Estimate the Probability Distribution of Imbalance-Induced Energy Loss with Minimal Data

Han Wu, *Student Member*, *IEEE*, Yue Yuan, and Kang Ma, *Member*, *IEEE*, Lurui Fang, *Student Member*

***Abstract*—Imbalance-induced energy loss accounts for a significant part of energy loss in low-voltage distribution networks. However, due to the prohibitive cost of full monitoring, high-resolution time-series data of low-voltage feeder currents are usually absent for accurate energy loss computation. This letter presents a novel method to estimate the probability distribution of imbalance-induced energy loss. In the proposed method, only the line resistance and statistical data (mean and covariance) of the phase currents are required to estimate imbalanced networks' energy loss. Furthermore, a Bayesian statistics-based data-driven method is also proposed to get the mean and covariance value of phase current with minimal measurements. Numerical tests on a real three-phase imbalanced circuit show the proposed approach obtains the probability distribution of imbalance-induced energy loss with a low estimating error and low field measurement frequency.**

***Index Terms*—Energy losses, distribution system, phase residual current, probability.**

I. Introduction

IMBALANCE-induced energy loss accounts for a significant part of energy loss in low-voltage distribution networks [1, 2]. Considering the rapid development of intermittent load and distributed power generations (e.g., electric vehicles and roof-top photovoltaic), the future distribution networks will see an explosive growth of imbalance-induced energy loss. Nevertheless, due to the prohibitive cost of equipping automatic monitoring (such as Micro-PMU) on all low-voltage feeders, most of the phase current data are measured manually by engineers in the field. Without a proper manner that reveals imbalance-induced energy loss with minimal data requirement, distribution network companies may be kept in the dark from actual energy loss.

To this end, this letter extends our work in [3] to three-phase imbalance networks and proposes an analytical formula for the moment generating function (MGF) of imbalance-induced energy loss with minimum information. Then, a moment matching method is developed to estimate the probability distribution. To make sure the proposed method is robust under a small sample set, we also give a Bayesian method to estimate the statistical parameter of phase currents. Compared with the work in [4], this paper aims to give an exact MGF formula for each distribution network with minimal data, which gives a better estimation of imbalance-induced energy loss.

This work was supported in part by the National Key Research and Development Program under Grant 2016YFB0900100.

H. Wu, and Y. Yuan are with the College of Energy and Electrical Engi neering, Hohai University, Nanjing 211100, China (e-mail: wuhanichina@vip.qq.com; yyuan@hhu.edu.cn).

K. Ma and L. Fang are with Department of Electronic and Electrical Engineering, University of Bath, Bath BA2 7AY, United Kingdom (e-mail: km670@bath.ac.uk; L.fang@bath.ac.uk).

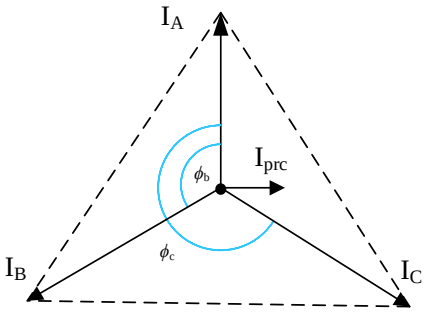


Fig. 1. A schematic diagram of the three-phase imbalanced current.

II. Methodology

A. Expectation for the Imbalance-Induced Energy Loss

The relation between phase current and phase residual current is shown in Fig. 1, where the phase residual current is the vector sum of the phase currents [4]:

$$\vec{I}_{prc} = \vec{I}_a + \vec{I}_b + \vec{I}_c \tag{1}$$

where $\vec{I}_{prc}$ denotes the phase residual current; $\vec{I}_a$, $\vec{I}_b$ and $\vec{I}_c$ denote the phase currents.

Take phase A as reference, equation (1) can be expended as,

$$\begin{aligned}\vec{I}_{prc} &= \vec{I}_a + \vec{I}_b + \vec{I}_c = I_a\angle 0 + I_b\angle\phi_b + I_c\angle\phi_c \\ &= I_a + I_b\cos\phi_b + jI_b\sin\phi_b + I_c\cos\phi_c + jI_c\sin\phi_c \\ &= I_a + I_b\cos\phi_b + I_c\cos\phi_c + j(I_b\sin\phi_b + I_c\sin\phi_c)\end{aligned} \tag{2}$$

where I_a, I_b and I_c are the magnitudes of phase currents; ϕ_b and ϕ_c are the phase angle of phase B and phase C, respectively.

