RESEARCH

Electrical transient modeling for appliance characterization

Mohamed Nait-Meziane[1*] , Philippe Ravier[1], Karim Abed-Meraim[1], Guy Lamarque[1], Jean-Charles Le Bunetel[2] and Yves Raingeaud[2]

Abstract

Transient signals are characteristic of the underlying phenomenon generating them, which makes their analysis useful in many fields. Transients occur as a sudden change between two steady state regimes, subsist for a short period, and tend to decay over time. Hence, superimposed damped sinusoids (SDS) were extensively used for transients modeling as they are adequate for describing decaying phenomena. However, SDS are not adapted for modeling the turn-on transient current of electrical appliances as it tends to decay to a steady state that is different from the one preceding it. In this paper, we propose a new and more suitable model for these signals for the purpose of characterizing appliances. We also propose an algorithm for the model parameter estimation and validate its performance on simulated and real data. Moreover, we give an example on the use of the model parameters as features for the classification of appliances using the Controlled On/Off Loads Library (COOLL) dataset. The results show that the proposed algorithm is efficient and that for real data the network fundamental frequency must be estimated to account for its variations around the nominal value. Finally, real data experiments showed that the model parameters used as features yielded a classification accuracy of 98%.

Keywords: Electrical appliances characterization, Harmonic signal, Parameter estimation, Transient modeling

1 Introduction

Studying transient phenomena is important and useful in many fields such as biomedical research for the analysis of heart rate variability [1], the extraction of detailed information of muscle behavior [2], and the detection and classification of epileptic spikes [3]; mechanics for the study of the susceptibility of structures to vibration issues [4]; and for seismic events detection and temporal localization [5, 6]. Monitoring electrical loads and systems is particularly one of the areas where transients play a central role. We cite as applications the analysis of disturbances affecting the quality of the electric power system [7, 8], fault detection in rotary machines [9, 10], and non-intrusive load monitoring (NILM) [11–13], a field concerned with extracting individual energy consumption (e.g., of different appliances) from measured total energy consumption (e.g., at main breaker panel).

Transients embed a decay or damping characteristic as they exist for short periods, and therefore, the superimposed damped sinusoids (SDS) model [14] was extensively used to model transients in many fields. For example, it was used for modeling electric disturbances [15], transient audio signals [16], and the free induction decay observed in nuclear resonance spectroscopy [17]. Along with the model, different algorithms were proposed for its parameter estimation [18]. The most known methods are Prony's [19], Pisarenko's [20], matrix pencil [21], Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [22], and MUltiple SIgnal Classification (MUSIC) [23]. Despite its success, the SDS model is inadequate for turn-on transient current signals. In fact, a lot of turn-on transient current signals are characterized by a quasi-stationary harmonic content (Fig. 1a) whereas the SDS model is best suited for modeling vanishing non-stationary content (Fig. 1b) because having different damping factors for each frequency produces a signal with non-stationary frequency content. Moreover, the turn-on transient current decays to a steady state that is different from the steady state preceding the turn-on

