Merging conventional and phasor measurements in state estimation: a multi-criteria perspective

Bruna Tavares
INESC TEC
Porto
Portugal
bruna.c.tavares@inesctec.pt

Victor Freitas
UFSC – Universidade Federal de Santa Catarina
Florianópolis (SC), Brazil
victor.silva@posgrad.ufsc.br

Vladimiro Miranda
INESC TEC and FEUP
Universidade do Porto
Portugal
vmiranda@inesctec.pt

Antonio Simões Costa
UFSC – Universidade Federal de Santa Catarina
Florianópolis (SC), Brazil
simoes@labspot.ufsc.br

***Abstract* — This paper presents a new proposal for sensor fusion in power system state estimation, analyzing the case of data sets composed of conventional measurements and phasor measurements from PMUs. The approach is based on multiple criteria decision-making concepts. The equivalence of an L_1 metric in the attribute space to the results from a Bar-Shalom-Campo fusion model is established. The paper shows that the new fusion proposal allows understanding the consequences of attributing different levels of confidence or trust to both systems. A case study provides insight into the new model.**

***Keywords* — state estimation, multiple-criteria, information fusion, correntropy**

I. Introduction

The addition of synchrophasors (PMUs – phasor measurement units) to power systems is in general admitted to lead to a better state estimation (SE), but the accuracy of the measurement values received at the control center depends not only on the device but also on a transmission chain, which has its own reliability and is prone to error injection. These units provide voltage measurements, which are usually adopted as state variables, as well as phasor current measurements. An estimation procedure based on voltage and current measurements, when expressed in the appropriate coordinate reference and admitting that the system is observable, may be designed to be a linear model – therefore, leading to easier calculations.

It may be foreseen that one day, all SE may be based solely on PMUs. However, in present days and for a number of years, one will have to live with both conventional and phasor measurements arriving at the control center. It is conceivable that one could profit from the combination of both sensor systems to perform a more accurate SE. Therefore, it is important to understand the properties and characteristics of an SE when information from both sources is merged.

Several processes have been proposed, in order to achieve a convenient contribution of both conventional and phasor measurements (also called *fusion*) to power system state estimation. The hybrid simultaneous state estimation methods include the PMU measurements straightforwardly in the SE task, processing simultaneously both conventional and PMU measurements [1],[2]. The decentralized estimation fusion methods perform SE for both systems separately and then combine the resulting state estimates [3],[4]. These latter models are based on Multisensor Data Fusion Theory, a field already explored in other areas such as Aerospace, Information Theory and Signal Processing [5].

In all those methods, cost functions are based on the least squares principle, i.e., on the minimization of the variance of the error distribution. However, this criterion is an optimizer only if the assumption that the error distributions are Gaussian is verified. In practice, it seldom is. This has motivated the proposal of other criteria based on Information Theory, such as Correntropy, which is particularly useful when the measurement set is contaminated with gross errors [6].

This paper provides an analysis of the properties of a state estimation procedure when a suitable combination of information provided by each sensor system (conventional and phasor measurements) is attempted. To avoid the contamination of a data set by any gross error contained in the other data set, an MCC – Maximum Correntropy Criterion – is adopted (instead of WLS – weighted least squares). Each sensor system data is affected by its own error weights but, on top of that, a coefficient is defined in order to assign variable levels of relative importance to each sensor system.

Adopting a multiple criteria point of view, the Pareto-optimal border resulting from considering an independent MCC criterion for each system may be explored. Fusion, in this sense, corresponds to a balance or compromise between the importance given to one and the other system. In this paper, some properties for such compromise are investigated, under distinct assumptions.

II. Conventional Power System State Estimation

A. Gauss-Newton Solution Method

The conventional power system SE solution relies on a system of overdetermined and inconsistent set of nonlinear equations given by:

A. Simões Costa and V. Miranda acknowledge the support given by CNPq (Brazilian National Research Council), under project no. 400799/2014-6. B.Tavares and V.Miranda acknowledge the support of the ERDF – European Regional Development Fund (COMPETE 2020 Programme) and of FCT - Fundação para a Ciência e a Tecnologia (project POCI-01-0145-FEDER-016731 INFUSE).