For simplicity's sake, assume the resistance and reactance of three-phase are the same, that is, the phase angles of three-phases are the same. We have ϕ_b=120°, ϕ_c=240°. Then equation (2) can be rewritten as,

$$\vec{I}_{prc} = \left(I_a - \frac{1}{2}I_a - \frac{1}{2}I_c\right) + j\left(\frac{\sqrt{3}}{2}I_c - \frac{\sqrt{3}}{2}I_b\right) \tag{3}$$

$$\left|\vec{I}_{prc}\right| = \text{sqrt}\left[\left(I_a - \frac{1}{2}I_b - \frac{1}{2}I_c\right)^2 + \left(\frac{\sqrt{3}}{2}I_c - \frac{\sqrt{3}}{2}I_b\right)^2\right] \quad (4)$$
$$= \text{sqrt}\left(I_a^2 + I_b^2 + I_c^2 - I_aI_b - I_bI_c - I_aI_c\right)$$

Note, if the phase angles are not the same, we can also get a similar equation through mathematical derivation.

Equation (4) suggests the phase residual current is a square root of a combination of product terms. Since we only focus on the energy loss of the phase residual current, the symbol of square root is omitted, as shown in equation (5).

$$I_{prc}^2 = I_a^2 + I_b^2 + I_c^2 - I_aI_b - I_bI_c - I_aI_c \quad (5)$$

The expectation of the phase residual current is easy to achieve by the definition of variance as,

$$E(I_{prc}^2) = E(I_a^2 + I_b^2 + I_c^2) - E(I_aI_b) - E(I_bI_c) - E(I_aI_c)$$
$$= \sigma_a^2 + \sigma_b^2 + \sigma_c^2 - \text{cov}_{ab} - \text{cov}_{bc} - \text{cov}_{ac} \quad (6)$$
$$+(\mu_a)^2 + (\mu_b)^2 + (\mu_c)^2 - \mu_a\mu_b - \mu_b\mu_c - \mu_a\mu_c$$

where μ_a, μ_b, μ_c are the mean value of phase currents; σ_a^2, σ_b^2, σ_c^2 denote the variance of each phase; cov_{ab}, cov_{bc}, cov_{ac} denote the covariance value, respectively.

Then, the estimated imbalance-induced energy loss is computed by equation (6),

$$E(W_{prc}) = \frac{rT}{\Delta t} \cdot E(I_{prc}^2) \quad (7)$$

where r denotes the line resistance, T denotes the total measurement time, Δt denotes the time interval.

Through equation (6) and (7), one can quickly and accurately derive the expected imbalance-induced energy loss. However, only the knowledge of the expectation is usually not enough to estimate the energy loss and its economic risks. Some further information, like the probability distribution of the phase residual current is required to make rational decisions.

B. Moment Generating Function of the Imbalance-Induced Energy Loss

For simplicity's sake, we rewrite equation (5) to a compact matrix form as,

$$I_{prc}^2 = \boldsymbol{I_\varphi}^T \boldsymbol{A} \boldsymbol{I_\varphi} \quad (8)$$

where $\boldsymbol{A}$ is a positive semidefinite matrix, $\boldsymbol{I_\varphi} = [I_a, I_b, I_c]^T$.

And the corresponding energy loss in equation (7) is, therefore, rewritten as,

$$W_{prc} = \frac{rT}{\Delta t} \cdot I_{prc}^2 = \frac{rT}{\Delta t} \cdot \boldsymbol{I_\varphi}^T \boldsymbol{A} \boldsymbol{I_\varphi} = \boldsymbol{I_\varphi}^T \boldsymbol{B} \boldsymbol{I_\varphi} \quad (9)$$

Equation (8) and (9) indicate the distribution of phase residual current and the corresponding energy loss are linear combinations of correlated quadratic terms. Although it is difficult to get the exact probability density for these quadratic terms, a closed-form formula for the MGF of equation (9) can still be derived under a Gaussian assumption.