*Correspondence: mohamed.nait-meziane@univ-orleans.fr
[1]PRISME Laboratory, University of Orléans, 12 rue de Blois, 45067 Orléans, France
Full list of author information is available at the end of the article
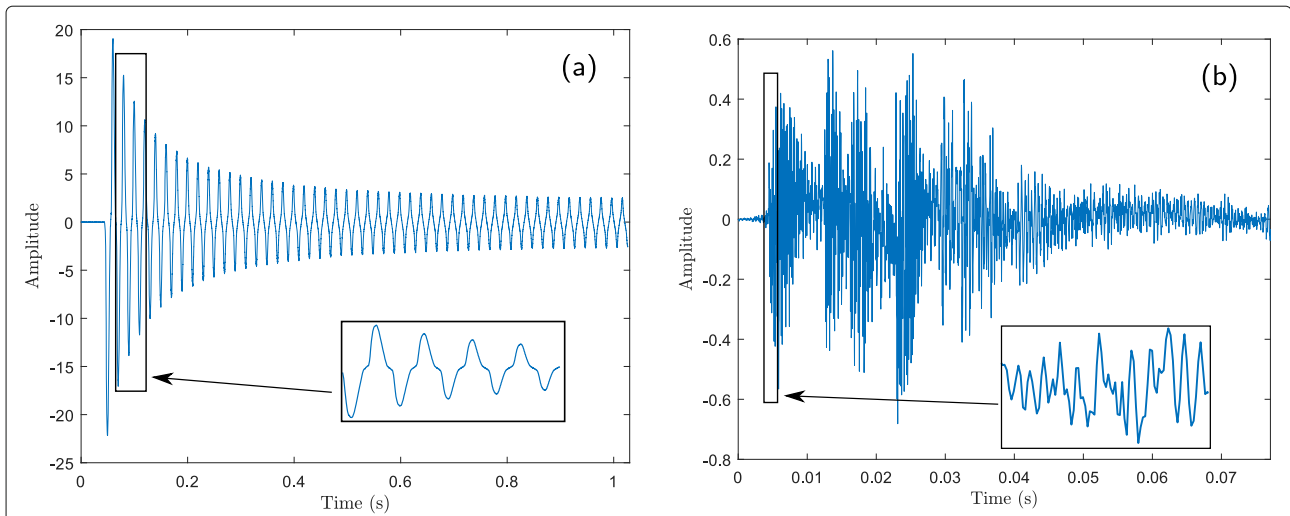
Fig. 1 Two examples of transients. **a** A (turn-on) transient of an electrical signal (drill) for which the superimposed damped sinusoids (SDS) model is inadequate, and **b** a transient of an audio signal (castanets) for which the SDS model is adequate. Note the non-stationary content of **b** and the quasi-stationary content of **a**

of the appliance, whereas in the SDS the transient model starts from one steady state and decays to the same one afterwards. The electrical current "turn-on" transient is the current that appears with the switching-on of an electrical appliance. This corresponds to a transition from one steady state to another. For example, if we consider a single appliance on the network, then the first steady state is the state of no consumption, the second steady state is the state of steady consumption, and the transient is the part in between. Note that the transient we are interested in modeling is the one related to the electrical consumption. This transient is different from the very high frequency transient (appearing as a short pulse preceding the turn-on transient) generated by the switching devices because of the closure of the circuit [24]. Turn-on transients are appliance-dependent and last usually from few power system cycles to few seconds. A turn-on transient is typically characterized by a high current amplitude (surge) at the beginning of consumption followed by a decrease (or damping) in the amplitude of the consumed energy until reaching a stable state (Fig. 1a).

In this paper, we propose a new model for turn-on transient current signals along with an efficient algorithm for the parameter estimation. The parameters are used to characterize electrical appliances and are shown to be useful for appliance classification. Several objectives are targeted in this paper including:

- Derivation of an efficient estimation algorithm for the model parameters.
- Assessment of the estimation performance via the computation of the Cramér-Rao bound (CRB).

- Validation of the proposed model on real transient signals and the evaluation of the modeling error when using the developed estimation algorithm.
- Exploitation of the model parameters for appliance characterization and assessment of their usefulness as relevant features for a classification task example.
- As a by-product, we also developed a full experimental setup (with a dedicated transient signal acquisition device) to build our dataset of real transient signals corresponding to different electrical appliances. This dataset is used for our model validation as well as for the performance assessment of the proposed appliance identification method.

The rest of the paper is organized as follows: Section 2.1 describes the proposed data model, Section 2.2 details the proposed parameter estimation algorithm, Section 3 gives the derivation of the parameters' CRBs, Section 4.1 gives the assessment of the proposed model and algorithm on simulated and real data, Section 4.2 shows the usefulness of the model parameters through an appliance classification example, and finally, Section 5 concludes the paper.

