Reviewer #1 (Required Remarks for the Author):   
  
Authors are proposing models for synchrophasor with step discontinuities in magnitude and phase. The work provides a possible solution for calibration of PMUs under such transients by including data during the transient.   
While the proposed parameters were estimated and results presented both from simulations and measured results, the proposed paper has the following major issues: 

We would like to thank Reviewer #1 for taking the time to read carefully the manuscript and to provide valuable comments on the content of our work.

I. INTRODUCTION   
UTC is Universal Time Coordinated and a time base could be synchronous with the UTC, not centered in.

Agreed. We changed the text as suggested.

II. MATHEMATICAL BACKGROUND   
Equations (1) and (2) both define y(t), but they describe different models. Please use different designation for each model.

Thanks for pointing this out. We added a subscript modifier to distinguish the models.

B. Reference phasor values   
Here Ve is defined but its estimation performance is not reported elswhere in the paper.

Following the changes in Eqs. (1) and (2), we have replaced with and in (4) and (5) to discern between the models for the intermediate phasors in the cases of magnitude and phase steps, respectively. Moreover, we clarified in the text that (4) and (5) are our propositions for reference phasors, which can be computed either via prior knowledge of the parameters of (1) and (2) or their estimation via suitable methods. We proposed the method based on the Hilbert’s instantaneous frequency, described in section II.C, for attaining estimates of . For estimates of the remaining parameters, we followed the footsteps of Frigo [13] and employed the LM.

Table II shows the performance figures of the LM phasor estimator for the step in magnitude case. Table III shows the performance figures of the LM phasor estimator for the step in phase case.

The only results are those presented in the Fig. 7 and 8, but not in the Fig. 11 and 12..

Figures 11 and 12 are related to additional tests in the laboratorial setup. In theory, in the test case of a step in magnitude (resp. phase), the phase (resp. magnitude) shows no discontinuity. However, in an actual laboratorial experimental setup, perturbations in phase (resp. magnitude) are produced when the AWG is forced to generate a step-like waveform in magnitude (resp. phase). We intent to quantify how well our AWG does its job in a challenging scenario of approximating a discontinuous signal in magnitude or phase. For example, during a phase step, the magnitude (which we expect to be constant) is not stable, varies as shown in Fig. 11. Likewise, during a magnitude step, the phase (expected to be constant) varies as shown in Fig. 12. Please note that as we inserted two more figures, the figures numbering changed.

It is not clear what is the purpose or the benefit of using these intermediate values. How do they add to the calibration of the PMU?

Reference phasors are undefined when the observed phasor waveform has steps in magnitude or phase. Our purpose is to propose well defined reference phasors in such situations, which can be used in the calculation of TVE and other related indicators. In previous works, reference phasors are simply obtained from adjacent windows with steady state signals, taken before or after the step occurrence, depending on its time location. This procedure depends on detecting with good accuracy the instant of occurrence , and is based on the assumption that the measurements taken from adjacent windows are suitable estimates of the window with a step. It should be noted that methods that use adjacent windows in the phasor estimate can be easily adapted to employ the models (4) and (5) of the intermediate phasors.

The standard prescribes step tests to measure the time response (TR), which is the time during which the TVE, FE or RFE exceeds the defined limit values. The calibration of P-class PMUs shall provide a very short TR (of the order of a few cycles).

We now report in the paper results of additional computational simulations to show how the proposed intermediate phasor values affect the TVE estimates of a DFT-based PMU. For that, we calculated TVE estimates with two different methods:

1. The reference phasors are taken from the nearest adjacent windows of 500 samples, as related in reference [4], with a Steady State estimator (SS). The value of is known a priori.
2. The reference phasors are taken feeding our intermediate reference phasor models with values obtained by an LM estimator, using 1000 samples. The value of is known a priori.

The simulated PMU receives a signal containing a positive 10% magnitude step, SNR = 50dB. The PMU window takes 1000 samples, with reporting rate of 20 frames/s.



The upper graph shows the TVE response of the DFT-PMU calculated with method 1 (black line), and method 2 (red line). The lower plot shows the TVE calculated with method 2 alone.

