Sobre os comentários do revisor:  
  
"- in a calibration architecture like the one proposed by the Authors,   
the location of the step change should be known in advance since the   
user defines the signal model of the test waveforms. "

R: Our proposed architecture is based on the existing designs already published by the few National Measurement Institutes that detain the technology for the calibration of PMU calibrators. In such designs, the reference values are estimated by processing sampled signals, instead of using the nominal values set in the generation module, because the sampling systems are more stable and accurate than the generation, amplifier and transducer circuit. There are two published solutions to calibrate PMU calibrators for the step tests, both demanding a reliable estimation of the step location in advance: 1) adjust the observation windows to take only the steady state signals before and after the step [ref NIST]; or 2) use a parametric model with iterative procedures [ref INMETRO]. As the IEEE Std does not specify the step location inside the frame, it is often different for each particular implementation. In addition, there is a growing demand for calibrating commercial calibration systems that do not allow the user to set all parameters of the generation signal model. Therefore, estimating the step location from the sampled signal is a need.

PE: Sim. E daí? Alguma informação de referência é necessária. Caso bastasse   
contar com a referência especificada nominalmente, o calibrador seria   
desnecessário! Ou estou entendendo errado a estória?

MBM: Na arquitetura proposta, que é a mesma filosofia empregada em todos os NMIs que se propõem a calibrar calibradores, os valores de referência (mais confiáveis) são obtidos pelo sistema de amostragem. Há sistemas que se baseiam nos valores nominais (como o calibrador da FLUKE), em que se confia nos valores configurados para o gerador. Todavia, mesmo estes sistemas devem ser periodicamente calibrados através de amostragem digital.

A localização do degrau para os calibradores não está definida em norma, podendo acontecer em qualquer instante, dependendo da implementação particular. Nem sempre se tem acesso ao valor nominal da localização do degrau. Estimar a localização a partir de um sinal amostrador é necessário, seja para utilizar o procedimento de ajuste das janelas (como o NIST faz), seja para alimentar os modelos paramétricos como o nosso.

" Moreover, the accuracy of the proposed estimation techniques and the   
reliability of the detectors is not compliant with the requirements of a   
calibration system whose accuracy should exceed the expected one of the   
DUTs by at least one order of magnitude;

R: We propose this technique to assess the generation modules of PMU calibrators, particularly for the estimation of step location inside a particular frame, not PMUs directly. PMUs are not required in the IEEE Std to provide any estimation of the step location inside the frame.

Nonetheless, for the best of our knowledge, commercial PMUs can detect non reliable phasor estimations with some transient detection technique, at rates up to 120 frames/s (half cycles for 60 Hz). It would give the lowest time resolution of 8.33 ms. Our procedure gives at least 0.416 ms (2 \* 1/4800) of time resolution for the typical SNR, one order of magnitude below.

PE: PMU comerciais estimam e entregam a localização do salto? Há   
referências publicadas sobre isso? De fato, nada foi verificado sobre a   
ordem de grandeza das estimativas. Mas se não houver informação sobre o   
erro de estimativa de localização de salto de um PMU típico, não há como   
afirmar nada sobre a questão.

MBM: Eu não tenho notícia de PMUs que localizem os saltos dentro da janela. No máximo implementam um detector de transiente genérico, que indica que a estimativa do fasor não é confiável naquele instante. Esse procedimento na melhor das hipóteses identifica com resolução de meio ciclo (8,3ms para 60 Hz), o que é uma ordem de grandeza maior que a nossa (0,416ms para fs = 4800 Hz).

Não há requisitos de exatidão para localização do salto dentro da janela, porque esta não é uma grandeza que se espera receber de um PMU. Esta é uma preocupação de projetistas de sistemas de calibração, ao menos por enquanto.

OBS: Vejo que um aspecto que pode confundir os revisores é que usamos a letra *n* para designar as amostras dentro de uma janela de observação, mas na norma a mesma letra é usada para designar o *frame*. Talvez valha a pena trocar por outra letra.