2 Methods
2.1 Data modeling
In this section, we propose and discuss a mathematical model for turn-on transient current signals. Strictly speaking, we model the turn-on transient including a small part of the following steady state regime; mainly because the transient end is not well defined and because estimating the harmonic content is easier on the steady state part.

The shape of the turn-on transient and the related amplitudes vary from one electrical appliance to another. To take into account these variations, we propose to model the noiseless electrical current turn-on transient $s(t)$ as the product of two signals

$$s(t) = e(t)s_s(t), \quad \forall t \in [\,0, +\infty) \tag{1}$$

where $e(t)$ represents an amplitude modulation (envelope) and $s_s(t)$ is a sum of d sinusoids given as follows

$$s_s(t) = \sum_{i=1}^{d} a_i \cos\left(2\pi f_i t + \phi_i\right), \quad \forall t \in [\,0, +\infty) \tag{2}$$

where $a_i\ (\geq 0), \phi_i \in [-\pi, \pi]$ and f_i are the sinusoids amplitudes, phases, and frequencies, respectively. The number d of sinusoids (current harmonics) is assumed known (typically $d = 5$ harmonics is enough to represent the sinusoidal signal $s_s(t)$) and the frequencies satisfy $f_i = (2i - 1)f_0, i = 1, \ldots, d$ where $f_0 \approx 50$ Hz. Indeed, because of the half-wave symmetry found in electrical signals (i.e., for a periodic signal $g(t)$ of period P, a half-wave symmetry is characterized by $g(t + P/2) = -g(t)$), the sinusoid frequencies f_i are odd-order harmonics of the fundamental frequency. Note that the nominal value (50 Hz) of the network fundamental frequency is a priori known, but due to its fluctuations around this value over time (i.e., $f_0 = 50 + \delta f$), we have observed that for a correct modeling, f_0 should be considered as unknown and hence one needs to re-estimate the fundamental frequency value for each transient signal[1].

The envelope $e(t)$ is chosen of the form $e^{u(t)} + 1$ such that $e^{u(t)} \xrightarrow[t \to +\infty]{} 0$. This exponential function was chosen for its usefulness in describing damped phenomena. A classical damped model is such that $u(t) = -\alpha t$ with $\alpha > 0$. For our model, we propose to extend this classical model in order to adapt it to real current signals. Specifically, we propose to model $u(t)$ as a polynomial function allowing more flexibility in describing the amplitude modulation of real signals

$$e(t) = e^{\mathbf{p}^T \mathbf{t}} + 1, \quad \forall t \in [\,0, +\infty) \tag{3}$$

where $\mathbf{p} = \left[p_0, p_1, \ldots, p_n\right]^T$ is a vector of $n+1$ polynomial coefficients and $\mathbf{t} = [1, t, \ldots, t^n]^T$ is a time vector such that $\mathbf{p}^T \mathbf{t}$ is a polynomial of degree n allowing the model adaptation to the real signal variations[2]. The polynomial is such that $\mathbf{p}^T \mathbf{t} \xrightarrow[t \to +\infty]{} -\infty$ leading to $e(t) \xrightarrow[t \to +\infty]{} 1$ (verified if $p_n < 0$).

---

[1]The European norm "EN 50160" [25] fixes the acceptable variation ranges for δf_0. For the synchronous grid of Continental Europe—the largest synchronous (same frequency) grid in the world linking most of Europe's countries and some countries of north Africa—these ranges are $\pm 1\%$ of f_0
($\delta f_0 = [-0.5, +0.5]$ Hz) 99.5% of a year and $-6\%/+4\%$ of f_0
($\delta f_0 = [-3, +2]$ Hz) 100% of the time. The latter range is made large to account for occasional high variations.
[2]Based on our real data measurements, we have observed that a polynomial order $n = 3$ is sufficient to model properly the considered transient signals.

We assume that the measured current signal $x(t)$ is corrupted by an additive white Gaussian noise (AWGN) $w(t)$ with zero mean and variance σ^2

$$x(t) = s(t) + w(t), \quad \forall t \in \mathbb{R}. \tag{4}$$

The passage from continuous to discrete-time notation is done using $t_k = kT_s$, where $T_s = 1/F_s$ is the sampling period and $k \in \mathbb{Z}$. This notation is used in the remainder of the paper.