For simplicity's sake, assume the three-phase currents have a 3-dimensional Gaussian distribution with mean $\boldsymbol{\mu} = (\mu_a, \mu_b, \mu_c)^T$ and covariance matrix $\boldsymbol{\Sigma}$. Let $\boldsymbol{Z} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{I_\varphi} - \boldsymbol{\mu})$, which converts the phase current $\boldsymbol{I_\varphi}$ to a zero expectation and identity variance matrix $\boldsymbol{Z}$. Thus, the imbalance-induced energy loss W_{prc} is reformulated as,

$$W_{prc} = \boldsymbol{I_\varphi}^T \boldsymbol{B} \boldsymbol{I_\varphi} = (\boldsymbol{Z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1/2} \boldsymbol{B} \boldsymbol{\Sigma}^{-1/2} (\boldsymbol{Z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}) \quad (10)$$

Then, we force the eigen decomposition on the middle term of equation (10), which is,

$$\boldsymbol{\Sigma}^{-1/2} \boldsymbol{B} \boldsymbol{\Sigma}^{-1/2} = \boldsymbol{P}^T \boldsymbol{\Lambda} \boldsymbol{P} \quad (11)$$

where $\boldsymbol{\Lambda}$ is a diagonal for eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of $\boldsymbol{\Sigma}^{-1/2} \boldsymbol{B} \boldsymbol{\Sigma}^{-1/2}$; $\boldsymbol{P}$ is a 3×3 orthogonal matrix ($\boldsymbol{P}\boldsymbol{P}^T = \boldsymbol{P}^T\boldsymbol{P} = \boldsymbol{E}$, $\boldsymbol{E}$ is an identity matrix). Then the energy loss is represented as,

$$W_{prc} = \boldsymbol{I_\varphi}^T \boldsymbol{B} \boldsymbol{I_\varphi} = (\boldsymbol{Z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1/2} \boldsymbol{B} \boldsymbol{\Sigma}^{-1/2} (\boldsymbol{Z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})$$
$$= (\boldsymbol{Z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})^T \boldsymbol{P}^T \boldsymbol{\Lambda} (\boldsymbol{Z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}) \boldsymbol{P} \quad (12)$$
$$= (\boldsymbol{P}\boldsymbol{Z} + \boldsymbol{P}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})^T \boldsymbol{\Lambda} (\boldsymbol{P}\boldsymbol{Z} + \boldsymbol{P}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})$$

Define new variables $\boldsymbol{U} = \boldsymbol{P}\boldsymbol{Z}$ and $\boldsymbol{b} = \boldsymbol{P}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$ to simplify the notations, where $\boldsymbol{U}$ is actually a standard normal distributed variable with zero expectation and identity variance matrix, $\boldsymbol{b}$ is a constant array. Finally, we get the imbalance-induced energy loss as follows,

$$W_{prc} = \boldsymbol{I_\varphi}^T \boldsymbol{B} \boldsymbol{I_\varphi} = (\boldsymbol{U} + \boldsymbol{b})^T \boldsymbol{\Lambda} (\boldsymbol{U} + \boldsymbol{b})$$
$$= \sum_{j=1}^{3} \lambda_j (U_j + b_j)^2 \quad (13)$$

Equation (13) indicates the imbalance-induced energy loss is a sum of weighted noncentral chi-square variables, which has a generalized noncentral chi-square distribution. However, exact formulas for its probabilistic density function (PDF) and cumulative distribution function (CDF) are very complicated [5] apart from certain simple cases (the PDF is $e^{-x\Delta t/3rT}/3$ when phase currents are independent standard normal distributions), while its MGF is much simpler. Thus, we turn to its moment and apply the moment matching method to estimate the shape of the PDF. The MGF is given as [6],

$$M_{W_{prc}}(t) = \exp\left(t\sum_{j=1}^{3}\frac{b_j^2\lambda_j}{1-2t\lambda_j}\right)\prod_{j=1}^{3}\left(1-2t\lambda_j\right)^{-1/2} \quad (14)$$

Although equation (14) is derivated from the Gaussian assumption, the proposed method can easily extend to non-normal distribution through the Gaussian mixture model (GMM), which can be applied on arbitrarily line current distributions.