Por exemplo:

No nosso caso: fs = 4800 Hz, =1/fs, ; n é o número da amostra

Na norma: fps = 60 frames/s, X(n) é o fasor do frame n, atualizado a cada 1fps segundos.   
  
"- the literature review has to be improved with specific reference to   
other existing PMU calibration architectures, the detection of power   
quality events (like step changes), and the analysis of PMU performance   
during transient conditions;"

R: We have improved the Introduction section accordingly.  
  
PE: Isso é relativamente fácil de fazer.

MBM: Sem problemas, trabalhando nisso.

> - the problem of defining phasor and frequency during transient    
> events is currently subject of many papers in the literature. ...

During a step change, the signal energy is not conveyed by the only fundamental component, but it is spread all over the spectrum. In such a scenario, the definition of phasor and frequency associated to the fundamental component looses its significance. Also for this reason, the IEEE Std introduces the concept of response time and almost neglects the PMUs' accuracy in the presence of a step change.

R: Yes, the problem was well presented in the literature for many years. Although there are no limits prescribed in the IEEE Std, recent works [refs] highlight the practical need for extracting some information about frequency (and ROCOF) under step transient events, even though the usual definitions loose significance, to avoid the real catastrophic events recorded in grids with high renewables penetration. The underlying frequency concept for these cases was introduced in [ref], and is not formally adopted in the IEEE Std yet. In this context, our work aims to contribute on the particular topic of estimating the underlying frequency under steps, not limited to the IEEE Std absence of prescriptions for FE during steps.

MBM: Trabalhos recentes, como do Paul Wright, ressaltam a importância de extrair informação útil neste cenário, principalmente relacionados à frequência, mesmo que tenha uma perda de significância pela presença do degrau. A discussão sobre frequência subjacente durante afundamentos foi inclusive tema de artigos recentes deste mesmo evento. Assim, mesmo que a norma não defina limites, não podemos nos restringir a ela, pois novos problemas estão surgindo. A norma também não definia limites para ensaios dinâmicos em 2005, quando vários estudos foram feitos e levaram à revisão de 2011.

> recent literature has introduced alternative performance indicators   
> like Normalized Root Mean Squared Error and Goodness of Fit;  
  
R: Yes, these are alternative performance indicators to the standard TVE, and they can be used to evaluate the goodness of fit of a phasor estimate. The scope of in this paper is not estimate the phasors, but only step locations inside a frame and underlying frequency.

PE: Bem, a estimativa da frequência do fasor, tal como proposta, é fraca   
mesmo e foi ressaltado que não serve para propósitos de calibração. Não   
é o proposito do artigo discutir o que o revisor sugere. Sendo assim,   
uma solução de meio termo seria mencionar a questão muito brevemente.

MBM: Até onde entendo, estes indicadores são para medir a estimativa do sincrofasor, não da frequência, como uma alternativa ao TVE. Não está no escopo deste trabalho. Pensei em comentar algo brevemente na Introdução como parte da revisão bibliográfica e ressaltar que estamos focados na estimação dos saltos e da frequência, mas o fasor mesmo seria estimado pelo método paramétrico.  
  
"> - the implementation of Hilbert Transform strongly affects the    
accuracy of the instantaneous estimates of amplitude, frequency and    
phase. The HT computation can be obtained via two main approaches:...  
  analytic signal might be not totally reliable for this application;  
"

R: We show by numerical simulations that for the detection of steps the proposed approach is accurate. For the underlying frequency, we report its limitations and noise dependence for typical SNRs, as a reference for future improvements. We consider that it is possible to achieve the lower uncertainty reported in our previous work [ref] by implementing a denoising technique.  
  