If one takes the TR from the TVE calculated with the phasor reference of method 1, it lasts about 250 ms. On the other side, with TVE calculated with the reference of method 2, the DFT-PMU never passes the standard limit. In the first case, this PMU would be rejected as a P-class, but in the second case not. It is the same PMU, the only difference is the definition of reference values for TVE. As TR is defined in function of TVE, this difference can be significative for the already standardized P-class PMU tests, as well as possible more stringent TR tests for distribution PMUs in the future.

D. Model Parameters Estimation via Levenberg-Marquardt   
Please define y(k) beforehand.

Thanks for pointing this out. We defined beforehand as requested.

III. NUMERICAL SIMULATIONS   
B. Parameter estimation with non-linear least squares   
Here, the simulations were caried out using SNR of 90 dB, 93 dB and 97 dB. All these SNRs are too good for any realistic measurement setup concerning PMUs. Consequently, achieved standard deviations are very small, useless for any assessment of their influence in th efinal uncertainty budget. In previous paragraph (A), a SNR of 40 dB and 75 dB were used, which are more realistic. Consider using SNR as expected or measured with the sampling system and sampling parameters used (AWG + DSVM). 

Yes, indeed.

These levels are desirable levels for calibration systems, and considered in the setup used in the reference (Frigo et al.). Although, in a simulated scenario, these high SNR values are useful to compare the performance of our method against alternative solutions, they are indeed unrealistic for the setup used to make measurements in this work.

The measured SNR with AWG + DSVM are estimated to be around 60 dB, according to the references (xxx). Considering SNR = 60 dB, we carried out simulations to assess a more realistic contribution of the parameter estimation in the final uncertainty.

IV. LABORATORY MEASUREMENTS   
... Both are triggered with a 1 PPS (pulse per second) signal. Both goes for AWG and DSVM here, but only trigger to AWG is shown on Fig. 5.

Of course, the figure is corrected now.

Also, if DSVM is triggered externally, there is a time delay and timing uncertainty (likely around 400 ns +- 100 ns) involved. How was this covered in the measured results? 

The DSVM trigger time delay is one component of the overall phase error (it is indeed one of the components of the absolute phase uncertainty, whose estimation is beyond the scope of this work). Therefore, phase uncertainty is part of the variations observed in the results. However, previous works (add ref) indicate time delay less than 50 ns and jitter less than 100 ps for the particular DSVM used.

V. CONCLUSION   
The authors state: "The proposed approach tackles the estimation of the step discontinuities in the phasor signal observed within an analysis window, instead of dodging the problem." However, they do not give any further benefit of doing so, nor they provide any comparison to numerous previous results already published. This is the major issue of this paper - it should provide a clear comparison to the work already published and clearly depict benefits of the proposed method in comparison with previous work.

The proposed method is capable of detecting, without initial guessing, the instant of the transient with time resolution of the sampling intervals and uncertainty of a few units of . If we simply detect the transient and ignore the window, the best time resolution achieved is of a few cycles.

Frigo et al. provide a very detailed comparison of the accuracy of an LM estimator (incorporating a model for the transient) with other representative estimators (IP-DFT and CS-TFM). They report two orders of magnitude less TVEs for the LM estimator for the tests prescribed by the IEEE Standard (in the order of 1e-3), while the others are not even compliant to the standard limit (>1% TVE) when the step rise time is fast (less than 0.2ms). For that, they use only 3 parameters (magnitude, frequency and phase, the others are supposed known *a priori*) for the LM estimator. Based on that, we chose LM as a suitable way to estimate models (1) and (2).

We propose, as an enhancement, incorporate the depth of the step in the LM estimator, to evaluate its performance with four parameters, along with the estimation based on the instantaneous frequency of the analytical signal (via Hilbert’s method). This can be considered a hybrid estimator. Now we include a comparison in section III.C, Table IV, where the performance of the hybrid estimator (HLM4) is compared to the previously published by Frigo et al. [13], (LM3). The results show a similar performance, but the HLM4 has the advantage of estimating more parameters of the underlying model.

For the assessment of PMU calibration systems, the proposed method that can bring more information about the system components, especially for the cases when one needs to investigate the behavior of the measurements during a few cycles.

Benefits:

1 – estimation of without any initial guessing of its value.