978-1-5090-4000- 1/17/$31.00 ©2017 IEEE

$$z = h(x) + \eta \quad (1)$$

where $\boldsymbol{x}$ is the $n \times 1$ state vector composed by bus voltage magnitudes and phase angles, and $\boldsymbol{z}$, $\boldsymbol{h}(\cdot)$ and $\boldsymbol{\eta}$ are $m \times 1$ vectors containing the measurements, nonlinear functions relating measured quantities and states, and measurement errors, respectively. In the absence of bad data, $\boldsymbol{\eta}$ is assumed to be normally distributed, zero mean, and uncorrelated, whose $m \times m$ diagonal covariance matrix is denoted by $\boldsymbol{R}$, with its i^{th} entry equal to $\sigma^2_{m,i}$, which is the variance of the error of measurement i.

The conventional approach for power system SE is based on the weighted least-squares (WLS) method, which minimizes the weighted of squared residuals (weighted variance of the error distribution):

$$\min_{\hat{x}} J(\hat{x}) = \frac{1}{2}[\boldsymbol{z} - \boldsymbol{h}(\hat{\boldsymbol{x}})]^T \boldsymbol{R}^{-1} [\boldsymbol{z} - \boldsymbol{h}(\hat{\boldsymbol{x}})]. \quad (2)$$

The above problem can be solved through the Gauss-Newton method, leading to an iterative process in which so-called normal equation is solved in each iteration [7]-[9]:

$$\boldsymbol{G}\Delta\boldsymbol{x} = \boldsymbol{H}^T \boldsymbol{R}^{-1} \Delta\mathbf{z} \quad (3)$$

where $\boldsymbol{G} = (\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H})$ is referred to as gain matrix; $\boldsymbol{H}$ is the $m \times n$ Jacobian matrix of $\boldsymbol{h}(\boldsymbol{x})$ calculated at a given point $\boldsymbol{x}^k$, and $\Delta\mathbf{z} = \boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}^k)$.

In the end of each major iteration, the solution of (3) yields vector $\Delta\boldsymbol{x}$ of increments to the states, so that the updated state vector is obtained as

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \Delta\boldsymbol{x} \quad (4)$$

This process goes on until $||\Delta\boldsymbol{x}||$ becomes smaller then a pre-specified tolerance. The estimation error covariance matrix can be obtained at the state estimator convergence as [7]

$$\boldsymbol{P} = \boldsymbol{G}^{-1} = (\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H})^{-1} \quad (5)$$

B. Power System SE Considering A Priori Information

Prior knowledge about the state variables is often available, and can be taken into account in the estimation process as *a priori* state information. For that purpose, a degree of confidence should be assigned to such data, under the form of a covariance matrix. In PSSE problem, *a priori* information on the state variables can be modeled as an extra quadratic term added to the weighted least-squares criterion [10]. Therefore the problem in (2) becomes:

$$\min_{\hat{x}} J(\hat{x}) = \frac{1}{2}[\boldsymbol{z} - \boldsymbol{h}(\hat{\boldsymbol{x}})]^T \boldsymbol{R}^{-1} [\boldsymbol{z} - \boldsymbol{h}(\hat{\boldsymbol{x}})] + \frac{1}{2}(\hat{\boldsymbol{x}} - \bar{\boldsymbol{x}})^T \boldsymbol{P}_0^{-1} (\hat{\boldsymbol{x}} - \bar{\boldsymbol{x}}) \quad (6)$$

where $\bar{\boldsymbol{x}}$ denotes the $n \times 1$ vector of *a priori* state values and $\boldsymbol{P_0}$ is the $n \times n$ corresponding covariance matrix.

The optimality conditions for problem (6) lead to the following extended version of the normal equation [10]:

$$(\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H} + \boldsymbol{P}_0^{-1})\Delta\boldsymbol{x} = \boldsymbol{H}^T \boldsymbol{R}^{-1} \Delta\mathbf{z} + \boldsymbol{P}_0^{-1} \Delta\bar{\boldsymbol{x}} \quad (7)$$

where $\Delta\bar{\boldsymbol{x}} = \bar{\boldsymbol{x}} - \boldsymbol{x}^k$.

III. Information Theoretic State Estimation

A. Correntropy and Induced Metric

In [6], a new performance criterion for State Estimation was proposed, inspired in developments in Information Theoretic Learning: this criterion is the Correntropy of the error distribution. Correntropy is a function that considers all even moments of a probability distribution function (pdf).

The Correntropy C estimation of an error distribution, representing ε_k errors in a *N* dimension error space in State Estimation, is

$$C = \frac{1}{N} \sum_{k=1}^{N} G(\varepsilon_k, \sigma^2 I) \quad (8)$$

where *G* is the Gaussian kernel with standard deviation σ (*I* is an identity matrix, implying independence among errors).