PE: Isso foi explicitamente dito no artigo, para o caso da frequência. Para   
o caso da localização de saltos, os resultados deveriam servir de prova   
da eficácia do método. Há alguma literatura sobre métodos paramétricos   
para localização de saltos, com resultados publicados, que possam servir   
de benchmark? Marcelo, chegou a fazer testes com o LM com tau   
desconhecido?

MBM: Estamos justamente mostrando pelos testes que, a despeito das distorções, o método é eficaz para a localização dos saltos. Não conheço literatura sobre método paramétrico para identificar saltos, posso tentar procurar algo. Me lembro que quando tentei utilizar o LM com o nosso modelo particular para estimar o tau inclusive, tive muitos problemas com a convergência, pois o perfil da função de custo não era suave.

MBM2: Estou procurando alguma referência, mas não encontrei ainda.

  "> - the performance of the proposed detectors should be tested also   
in off-nominal or other dynamic conditions, in order to verify the    
robustness of the proposed thresholds.

R: The proposed detectors are not intended to be used in other dynamic conditions, but only the step-like ones based on the models prescribed by the IEEE Std. Off-nominal frequencies are not prescribed in the step tests. We run tests for other frequencies and there is no significant change. That happens because, for the detection signal given by eq. (8), the influence of the fundamental frequency is strongly reduced by the subtraction of the median of fi.

OBS: será que ele se refere a outras grandezas off-nominal? Como maior/menor profundidade dos saltos?

PE: A proposta tem um escopo bem definido. Testes com sinal off-nominal não   
vão produzir diferença alguma pois a nominal é removida nas contas   
(citar a equação). Outras condições dinâmicas é muito vago como escopo.   
No mais, foram avaliados os cenários de teste do IEEE Std.

MBM: Os testes de degrau da IEEE Std não preveem frequências off-nominal. Estamos projetando um calibrador, não um PMU para receber qualquer sinal ! No máximo podemos variar ligeiramente a frequência do sinal em torno da nominal, para refletir alguma incerteza na geração, o que não é exatamente off-nominal. Não espero alteração significativa no resultado da detecção. Para os sinais de detecção, (eq (8)), a subtração da mediana de fi, que é uma estimativa de fu, praticamente elimina a influência do valor de fu na detecção.

"> - the adopted sampling rate is rather low (4.8 kHz). This means that   
the spectral bandwidth of the observed phenomena is limited to 2.4 kHz    
(assuming that an anti-aliasing filter is employed). Furthermore, this  
> low sampling rate limits also the resolution of the detectors as the   
> step change is associated to a given sampling time;  
"  
R: We report a worst case for the step detection, that is already sufficient for the assessment of step locations. We can run for higher sampling rates, what will increase the time resolution and consequently the performance, at a higher computational cost.

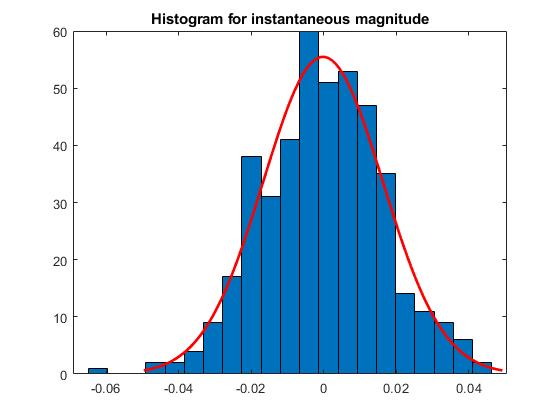
PE: Aumentar a taxa de amostragem só vai melhorar os resultados: tanto em   
resolução temporal quanto em desempenho de detecção. Isso não é   
problema. Só aumenta o custo computacional.