2.2 Parameter estimation algorithm

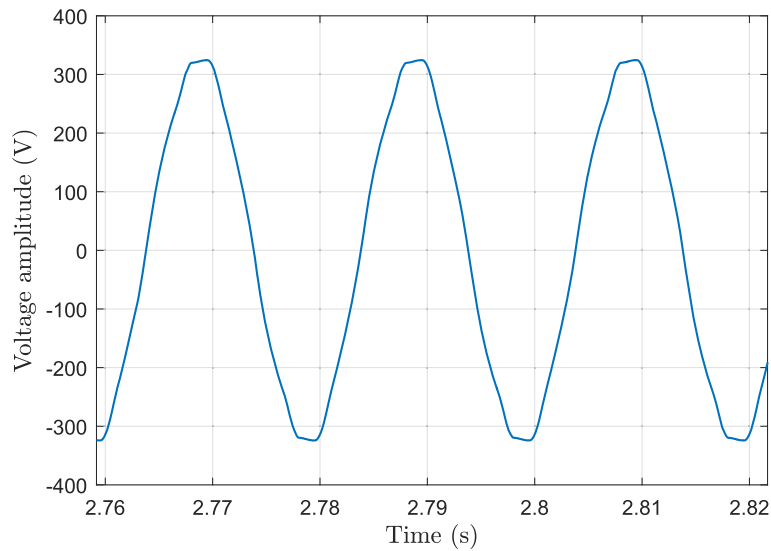The proposed parameter estimation algorithm proceeds in two steps:

- Estimation of the fundamental frequency f_0.
- Estimation of the other signal parameters using the a priori estimated frequency $\hat{f}_0$.

2.2.1 Fundamental frequency estimation

We assume that the fundamental frequency is unknown but quasi-constant over the transient duration, typically less than 5 s, and we propose to estimate it using the voltage signal, which is almost perfectly sinusoidal (Fig. 2). Indeed, the stability of f_0 (i.e., its rate of change) is an important issue that was discussed in depth in the literature. This can be seen from the plot in Fig. 3 (borrowed from http://wwwhome.cs.utwente.nl/~ptdeboer/misc/mains.html) which represents the "Allan deviation" of f_0 for a measurement over a period of 69 days. As explained in this reference, if the Allan deviation at an averaging duration of 10 s is 10^{-4}, it means that if one measures the frequency during 10 s and once more during the next 10 s, these measurements will differ on average by 0.01%. Based on this, we consider the frequency variation over our 5-s measurement period as negligible. Hence, the fundamental frequency estimation problem turns into the classical problem of estimating the frequency of a monotone signal in noise. It is known that the CRB of the frequency parameter of a monotone signal decreases with a rate of $1/N^3$ [26]

$$\text{var}(\hat{f}_0) \geq \frac{12}{(2\pi)^2 \eta N(N^2 - 1)}, \tag{5}$$

where η is the signal-to-noise ratio (SNR) and N the number of signal samples. This allows for highly precise frequency estimates. Practically, we use the algorithm proposed by Aboutanios and Mulgrew [27] which is shown to provide a precise frequency estimate reaching the CRB. Indeed, the voltage signal is modeled here as a pure sinusoid of frequency f_0 corrupted by an AWGN. In such a case, the optimal maximum likelihood (ML) solution coincides with the peak location estimation of the Fourier transform of the signal. This estimation is achieved by the low cost efficient numerical method in [27].

Fig. 2 A real voltage signal

In the second step (next section), we estimate the remaining parameters using the frequency $\hat{f}_0$.

2.2.2 Transient current estimation (TCE) algorithm

Given the estimated frequency $\hat{f}_0$, the second step of our estimation algorithm operates in two phases: (i) initialization and (ii) parameter estimation.