Maximizing Correntropy has the effect of searching for a pdf that has a maximum value at the origin (*x = 0*). Therefore, by maximizing Correntropy, one achieves a result usually very close to minimizing the Entropy of a same distribution with mean equal to zero. The Entropy of an error distribution is a measure of the information contained by the pdf of the errors, in a non-parametric way. The classic approach to State Estimation uses regression in a Least Squares (LS) sense, but this model, in fact, is assuming that the error distribution is Gaussian. Actually, this classical approach relies solely on minimizing the Variance of the error distribution – and this is only an optimal model if the Gaussianity assumption remains valid, which is usually not the case.

It has been shown that a SE procedure, under a Correntropy criterion, is very good at treating gross errors in data as outliers (and therefore ignoring them), contrary to any LS formulation, where any gross error may severely contaminate a subset of the estimates. It has also been shown that there is a geometric interpretation for the use of a Correntropy criterion, which links this function to a specific metric in space called CIM – Correntropy Induced Metric. CIM, in the form of distance to the origin, has the expression

$$CIM(\varepsilon) = \left(G(0, \sigma^2 I) - V(\varepsilon)\right)^{1/2} \quad (9)$$

It becomes obvious that minimizing CIM is tantamount to maximizing Correntropy. Therefore, a State Estimation procedure adopting a Maximum Correntropy Criterion (MCC) is just searching for a an error vector as small as possible, in the CIM sense, while a procedure adopting a LS criterion is searching for a vector as small as possible in the Euclidean sense (sum of the squares of the coordinates).

B. State Estimation Based on Correntropy

If instead of attempting to minimize the variance, one goes for minimizing the correntropy of the error distribution, one has an Information Theoretic State Estimation procedure [11], which can be described as follows.

The MCC method aims act to extracting the maximum amount of information from the residual distribution with regard to the available measurement set. This is accomplished through correntropy concepts, which measures the similarity between measured and the estimated values within a given "observation window". The latter is defined on the basis of the Parzen window technique [12]. The state estimates are obtained by maximizing the following objective function:

$$J_{MCC} = \max_{\hat{x}} \sum_{i=1}^{m} \mathrm{G}(z_i - h_i(\hat{x}), \sigma_i^2) \quad (10)$$

A desirable feature of correntropy is that it is able to take into account all even moments of the residual function $z - h(\hat{x})$, depending on the width σ of the Parzen window. Such parameter plays an important role to the optimization process. For initially large σ values, the MCC method is basically equivalent to the WLS approach, since the correntropy reduces itself to the familiar Euclidian norm (L_2) of the residuals. When σ is progressively reduced during the optimization procedure, the residuals with large magnitudes are most affected, substantially reducing their influence on the final estimation solution.

In this paper, one will not enter into technical details on how to reach the solution. It is important to notice that the structure of matrices is preserved and the MCC solution can be guaranteed by solving equations whose form is similar to the conventional Gauss-Newton solution [11],[13].

IV. BAR-SHALOM-CAMPO FUSION IN STATE ESTIMATION

A. Bar-Shalom-Campo fusion

The Bar-Shalom-Campo fusion formula [14] may be applied when there are two distinct sensor systems. Its application is particularly attractive when the correlation between the estimation errors relative to the sensor systems can be neglected. Both conditions apply to the problem addressed in this paper, since we are considering only two systems: conventional and phasor measurements. Simply said, the optimal global state estimates $\hat{\boldsymbol{x}}^*$ may be calculated [5] as a linear combination of the individual estimates $\hat{\boldsymbol{x}}_i$ as

$$\hat{\boldsymbol{x}}^* = \boldsymbol{W}_C^T\, \hat{\boldsymbol{x}}_C + \boldsymbol{W}_P^T\, \hat{\boldsymbol{x}}_P \quad (11)$$

where $\boldsymbol{W}_C, \boldsymbol{W}_P$ are weighting matrices and the indices C, P relate to conventional and phasor measurements. These weighting matrices $\boldsymbol{W}$ are calculated from solving the following optimization problem:

$$\begin{aligned} &\min_{\boldsymbol{W}} \quad E[(\boldsymbol{W}^T\hat{\boldsymbol{x}}_a - \boldsymbol{x})(\boldsymbol{W}^T\hat{\boldsymbol{x}}_a - \boldsymbol{x})^T] \\ &s.t. \quad \sum_{i=1}^{N_S} \boldsymbol{W}_i = \boldsymbol{I} \end{aligned} \quad (12)$$

where $E[\cdot]$ is the expectation operator, $\boldsymbol{x}$ is the vector of true values for the state variables and $\boldsymbol{I}$ is an identity matrix.