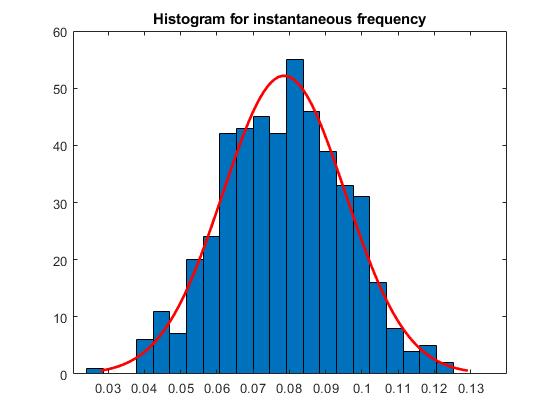
MBM: Sim, podemos fazer com uma frequência maior, que vai melhorar tudo…  
  
"> - the adopted thresholds rely on the assumption of normal   
distribution. How are the mean and standard deviation computed, i.e.   
over which time-window? Or is the algorithm run off-line? The Authors   
should provide a qqplot or a histogram in order to verify this   
assumption;  
"  
R: We use frames of 6 cycles (480 samples), and all calculations are done with the samples within the frame. The thresholds for the detection signals given by eqs. (7) and (8) are proportional to the median, which is a robust estimate of the standard deviation of the signal in the presence of steps. As the calculation of f\_i[n] involves only linear operations, provided the noise in x[n] is Gaussian, the noise in f\_i[n] will be also Gaussian. We verified that, as shown in the histograms below, for SNR=30dB.

PE: Sim, é uma assunção feita (sem verificação formal, porque vai falhar)   
em 90% de trabalhos de engenharia)! Aqui, basta explicar que a conta   
feita (sobre uma janela de observação) para obter o limiar (citar as   
equações) já é uma estimativa robusta do std da distribuição do sinal de   
detecção. É rodado online. Se for para mostrar histograma (o que eu acho   
ridículo), faz com SNR baixo. Talvez, baste mostrar para o revisor na   
carta de resposta.

MBM: de novo, me parece que o revisor não está entendendo que estamos tratando das amostras dentro de uma janela. Se ele se refere aos limiares dos detectores, usamos a mediana, não o desvio padrão, com as amostras da mesma janela, excluídas 5% no início e no final. Além do mais, essas estatísticas podem ser aplicadas a qualquer distribuição.

De qualquer forma, é aproximadamente normal. Figuras geradas com SNR = 30dB:





No caso, entendo que estou rodando off-line, pois eu coleto as amostras e processo no Matlab depois. A implementação será feita online mais tarde. Mas qual a diferença, para esse aspecto, se o algoritmo roda online ou off-line?

> - the tested SNR levels prove the robustness of the adopted technique   
> but are not realistic for PMU applications where a lower level of    
> noise is expected. In this context, what is the delay and distortion  
> introduced by the measurement chain: Amplifier + Divider?  
  
R: Indeed, SNRs lower than 60 dB are not realistic for calibration purposes. The performance reported for 60 dB is a worst case for calibration, expected to be better as SNR increases. The analysis for SNRs lower than 60 dB gives us limits for other potential contexts. Regarding the delay and distortion introduced by the Amplifier + Divider, we expect to have some information when we run tests with these components, but it is beyond the scope of this paper to have an accurate characterization of them. That is why they were wiped out of the circuit for the validation of the method.

PE: De fato, SNR abaixo de 60dB é irrealistico no contexto de calibração.   
Mas pode-se dizer que a análise é útil para outros contextos.

MBM: Para a localização dos saltos eu diria que, se eu comprovo a eficácia com ruído alto, a tendência é melhorar para níveis de ruído mais baixos. Para estimação da frequência faria sentido reduzir o ruído, seja com hardware melhor, seja com método de filtragem de ruído.

"what is   
the delay and distortion  
> introduced by the measurement chain: Amplifier + Divider?" Não está   
> registrado, de forma consolidada, nos erros reportados de localização   
> de saltos?

MBM: Seção IV.B começa com: “For a preliminary investigation, we connected the analog

output of the Generation module directly to the analog input

of the Sampling module, as to eliminate the influence of the

Voltage Amplifier and the RVD.”

