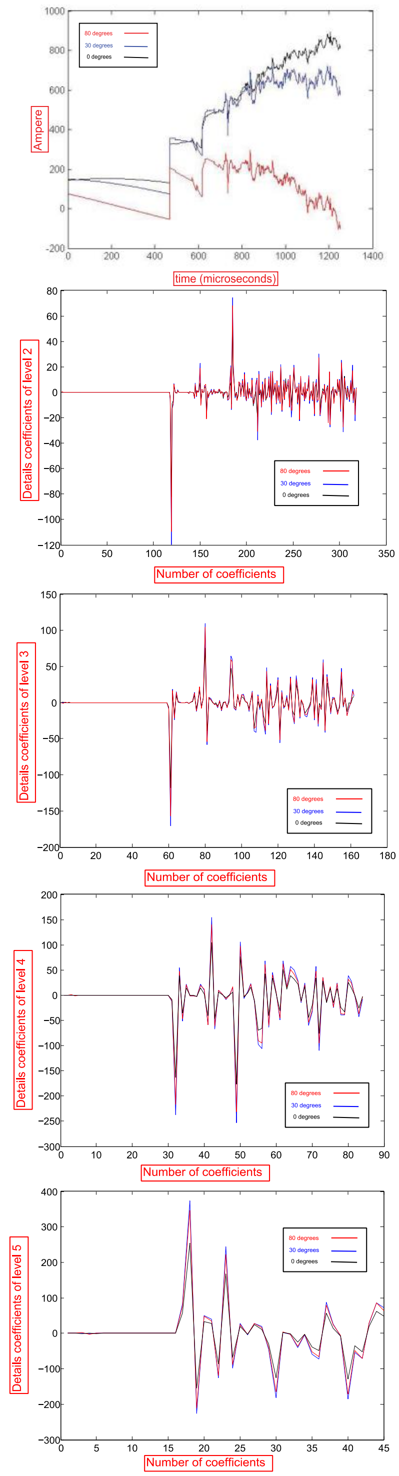

Fig. 4. Figure shows the aerial mode current of a certain fault case at different fault incipient angles. The same fault type, fault location and resistance is evaluated but at different fault incipient angles.

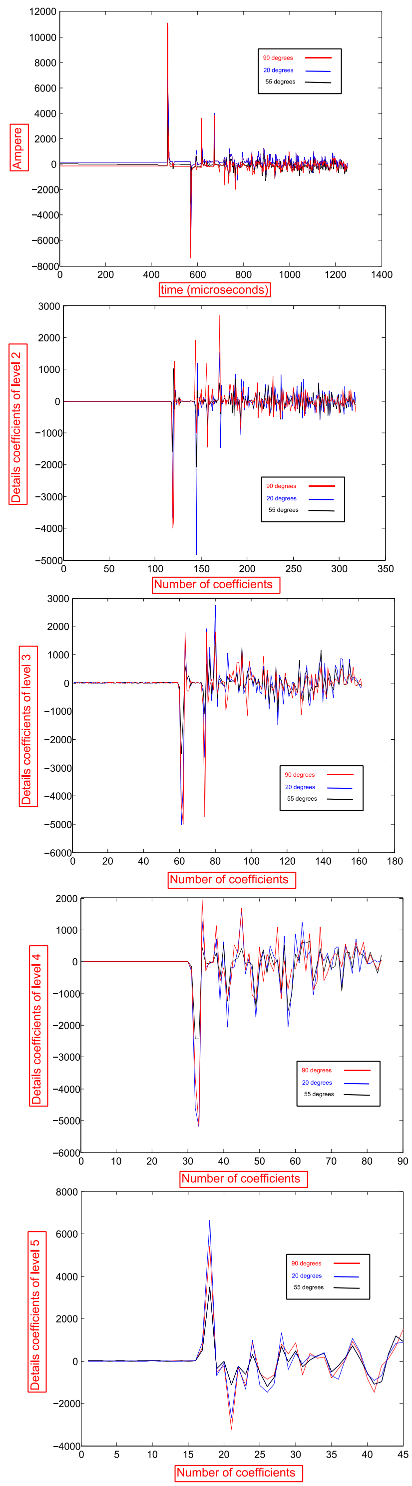

Fig. 5. Figure shows the aerial mode current of a certain lightning case at different incipient angles. The same lightning amplitude, location and phase strike is evaluated but at different incipient angles.

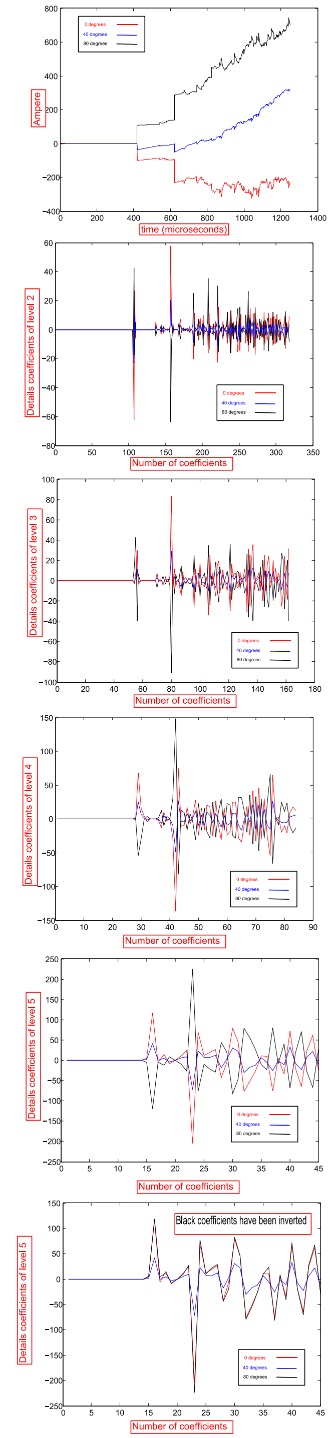

Fig. 6. Figure shows the aerial mode current of a certain line switching case at different incipient angles. The same line terminal is energized but at different incipient angles. Note that the bottom graph is the same as the one on top of it but with the black curve inverted to show the “in-betweenness” of the three curves.

Bibliography

[1] J. Duan, B. Zhang, S. Luo, and Y. Zhou, “Transient-based Ultra-high-speed Directional Protection Using Wavelet Transforms for EHV Transmission Lines,” in *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*, 2005, pp. 1–6.

[2] E. O. Schweitzer, B. Kasztenny, and M. V Mynam, “Performance of Time-Domain Line Protection Elements on Real-World Faults,” in *42nd Annual Western Protective Relay Conference, Spokane, Washington USA*, 2015.

[3] J. J. Grainger and W. D. Stevenson, *Power system analysis*. McGraw-Hill, 1994.

[4] J. M. Zurada, *Introduction to artificial neural systems*, vol. 8. West St. Paul, 1992.

[5] A. Ametani, N. Nagaoka, Y. Baba, and T. Ohno, *Power system transients: Theory and Applications*. CRC Press, 2013.