It is possible to show that the optimal estimate is given by:

$$\begin{aligned} \hat{\boldsymbol{x}}^* = &(\boldsymbol{P}_{PP} - \boldsymbol{P}_{PC})(\boldsymbol{P}_{CC} + \boldsymbol{P}_{PP} - \boldsymbol{P}_{CP} - \boldsymbol{P}_{PC})^{-1}\, \hat{\boldsymbol{x}}_C \\ &+ (\boldsymbol{P}_{CC} - \boldsymbol{P}_{CP})(\boldsymbol{P}_{CC} + \boldsymbol{P}_{PP} - \boldsymbol{P}_{CP} - \boldsymbol{P}_{PC})^{-1}\hat{\boldsymbol{x}}_P \end{aligned} \quad (13)$$

where $\boldsymbol{P}$ are the estimation error covariance matrices as in (5). This expression is known as the Bar-Shalom-Campo formula.

With uncorrelated errors, this expression reduces to:

$$\hat{\boldsymbol{x}}^* = \boldsymbol{P}_P(\boldsymbol{P}_C + \boldsymbol{P}_P)^{-1}\, \hat{\boldsymbol{x}}_C + \boldsymbol{P}_C(\boldsymbol{P}_C + \boldsymbol{P}_P)^{-1}\hat{\boldsymbol{x}}_P \quad (14)$$

However, (14) is not amenable for application to large networks. Therefore, an alternative form to prevent those computational difficulties is using (5) to define the conventional and PMU measurement based state estimators gain matrices $\boldsymbol{G}_S$ and $\boldsymbol{G}_P$. Consequently, it is possible to rewrite (14) as proposed by Simoes Costa et al. in [4].

$$(\boldsymbol{G}_S + \boldsymbol{G}_P)\hat{\boldsymbol{x}}^* = \boldsymbol{G}_S\, \hat{\boldsymbol{x}}_S + \boldsymbol{G}_P\hat{\boldsymbol{x}}_P \quad (15)$$

Since the gain matrices $\boldsymbol{G}_S$ and $\boldsymbol{G}_P$ are available from the individual estimator solutions, (15) can be efficiently solved by sparse triangular factorization and forward/backward substitution. Equation (9) will be referred to as *Gain matrix-based fusion formula.*

B. Equivalence of the Gain matrix-based fusion formula and Hybrid SE

Expression (15) may be proven equivalent to the classical SE such as in (2) if one considers that both conventional and phasor data are simultaneously processed by a single LS state estimator, often referred to as a *hybrid estimator*. In fact, replacing in (15) the previous definition of matrices **G** with the appropriate subscripts, one has

$$\begin{aligned} &(\boldsymbol{H}_C^T\boldsymbol{R}_C^{-1}\boldsymbol{H}_C + \boldsymbol{H}_P^T\boldsymbol{R}_C^{-1}\boldsymbol{H}_C)\hat{\boldsymbol{x}}^* \\ &\qquad = \boldsymbol{H}_C^T\boldsymbol{R}_C^{-1}\boldsymbol{H}_C\hat{\boldsymbol{x}}_C + \boldsymbol{H}_P^T\boldsymbol{R}_C^{-1}\boldsymbol{H}_C\hat{\boldsymbol{x}}_P \end{aligned} \quad (16)$$

Furthermore, one may write, from (3), that

$$\begin{aligned} \hat{\boldsymbol{x}}_C &= (\boldsymbol{H}_C^T\boldsymbol{R}_C^{-1}\boldsymbol{H}_C)^{-1}\boldsymbol{H}_C^T\boldsymbol{R}_C^{-1}\boldsymbol{z}_C = \boldsymbol{G}_C\boldsymbol{H}_C^T\boldsymbol{R}_C^{-1}\boldsymbol{z}_C \\ \hat{\boldsymbol{x}}_P &= (\boldsymbol{H}_P^T\boldsymbol{R}_P^{-1}\boldsymbol{H}_P)^{-1}\boldsymbol{H}_P^T\boldsymbol{R}_P^{-1}\boldsymbol{z}_P = \boldsymbol{G}_P\boldsymbol{H}_P^T\boldsymbol{R}_P^{-1}\boldsymbol{z}_P \end{aligned} \quad (17)$$

Then, replacing (17) into (16), we get an expression with the general format

$$(\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H})\hat{\boldsymbol{x}}^* = \boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{z} \quad (18)$$

where matrix $\boldsymbol{H}$ ($\boldsymbol{R}$) contains matrices $\boldsymbol{H_C}$ and $\boldsymbol{H_P}$ ($\boldsymbol{R_C}$ and $\boldsymbol{R_P}$) as partitions. Likewise, $\boldsymbol{z}$ contains vectors $\boldsymbol{z_c}$ and $\boldsymbol{z_p}$. Equation (18) is the basic equation behind the WLS method considering a hybrid estimation structure. Therefore, the fusion process based on the Bar-Shalom-Campo formula provides the same solution as a hypothetical hybrid estimator which would simultaneously process both conventional and phasor measurements, with error distributions assumed Gaussian.