According to equation (14), the expectation and variance of the generalized noncentral chi-square distribution can be written as,

$$\mu_{W_{prc}} = E[W_{prc}] = tr(\boldsymbol{B}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^T \boldsymbol{B} \boldsymbol{\mu} \quad (15)$$

$$\sigma_{W_{prc}}^2 = Var[W_{prc}] = E[W_{prc}]^2$$
$$= 2tr(\boldsymbol{B}\boldsymbol{\Sigma})^2 + 4\boldsymbol{\mu}^T \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}\boldsymbol{\mu} \quad (16)$$

Note, equation (15) is the same as equation (7). For simplicity's sake, we only test equation (15).

C. Data-Driven Phase Current Distribution Estimation Based on Bayesian Inference

High-resolution phase current data are usually absent in the distribution system due to the low penetration of smart meter.

Valuable phase current data are usually obtained by field measurements, which brings great difficulty to the estimation of energy loss. To address the problem with small sample sizes, the Bayesian inference is used here to maximize the utility of filed measurements.

Compared with the Frequency inference, Bayesian inference is more robust in the small data circumstance. It starts from a prior and iteratively updates its parameter using new measurements. For the energy loss estimation problem at hand, the conjugate prior for the parameter of multivariate Gaussian distribution is a normal-inverse-Wishart (NIW) distribution [7]:

$$(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \sim \mathrm{NIW}(\mu_0, k_0, \lambda_0, \nu_0) \tag{17}$$

The NIW distribution is actually a product distribution of inverse Wishart (IW) distribution and normal distribution as,

$$\boldsymbol{\mu}|\boldsymbol{\Sigma} \sim \mathrm{N}(\mu_0, \boldsymbol{\Sigma} / k_0) \tag{18}$$

$$\boldsymbol{\Sigma} \sim \mathrm{IW}(\lambda_0, \nu_0) \tag{19}$$

where μ_0 is the prior mean for $\boldsymbol{\mu}$; k_0 is a scalar presents how strongly we believe this prior μ_0; λ_0 is the prior mean for the covariance matrix $\boldsymbol{\Sigma}$; ν_0 denotes the degree of freedom, which also represents how strongly we believe the prior λ_0.

Since NIW is the conjugate prior for multinormal distribution, its parameters: μ_0, k_0, λ_0, and ν_0 can be upgraded through the following equations (20)-(23), when new data from field measurement are available [8].

$$\mu_0{}^* = \frac{\kappa_0 \mu_0 + n\bar{x}}{\kappa_0 + n} \tag{20}$$

$$k_0{}^* = k_0 + n \tag{21}$$

$$\nu_0{}^* = \nu_0 + n \tag{22}$$

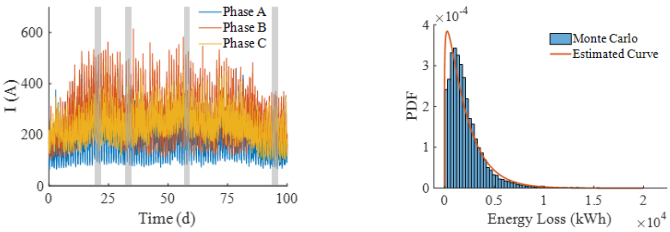
$$\lambda_0{}^* = \lambda_0 + \sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T + \frac{k_0 n}{k_0 + n} (\bar{x} - \mu_0)(\bar{x} - \mu_0)^T \tag{23}$$

where x_i denotes the new measurements; $\bar{x}$ denotes the statistical mean of the new measurements; n denotes the number of the new measurements; $\mu_0{}^*$, $\nu_0{}^*$, $k_0{}^*$ and $\lambda_0{}^*$ are the upgraded parameters of *NIW* distribution.

In practice, a weakly informative data-dependent prior is often employed to start up the Bayesian inference. We follow the recommendation in [8] and set $\mu_0 = \bar{x}$, $k_0 = 0.01$, $\lambda_0 = \mathrm{diag}(\mathrm{cov}(\bar{x}))/n$, and $\nu_0 = 5$. Then, the posteriors of the mean and covariance for the phase current under the Bayesian framework is obtained by iteratively taking field measurements into equation (20)-(23).