The initialization phase provides initial estimates of the parameters p_j, a_i, and ϕ_i ($j = 0, \ldots, n$ and $i = 1, \ldots, d$) to be used in the parameter estimation phase, during which these estimates will be refined. This two-phase

structure of the algorithm is motivated by the difficulty and high computational cost of the nonlinear maximum-likelihood-based estimation criterion (see (15)). In such a case, we usually seek a good initial estimate and then refine it in order to alleviate the ill-convergence and high computational cost of the problem.

Note that the algorithm needs some pre-specified values for n, d, and f_i and also needs a pre-defined steady state portion used in the initialization step for the estimation of the amplitudes a_i and phases ϕ_i of the sinusoids. Hereafter, we start by discussing these
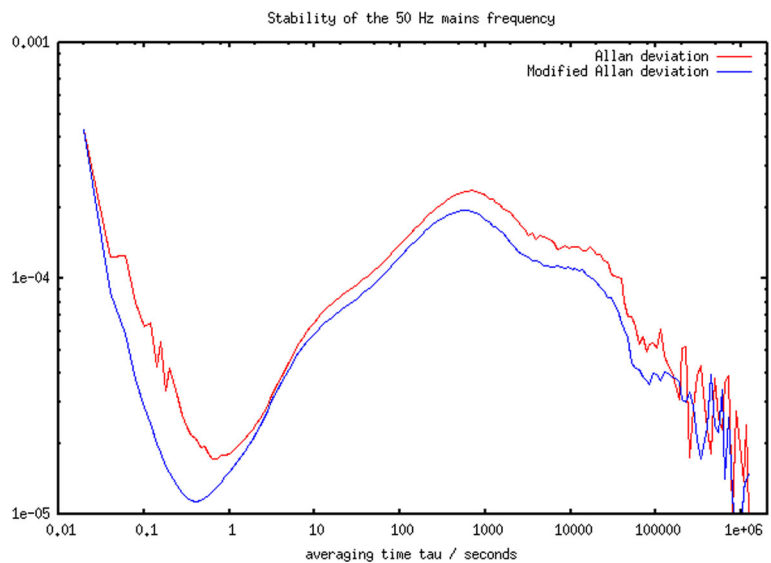
Fig. 3 Stability of network frequency (log-log plot of Allan deviation). Source: http://wwwhome.cs.utwente.nl/~ptdeboer/misc/mains.html

"pre-specified quantities," then we proceed to detailing the algorithm.

Pre-specified quantities These quantities are f_i ($i = 1, \ldots, d$), n, d, t_{ss1}, and t_{ss2} ("ss" stands for steady state). As mentioned in Section 2.1, f_i are odd-order harmonics of f_0. Taking into account the estimated fundamental frequency $\hat{f}_0$, the sinusoids' frequencies are given by $f_i = (2i - 1)\hat{f}_0$. The rest of the parameters are chosen in an ad hoc way. The polynomial degree n and the number of sinusoids d were chosen based on experimental observations made on the real data we used. The chosen values were $n = 3$ and $d = 5$ (see the discussion of assumption **A3** in Section 4.1.2).

Quantities t_{ss1} and t_{ss2} define the time instants that delimit a portion, noted $x_{ss}(t_k)$, of the steady state of the current signal (Fig. 4). This portion is used in the initialization phase for the estimation of the amplitudes and phases. On this steady state portion, $t_k \in [t_{ss1}, t_{ss2}]$, we can write $x(t_k) = x_{ss}(t_k) = s_s(t_k) + w(t_k)$ where $s_s(t_k)$ is the sum of sinusoids signal (2); we neglect the envelope influence by assuming $e(t) = 1$ on this portion.

We define t_{ss1} and t_{ss2} using the High Accuracy NILM Detector (HAND) algorithm [28] found in the literature. Applying HAND on a turn-on transient signal provides the time-instants t_{beg}^{on} and t_{end}^{on} defining the beginning and end of the turn-on transition (Fig. 4). Practically, we define t_{ss1} a few (typically ten) time cycles after t_{end}^{on} and t_{ss2} such that the duration of $x_{ss}(t_k)$ is 25 time-cycles (i.e., half a second, sufficient to get good initial estimates of amplitudes and phases).