V. MULTI-CRITERIA SENSOR SYSTEM FUSION

A State Estimation procedure, acting on measurements captured by a sensor system, is a form of building an inner (coherent) map of the reality, as perceived by the noise-affected sensor signals. When two sensor systems capture signals, two different images may be built – and which one is closer to the reality it is supposed to represent? If one has two sources of information, in which one to trust the most? Eventually, on some weighted combination of both.

While recent proposals for hybrid simultaneous state estimation methods focus only on sensor precision, the method discussed in this paper makes a distinction at the level of sensor systems, assigning more or less credibility (or trust) to a measurement depending on its original system and keeping explicit the performance of each system, in order to contribute to the global state estimation.

Assume that this performance is evaluated for system x by a cost criterion $J_x(\boldsymbol{\varepsilon})$, with $\boldsymbol{\varepsilon}$ being the vector of measurement errors associated with such system x. A global performance criterion J, for two measurement systems (assumed based on conventional and phasor measurements) will be given by

$$J = \alpha \sum_{i=1}^{C} J_C(\varepsilon_i) + \beta \sum_{k=1}^{P} J_P(\varepsilon_k) \tag{19}$$

with C,P being the number of conventional and phasor measurements and α,β being the credibility factors for each system, such that $\alpha + \beta = 1$. Without loss of generality, let's assume that each performance criterion J_x should be minimized.

The variation of the parameters α and β will influence the compromise reached. Thus, when α is 1 only the measurements from System 1 are considered. By optimizing J with varying α values, we may explore the Pareto-optimal border, defined by the two criteria J_C and J_P, such as illustrated in Fig. 1.

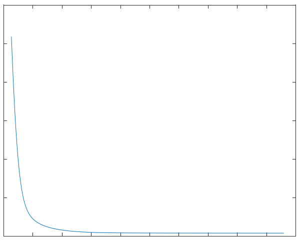

Fig. 1 – Illustration of a Pareto front obtained with the minimization of J under α variation, in an attribute space defined by J_C and J_P.

The ideal point [15], in the attribute space defined by J_C and J_P, would be a point with coordinates (min J_C, min J_P), i.e. the coordinates obtained when $\alpha = 1$ or $\beta = 1$. Such point is unfeasible, as it represents "the best of two worlds": the best image provided solely by the first system and the best image provided by the second system, when considered separately.

A compromise solution, taken as the fusion result, should therefore be a point over the Pareto front yet somewhere close to the Ideal. The concept of "closeness" is related to the metric adopted to define such distance d_n. If any metric L_n is adopted, with $n \in \{1,2,\ldots,\infty\}$, the compromise solution will result from

$$\begin{aligned} \min d^n = (J_C(\varepsilon) - min(J_C(\varepsilon))^n \\ + (J_P(\varepsilon) - min(J_P(\varepsilon))^n \end{aligned} \tag{17}$$

For metric L_1 ($n = 1$), the result is simply given by

$$\min d = J_C(\varepsilon) + J_P(\varepsilon) \tag{18}$$

which is the result obtained with setting $\alpha = \beta = 0.5$. This means that a classical State Estimation procedure, acting upon a single objective function by putting together both sensor systems, is tantamount to finding the point over the Pareto front nearest to the Ideal, in the L1 metric sense. The question then arises as whether this is the best fusion point.

Such solution, minimizing (according to some metric) the distance between the Ideal and the Pareto front, will correspond to a specific pair of (α, β) values. In concept, such values could be derived from historical data providing some measure of probability, denoting on which measurement system in a better image of the real system resulted. This may not be easy to define – it is therefore important to investigate some of the properties of the Pareto front and its relation with the credibility coefficients.

In the formulation above, the coefficient α represents the level of trust or credibility given to the inner map (or image of the reality) built from the conventional measurements – and β = 1 - α is associated with the complementary trust assigned to the phasor measurements. This model is flexible, in the sense that it does not require observability for any of the sensor systems, only for the set of both systems.