2 – assessment of a hybrid estimator (HLM4) under step conditions including

a) uncertainties in tau estimation;

b) LM with 4 parameters (Frigo uses 3 parameters).

c) considering lower levels of SNR (60dB)

d) models for intermediate phasors that can be used in conjunction with previous approaches that employ adjacent windows.

Besides, with the models we proposed, one can continuously measure the parameters that are supposed to remain constant during a step change, and check if they really are.

Also, what would be the deficit of the measurement if the 5th window shown in Fig. 6 would be excluded from the analysis, related to the calibration outcome of the PMU involved? 

Recent works [10] point out the need of a better understanding on how PMUs behave during fast transients. Events with duration of a few cycles cannot be observed if one simply ignore the windows that contain the transients. Even for some cases already prescribed in the IEEE standard, the time response of P-class PMUs needs to be calibrated within a few cycles. Moreover, the usual approach performed by numerous works, (among them NIST calibrations, described in [4], and Frigo et al. [13]), does not exclude all the samples of the window containing the step, but take the samples partially, depending on the step occurrence. Our proposal shows the possibility of considering all the samples in a single window.

In addition, the paper need a thorough proofreading

We did our best to proofread and correct the errors found in the paper.  
  
  
  
Reviewer #2 (Required Remarks for the Author):   
  
The language is quite ok.   
The mathematics and explanations of the mathematics and tests are quite ok.   
The conclusions are also good.   
The problems lay in the problem definition, and reasons for the work.   
The proposed tests are for signals for which a PMUs response is undefined and as such a real calibration can not be done, since there are no requirements set in the PMU standard for these kinds of signals. In which situation/for what purpose do they propose/anticipate it to be used?

A test can still be valuable but not for calibration purposes, this should be reflected in the first section where reasons for the work is described. 

Further, the benefits of such a test is not very well described in the first section (while it is in the conclusion)   
In summary the first section lacks in clarity regarding the reason for and the practical use of the method proposed.   
  
The same responses to rev1 can be easily adapted/referred here.

Reviewer #3 (Required Remarks for the Author):   
  
(Note, due to the limitations in text formatting, mathematical symbols in this review are written in LaTex)   
  
Overview: Calibration of PMUs and PMU calibration systems is a very important topic and improvements in analysis techniques are essential for the technology to move forward since future requirements for PMU performance are expected to be more stringent than they are in the present standard. This paper is timely and important and it must be clear about what it can and cannot provide to the community.   
  
In the first paragraph of the introduction states: "Recent developments towards the calibration of PMUs for distribution grids demand lower uncertainty levels than the current systems, which were designed for the context of transmission grids [3]" This statement implies that the proposed method will provide higher uncertainty levels than existing methods. The statement must be substantiated or else state that future work is needed to determine if this method is more accurate. Accuracy is especially important in the determination of signal phase, which is assumed to be crucial parameter for distribution PMUs. 

We shall emphasize that the statement we have made cites a reference. Nowhere in the paper we have claimed that our proposition meets the demands reported in [3] or outperform other methods available in the literature. As pointed out by Reviewer #1, we did not even run comparative performance tests against other solutions and just reported the performance of our package of methods. In the conclusions we stated “Within the limits reported, the proposed method can give reliable and accurate results to assess PMU calibration systems.” We find it surprising that the distinguished Reviewer could make such a longshot implication without any factual basis whatsoever. As distribution PMUs are subjected to signal quality that is poorer than in a well-defined calibration scenario, quantifying the accuracy of these PMUs is a metrological challenge which is beyond the scope of this work in particular, but we mentioned it in the introduction to give a general context of the recent developments and perspectives.

Anyhow, by understanding the importance of comparing the performance of our proposition against well-stablished solutions available in the literature, we now report in the paper results of additional computational simulations to show that:

1. in comparison to previous work [13], which employed an LM estimator with 3 parameters, our proposed estimator, with the addition of one more parameter in the LM estimator and of uncertainties in the estimation of , have comparable performance. (section III.C, Table IV);
2. The SNR has great influence on the performance of estimators, and one shall take this into account when assessing a calibration system (section III.C, Fig.5);
3. The definition of the reference phasor influences the TVE and TR estimates, especially for transients of a few cycles (P-class step tests).