Initialization phase

1. Estimation of $s_s(t_k)$: Using the least squares (LS) criterion, parameters a_i and ϕ_i are estimated using
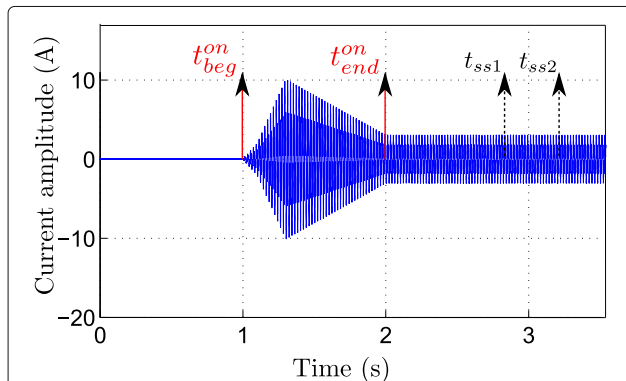
Fig. 4 Output of the High Accuracy NILM Detector (HAND) when applied to a simulated single-appliance signal. HAND outputs the time-instants: t_{beg}^{on} and t_{end}^{on}. Blue: current signal. t_{ss1} and t_{ss2} define the steady state portion (25 time-cycles long) used for the estimation

($t_k \in [t_{ss1}, t_{ss2}]$)

$$x_{ss}(t_k) = s_s(t_k) + w(t_k)$$

$$= \sum_{i=1}^{d} a_i \cos\left(2\pi f_i t_k + \phi_i\right) + w(t_k)$$

$$= \sum_{i=1}^{d} [\, a_i \cos \phi_i \cos\left(2\pi f_i t_k\right)$$

$$- a_i \sin \phi_i \sin\left(2\pi f_i t_k\right)\,] + w(t_k). \quad (6)$$

Writing (6) in vector form gives

$$\mathbf{x_{ss}} = \mathbf{Mc} + \mathbf{w}, \quad (7)$$

where $\mathbf{x_{ss}} = [x_{ss}(t_{ss1}), \ldots, x_{ss}(t_{ss2})]^T$, $\mathbf{c} = [a_1 \cos \phi_1, a_1 \sin \phi_1, \ldots, a_d \cos \phi_d, a_d \sin \phi_d]^T$, $\mathbf{w} = [w(t_{ss1}), \ldots, w(t_{ss2})]^T$ is the noise vector and $\mathbf{M}$ s the matrix given in (8).

$$\mathbf{M} = \begin{bmatrix} \cos\left(2\pi f_1 t_{ss1}\right) & -\sin\left(2\pi f_1 t_{ss1}\right) & \cdots & \cos\left(2\pi f_d t_{ss1}\right) & -\sin\left(2\pi f_d t_{ss1}\right) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \cos\left(2\pi f_1 t_{ss2}\right) & -\sin\left(2\pi f_1 t_{ss2}\right) & \cdots & \cos\left(2\pi f_d t_{ss2}\right) & -\sin\left(2\pi f_d t_{ss2}\right) \end{bmatrix}$$

$$(8)$$

The LS criterion, used to find an estimate for $\mathbf{c}$, aims to minimize the square of the Euclidean norm ($\| \cdot \|_2$) of the difference between the measured signal and the data model, i.e.,

$$\hat{\mathbf{c}} = \arg \min_{\mathbf{c}} \frac{1}{2} \|\mathbf{x_{ss}} - \mathbf{Mc}\|_2^2$$

$$\text{subject to} \quad a_i \geq 0 \text{ and } \phi_i \in [-\pi, \pi], \forall i. \quad (9)$$