If one of the systems still makes the power system observable, say the conventional measurement, then one may set $\alpha = 1$ and still solve the problem. However, if the phasor system is insufficient to render the power system observable, one cannot solve the problem for $\alpha = 0$ (or $\beta = 1$) – but we still can define an image of reality, given by this system, as the limit of the State Estimation when $\alpha \rightarrow 0$, because the power system will be mathematically observable for any value of $\alpha > 0$. Another way to define the Ideal, for the case when one sensor system by itself does not provide observability, is to consider *a priori* information, such as presented in Section V.B. It is in this sense that one may define an Ideal point for the 2-criteria problem, as referred to previously.

VI. Quality of an Estimated Vector

One way to assess the quality of a State Estimation result is to establish an index denoting the similarity between the vector $\boldsymbol{V}$ of estimated nodal complex voltages, and the vector $\boldsymbol{V}^{calc}$ of "exact" voltages (calculated by a power flow routine). This is possible in case studies, when the measurement vectors are built from power flow cases, by adding noise to the calculated values. According to this idea, the following quality criterion M_V may be used [16]:

$$M_V^2 = \frac{1}{N} \sum_{i=1}^{N} \left| \boldsymbol{V}_i - \boldsymbol{V}_i^{calc} \right|^2 \tag{19}$$

It is the length, in a Euclidean norm, of the vector that has, in each component, the module of the difference of the complex voltages (estimated and real) in each node. This index will be used in the following section.

VII. Case Studies

In the following sections, numerical examples will be presented, based on the IEEE 30-bus system. An assumed measurement set includes 97 conventional devices measuring active, reactive and current values and 8 PMUs, measuring voltage and current phasors. The measurement values were built from a power flow solution by adding Gaussian errors to the calculated values. The error variances for conventional measurements are defined as 1 x 10^{-3} p.u. and for phasor measurements as 1 x 10^{-6} p.u., to present a case where two sensory systems display radically different error distributions.

In this paper, it is assumed that all Gaussian errors related to each sensor provide sound measurements. However, it is well known that gross errors in the measurement sets may interfere in the fusion process and will be the subject of further studies to assess bad data effects in the fusion method.

SE is performed using both methods (WLS criterion and MCC). It is a theoretical result that both methods will give the same solution, if no gross errors or outliers are present in the data set. In all tested cases of this nature, this has been confirmed – therefore, the distinction of which method is used to calculate which result is a secondary matter.

A. Pareto front and the Ideal Solution Point

The Pareto front was obtained by successively solving the SE problem, varying α from 0 to 1, in steps of 0.001. Fig. 2 and Fig. 3 show the Pareto front in the plane (J_C, J_P) for both WLS-based and MCC-based estimators. This latter figure would be the same as the former, if a representation of the minimization of the CIM metric had been used, instead of the maximization of correntropy, because the data set was prepared without gross errors. The figures confirm that, in the absence of gross errors, the Pareto front is smooth and convex.

The location of the ideal solution point may be also seen in both figures. In this case, its coordinates correspond:

a) In the J_C axis, to the best estimate with $\alpha = 1$

b) In the J_P axis, the best estimate with $\alpha = 0.001$ (not 0, because the system is not observable only with PMU measurements).

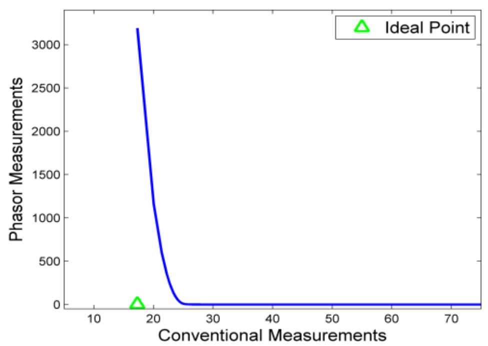

Fig. 2. Pareto front in the plane (J_C, J_P), for the WLS-based estimator

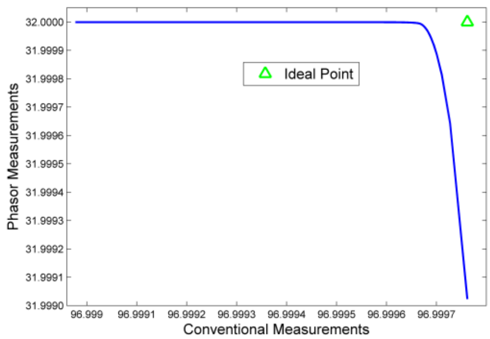

Fig. 3. Pareto front in the plane (J_C, J_P), for the MCC-based estimator

B. Fusion points for distinct metrics

Based on Section V, the location of a compromise solution over the Pareto front, closest to the Ideal, was attempted for L_1, L_2 and L_∞ metrics – the set [L_1, L_∞] on the Pareto front is called "the compromise set" by Zeleny [15]. Also, the evaluation of the solutions on the Pareto front was made using the M_V performance criterion. The functions J_C and J_P evaluate solutions based on the similarity between the perceived image from the reality and the built inner map – i.e. between the measurement vector and the estimated results (usually called residuals and many of them related to power values). The M_V criterion (that can be computed only in experimental studies) evaluates solutions based on the similarity between the "real world" voltage values and the inner map constructed by the SE procedure.