In the forth paragraph: "...and sets the phasor estimates where the discontinuity occurs as those of obtained from the previous or following window." change: "those of obtained" to "those obtained"

Agreed. We changed the text as suggested.

Equation number 4 is skipped, equation 3 is followed by equation 5. Equation 4 appears on page 3 in section D and should be equation 14 once the numbering is corrected. 

Thanks for pointing this out. We made a general revision on cross-references.

Following Figure 1, it is unclear what the author means by "intermediate phasor estimates". I think this refers to the phasor values during the rise time of the step itself. the step generally rises faster than the sampling rate so (5) and (6) can be used to find intermediate values between two samples. This is important to know as a reference value if the step occurs exactly coincident with a synchrophasor report but not exactly coincident with a sampled value. The equations work, however "T" and \tau are not the correct values to use. "T" should be "dt" and another symbol should replace "\tau" to denote the length of time following the sample time at which we are trying to interpolate the value. For example, if the sampling period "dt" is 1/5000 s and the step occurs half way between two samples then "\tau" is 1/10,000 s. "T" is used elsewhere in the paper to represent the total duration of the sampled window and \tau is the time from the beginning of the window until the step. Using these values in (5) and (6) would basically mean that the step rise time is the duration of the window. Using those values, (5) and (6) would only provide some kind of average value of amplitude or phase over the entire sampling window, which I fail to see as a useful metric. I recommend changing "T" to "dt" and providing a new symbol to represent the distance between the samples where the step actually occurred. 

The reference phasor values are those given by equations (4) and (5), whose parameters can be estimated by several means or known a priori. In our case, we have chosen to employ the models (1) and (2), sampled at , in conjunction with the LM method to estimate the needed parameters, over window of 1000 samples inside which the step happens. The value of is estimated by a detector based on the instantaneous frequency of the Hilbert’s analytical signal related to (1) and (2). Therefore, intermediate reference phasor values given by (4) or (5) hold for the whole analysis window, including the rise time of the step itself. It is a proposal to be used as a well-defined reference phasor, compatible with TVE calculation, instead of using reference values from adjacent windows.

However, nothing prevents a more detailed view of the phasor estimates within the window, since the transient location can be accurately determined, along with the phasor parameters before and after the short transient.

Equation 4 is mis-numbered (should be 14),

Thanks for pointing this out. We made a general revision on cross-references.  
  
Top of page 5: The PMU standard does not state that synchrophasors must be obtained relative to the center of an analysis window. Synchrophasors may be obtained in any manner as long as the performance meets the requirements of the standard.   
  
Agreed. Our particular choice was taking the synchrophasors related to the center of the window. We changed the text accordingly.

Conclusion: Recommend mentioning that Levenburg-Marquadt initial parameter sensitivity should be studied in future work. When calibrating PMU calibration systems, actual parameters are not available to the calibrator calibration system. Some anecdotal testing has shown that due to local minima or possibly phase wrap issues, Lebenburg-Marquadt is sensitive to errors in initial phase error as small as 3 degrees: estimated parameters with a negative magnitude and phase error of pi resulted (which actually provides the same signal). Also recommend that future work be suggested to test sensitivity to \tau estimation when \tau of a magnitude step near or coincident with a zero crossing. The PMU standard says nothing about the required phase at which the steps may or may not occur but this and previous methods may not work when the step occurs at or near particular points in the phase, such as zero crossings.

We agree that this study should be extended to assess the behavior of the estimator when subjected to large errors in the initial guessing.

The sensitivity of estimation in special situations can be high with lower SNR and vary depending on the location of inside the window. However, the use of a more elaborated detector can be implemented to deal with these cases. For example, we have already run computational simulations that show that it is feasible to attain reliable estimates in a scenario where a magnitude step occurs near or coincident with a zero crossing, for SNR as low as 50 dB, by using a detector based on the instantaneous magnitude of the Hilbert’s analytical signal. Similarly, phase steps that occur near a local maximum (peak) or minimum (valley) of the waveform can be reliably detected using a detector based on a suitable combination of the instantaneous magnitude and phase of the Hilbert’s analytical signal.

Add the experimental results; share the matlab codes.