Só teremos essa informação quando forem feitos testes com os circuitos inseridos.

"As a main remark, the reference section should be largely extended,  
> to reflect the significant work currently under development in   
> several  
> labs on the subject of PMU calibration systems."

R: We modified the text accordingly.

Ok. Fácil de resolver.

MBM: já comentado anteriormente  
  
"More tests should be provided in the numerical simulations to assess  
> the performance of the detectors. In particular, it seems to have  
> little sense testing each detector only against the specific steps it  
> looks for. On the contrary, the performance of the three detectors  
> should be compared in all situations. Concerning the simultaneous  
> presence of amplitude and phase steps, considering them as  
> simultaneous represents again a too favorable situation for the   
> hybrid  
> detector. Test with non-simultaneous step would represent a “fairer”"

> ======= Review 2 =======  
>  
>> \*\*\* Guidance: Please provide specific recommendations that you would   
>> like to suggest to the author for improving the paper.  
>  
> The authors may want to explicitly define the contribution of this  
> paper with respect to reference [6] and possibly compare (or rather  
> reference instead of duplicating) the results to [6] for the common  
> cases.

R:

MBM:

É preciso explicitar que em [6] nós testamos o desempenho de um detector baseado em fi, em apenas alguns casos particulares de phi0, que não coincidiam com os casos críticos. Não reportamos avaliação do desempenho para os casos críticos em [6]. Valeria a pena incluir esta informação neste trabalho?

> ======= Review 3 =======  
>  
>> \*\*\* Guidance: Please provide specific recommendations that you would   
>> like to suggest to the author for improving the paper.  
>  
> The topic is interesting and coherent to AMPS’s scope. The paper is  
> well written.  
> As a main remark, the reference section should be largely extended,  
> to reflect the significant work currently under development in   
> several  
> labs on the subject of PMU calibration systems.

R: We extended the references accordingly.

> More tests should be provided in the numerical simulations to assess  
> the performance of the detectors. In particular, it seems to have  
> little sense testing each detector only against the specific steps it  
> looks for. On the contrary, the performance of the three detectors  
> should be compared in all situations. Concerning the simultaneous  
> presence of amplitude and phase steps, considering them as  
> simultaneous represents again a too favorable situation for the   
> hybrid  
> detector. Test with non-simultaneous step would represent a “fairer”  
> test

R: As we are designing calibration systems, our intent is to assess primarily the tests prescribed in IEEE Std. The simultaneous test is not prescribed in the standard, but is included as a first stage to reproduce real voltage sags in future works.

We now report the results for the hybrid detector for the 3 cases (phase step only, magnitude step only, simultaneous steps). The primary detectors perform slightly better when used to the specific scenarios for which each of them were designed.

PE: A resposta deve ser no sentido de replicar testes especificados na   
norma IEEE. E sim, faz sentido testar detectores projetados   
especificamente para um cenário de teste. O que não faz sentido é testar   
um detector de saltos em mag para casos de salto em fase. Cada   
maluquice. Falta do que comentar. A ideia do detector híbrido é   
justamente prescindir do conhecimento prévio do tipo de salto envolvido.  
Fazer o que o revisor quer só vai poluir o artigo com informação   
desnecessária. Para ser franco, eu só reportaria testes com o detector   
híbrido, nas 3 condições de saltos. Testes com saltos não-simultâneos   
ficam para o futuro. Ou então mostraria um caso com voltage sag (2   
saltos).

MBM: Farei então com o detector híbrido: 1) salto somente de fase, 2) somente de magnitude, 3) simultâneo; Reporto os resultados em breve.

MBM: Para mim, o que não faz sentido é usar um detector em uma situação para a qual ele não foi projetado. Se fizermos isso, teremos 9 combinações. De qualquer forma, se substituirmos pelo detector híbrido, como comentado anteriormente, podemos tratar esta questão...