Fig. 4 provides a plot of the value of the M_V criterion over the range $\alpha \in [0,1]$, i.e., over all the Pareto front. Remember that $\alpha = 0$ means that only phasor measurements are considered, while $\alpha = 1$ means that only conventional measurements are taken in account. The most similar estimated voltage vector, in relation to the calculated voltages (assumed "exact" values) is obtained at a value of $\alpha = 0.3650$ – this means that a stronger than average contribution of phasor measurements is being used to build the estimated image of reality.

Table I includes the best compromise solution, in terms of the mix of sensor systems defined by the α credibility coefficient, according to the metric or criterion used.

TABLE I. BEST FUSION SOLUTION, ACCORDING TO DISTINCT CRITERIA

Metric/criterion	Credibility coefficient α
L_∞	0.945
L_2	0.768
L_1	0.5
M_V	0.365

Fig. 5 shows the location, over the Pareto front, of the three compromise solutions and the M_V solution, or four alternative fusion outcomes. Recall that the fusion obtained at $\alpha = 0.5$ is the same as the one given by the Bar-Shalom-Campo fusion.

It becomes clear that the criterion M_V is proposing a fusion point that is outside the compromise set and on the side of a "fractional metric" ($n < 1$). Mind that fractional norms do not induce proper metrics, because the triangular inequality is not observed – however, it does not seem relevant in this case. What is obvious is that, depending on the L (or other) norm used to estimate the minimum distance of the Pareto front to the Ideal, the proposed fusion point will change.

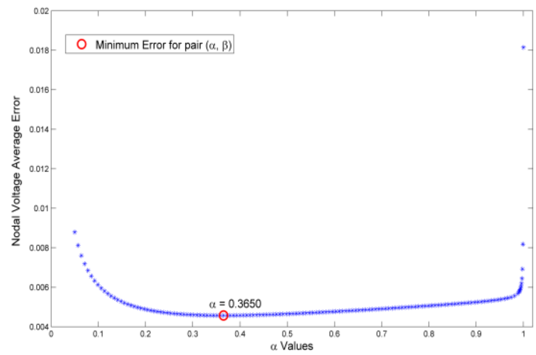

Fig. 4. Plot of the value of Criterion M_V for the whole range of $\alpha \in [0,1]$. The best solution under this criterion is marked on the graph with $\alpha = 0.3650$.

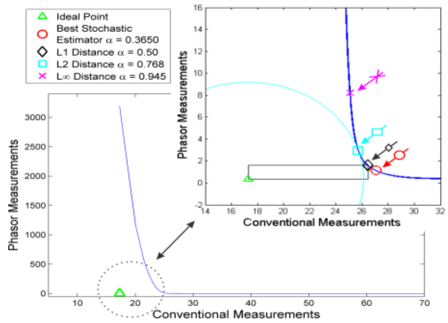

Fig. 5. Location over the Pareto front of the fusion or compromise solutions proposed by the different criteria

The compromise set lies between the solution for L_1 and for L_∞. In this case, it is suggesting that more relevance should be given to the conventional measurements, if one wishes to have an image of the reality with an evaluation as close as possible to the best that could be provided by any of the systems (conventional or PMUs). This may be reasonable, because only 8 PMUs are being used and the image of the system state will be foggy for nodes other than the ones being monitored. However, in the sense of the M_V index, something else is being suggested. The relation between the characteristics of the measurement sets and the Pareto front deserve further investigation.

VIII. CONCLUSIONS

Fusion is a process of combining information from different sensor systems, so that a consistent inner map of an outer reality, perceived with added noise, may be built. Sensor systems may display very diverse characteristics – a trait found in robotics and navigation, where images provided by acoustic and optic sensors must be combined to build a map of the landscape. The growing existence of conventional analog devices and PMUs in a power system provides inspiration to treat the information independently collected as a process requiring information fusion. This can be traditionally dealt with the Bar-Shalom-Campo fusion formula, which has recently been restated in terms of more familiar power system State Estimation quantities.