III. Case Study

A three-phase imbalanced circuit from Western Power Distribution (a UK distribution network operator) is employed to demonstrate the proposed energy loss estimation method. In the three-phase imbalanced circuit, phase B burdens the "heaviest" load and phase A has the "lightest" load. The neutral line resistance is set to 0.1 Ω. The data set given in Fig. 2 (a) contains a measurement of line current for 100 days with a time interval of 10 min (144 data points per day for each phase). To mimic the real phase current measurement scene, we randomly choose four sample periods, each period contains 10 successive sample points (see grey bars in Fig.2 (a)).



(a) (b)

Fig. 2. (a) The testing phase current and samples. (b) The estimated probability distribution and the Monte Carlo simulation result.

TABLE I
Statistics of the Phase Currents

Phase	Mean (A)			Standard Variance (A^2)		
	Bayesian Inference	Frequency Inference	True	Bayesian Inference	Frequency Inference	True
A	195.46	225.40	187.05	46.73	70.29	74.16
B	277.82	301.86	262.92	53.86	68.79	87.63
C	206.20	222.12	229.56	56.29	62.90	54.64

TABLE II
Computing Results of Imbalance-induced Energy Loss

Cases	Energy Loss (kWh)	Estimated Error (%)
True Energy Loss	2085.85	-
Equation (15) with True Parameter	2085.92	3.36×10^{-3}
Frequency Inference	2292.21	9.89
Bayesian Inference	2133.91	2.30

The case study is organized as follows. Firstly, the performance of Bayesian statistics is demonstrated. Then the energy loss expectation from equation (7), equation (15) and the true energy loss from all measurements are compared. Finally, the probability density through moment matching is compared with the Monte Carlo method. The whole case study is performed on a laptop with Intel Core i7 8650U (1.90 GHz) and 16 GB RAM, while the program is implemented using Mathematica 12.1. The original data and code are available in [9] for readers to verify this method.

The statistics of phase current through Bayesian inference and Frequency inference are given in Table I. The estimated imbalance-induced energy loss is calculated by equation (15). As a comparison, we also calculate the practical imbalance-induced energy loss based on the line current curves. Computing results and related estimating error (abs(A-B)/A×100, where A denotes practical results, B denotes proposed method results) are shown in Table II.

As shown in Table II, the proposed analytical energy loss formula calculates the exact energy loss with negligible estimated error. In the most data-scare distribution networks, the proposed energy loss formula with Bayesian inference method can get a 2.3% estimated error with only four measurements. The estimation result will be more accurate if more measurements are taken.

Fig. 2(b) compares the PDF matching by a Gamma distribution and histogram from Monte Carlo simulation. The results indicate that the proposed method successfully compute the probability distribution for the square of phase residual

current, and thus gives a good fit of energy loss.

IV. Conclusion

This letter presents an explicit formula for the expectation value and the MGF function of imbalance-induced energy loss in distribution network. To address the problem of low smart meter penetration, we also proposed a Bayesian framework to maximize the utility of filed measurements. In the case study, proposed method obtains an exact energy loss result with only four times of measurement. Further study can focus on different load distributions and the impact of dynamic line resistance.

References

[1] J. D. Watson, N. R. Watson, and I. Lestas, "Optimized Dispatch of Energy Storage Systems in Unbalanced Distribution Networks," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 2, pp. 639-650, 2018, doi: 10.1109/TSTE.2017.2752964.

[2] Goran Strbac *et al.*, "Management of Electricity Distribution Network Losses," 2014. [Online]. Available: https://www.westernpower.co.uk/downloads/4847.

[3] H. Wu, Y. Yuan, and K. Ma, "A Novel Probabilistic Method for Energy Loss Estimation Using Minimal Line Current Information," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4928-4931, 2020.

[4] L. Fang, K. Ma, R. Li, Z. Wang, and H. Shi, "A Statistical Approach to Estimate Imbalance-Induced Energy Losses for Data-Scarce Low Voltage Networks," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2825-2835, 2019.

[5] T. Cacoullos and M. Koutras, "Quadratic forms in spherical random variables: Generalized noncentral x2 distribution," *Naval Research Logistics Quarterly*, vol. 31, no. 3, pp. 447-461, 1984.

[6] A. Mathai and S. Provost, "Quadratic Forms in Random Variables: Theory and Applications," vol. 87, 1992.

[7] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*. CRC press, 2013.

[8] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012.

[9] "Data and code for three phase imbalance network." 2021. [Online] Available: https://github.com/wuhanichina/Imbalance-3-Phase.