This paper studies some consequences of adopting a different point of view to achieve sensor signal fusion: a multiple criteria framework where one seeks to maximize the similarity of the perception (measurements) and inner map (estimation of state variables, usually voltages) for both sensor systems simultaneously. This problem displays a domain in the attribute space of the criteria related to the residuals, which exhibits a Pareto optimal front of non-dominated solutions: to improve the similarity of the estimated state to the measurement set of one sensor system, one must accept a reduction in the similarity of the same estimated state to the other sensor system. To each system, a credibility coefficient is associated, reflecting factors like the trust or confidence in such measurement process or historical analysis on which system is providing more accurate estimations. This confidence, therefore, is a factor independent of statistical modeling of error distributions in each device. The variation of these coefficients allows exploring the Pareto front.

The new fusion concept becomes defined as follows: find, in the Pareto front, the fusion solution that is closest to the Ideal (best image from each system). The paper shows that the conventional SE procedure, as well as the information fusion in the Bar-Shalom-Campo sense, are just expressions of having equal confidence in both sensor systems – and that such fusion point corresponds to the minimum distance to the Ideal when a L_1 metric is used.

Other metrics were explored. Of particular interest has been the observation that the best result, evaluated under a measure of similarity between the estimated state and the "exact" or "real" values, diverges from the conventional solution and solutions proposed by other L metrics – in fact, suggesting a fractional metric. The reasons for this behavior will be analyzed in a future work.

REFERENCES

[1] M. Zhou, V. A. Centeno, J. S. Thorp, and A. G. Phadke, "An alternative for including phasor measurements in state estimators," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1930–1937, Nov. 2006.

[2] Vide, P. Castro; Barbosa, F. Maciel; Carvalho, J. Beleza. "Performance metrics for evaluation of a mixed measurement based state estimator." In: *ELEKTRO*, 2012. IEEE, 2012. p. 274-279.

[3] J. Zhou, Y. Zhu, Z. You, and E. Song, "An efficient algorithm for optimal linear estimation fusion in distributed multisensor systems," *IEEE Trans. Syst., Man, Cybern. A*, vol. 36, no. 5, pp.1000–1009, 2006.

[4] A. S. Costa, A. Albuquerque and D. Bez. "An estimation fusion method for including phasor measurements into power system real-time modeling." *IEEE Trans. Power Systems*, *28*(2), 1910-1920, 2013.

[5] H. Mitchell, "Multi-Sensor Data Fusion: An Introduction." New York: Springer, 2007.

[6] V. Miranda, A. Santos, and J. Pereira, "State Estimation Based on Correntropy: A Proof of Concept," *IEEE Trans. Power Syst*., vol.24, no.4, pp.1888,1889, Nov. 2009.

[7] F. C. Schweppe and J. Wildes. "Power System Static-State Estimation, Part I: Exact Model". IEEE Trans. on *Power Apparatus and Systems*, 3(4):120–135, Jan. 1970.

[8] A. Monticelli. "State Estimation in Electric Power Systems: A Generalized Approach". Kluwer Academic Publishers, 1999.

[9] Ali Abur and Antonio Gmez Expsito. Power system state estimation: theory and implementation. Marcel Dekker, 2004.

[10] P. Swerling, "Modern state estimation methods from the viewpoint of the method of least squares," *IEEE Trans. Autom. Control*, vol. 16, no. 6, pp. 707–719, Dec. 1971.

[11] Jakov Krstulovic Opara, "Information Theoretic State Estimation in Power Systems", PhD thesis, Faculty of Engineering, University of Porto, April 2014

[12] E. Parzen, "On the estimation of a probability density function and the mode", Annals Math. Statistics, v. 33, 1962, p. 1065.

[13] W. Wu, Y. Guo. B. Zhang, A. Bose and S. Hongbin, "Robust state estimation method based on maximum exponential square", *IET Generation, Transmission & Distribution,* Vol. 5, no. 11, pp. 1165–1172, Nov. 2011

[14] Y. Bar-Shalom, and L. Campo, "The Effect of the Common Process Noise on the Two-Sensor Fused Track Covariance," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-22, pp. 803 - 805, Dec. 1986

[15] M. Zeleny, J. L. Cochrane, "Multiple Criteria Decision Making", McGraw-Hill, 1982

[16] KEMA, "Metrics for Determining the Impact of Phasor Measurements on Power System State Estimation", Eastern Interconnection Phasor Project, Mar. 2006.