In that case, the solution $\hat{\mathbf{c}}$ is given by

$$r\hat{\mathbf{C}}l\hat{\mathbf{c}} = \mathbf{M}^+ \mathbf{x_{ss}}, \quad (10)$$

where $\mathbf{M}^+ = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T$ is the (Moore-Penrose) pseudo-inverse of $\mathbf{M}$. We extract from $\hat{\mathbf{c}}$ two vectors $\hat{\mathbf{cs}} = \left[\hat{a}_1 \cos \hat{\phi}_1, \ldots, \hat{a}_d \cos \hat{\phi}_d\right]^T$ and $\hat{\mathbf{sn}} = \left[\hat{a}_1 \sin \hat{\phi}_1, \ldots, \hat{a}_d \sin \hat{\phi}_d\right]^T$, and we compute $\hat{a}_i$ and $\hat{\phi}_i$ as follows

$$\hat{\mathbf{a}} = \begin{bmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_d \end{bmatrix} = \sqrt{\hat{\mathbf{cs}} \odot \hat{\mathbf{cs}} + \hat{\mathbf{sn}} \odot \hat{\mathbf{sn}}}, \quad (11)$$

$$\hat{\boldsymbol{\phi}} = \begin{bmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_d \end{bmatrix} = \arctan\left(\hat{\mathbf{sn}} \oslash \hat{\mathbf{cs}}\right), \quad (12)$$

where $\odot$ and $\oslash$ are the element-wise product and division operators, respectively.

2. Estimation of $e(t_k)$:
   To estimate $e(t_k)$, we use the trust-region-reflective (TRR) algorithm [29, 30], an efficient nonlinear optimization algorithm that belongs to the "trust

region" class of algorithms [31]. This algorithm allows constraints to be imposed on the values of the parameter estimates enabling us to satisfy our constraints ($a_i \geq 0$, $\phi_i \in [-\pi, \pi]$ and $p_n < 0$). Having the estimates $\hat{\mathbf{a}}$ and $\hat{\boldsymbol{\phi}}$ obtained from the previous step and the (overall) measured signal $\mathbf{x} = [x(t_0), \ldots, x(t_{N-1})]^T$, we estimate an initial value of $\mathbf{p} = [p_0, \ldots, p_n]^T$ (the remaining unknown) using

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} \frac{1}{2}||\mathbf{x} - (\mathbf{e} \odot \mathbf{M}_{ov}\hat{\mathbf{c}})||_2^2$$

$$\text{subject to} \quad p_n < 0, \tag{13}$$

where $\mathbf{e} = [e(t_0), \ldots, e(t_{N-1})]^T$, N is the number of samples of $\mathbf{x}$, and $\mathbf{M}_{ov}$ is the matrix $\mathbf{M}$ (8) with $t \in [t_0, \ldots, t_{N-1}]$ instead of $t \in [t_{ss1}, \ldots, t_{ss2}]$. Moreover, t_0 is practically chosen as the time-instant corresponding to the maximum current amplitude; that way, we model the damped part of the turn-on transient starting from the maximum amplitude. In the end of this phase, we obtain the initial estimated parameter vector $\hat{\boldsymbol{\theta}}_0 = \left[\hat{\mathbf{p}}^T, \hat{\mathbf{a}}^T, \hat{\boldsymbol{\phi}}^T\right]^T$.

Parameter estimation phase As the estimation of a_i and ϕ_i is done using only a portion of the total measured signal $x(t_k)$, the aim of this parameter estimation phase is to improve the estimation of all the parameters by considering all the samples of $x(t_k)$. We use for this the same TRR algorithm considered for the estimation of $e(t_k)$ by taking as initial value the result of the estimation phase $\hat{\boldsymbol{\theta}}_0$. The unknown, to be estimated this time, is the global parameter vector $\boldsymbol{\theta} = \left[\mathbf{p}^T, \mathbf{a}^T, \boldsymbol{\phi}^T\right]^T$ estimated as

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2} \|\mathbf{x} - (\mathbf{e} \odot \mathbf{M}_{ov}\mathbf{c})\|_2^2 \tag{14}$$

The TCE algorithm is summarized in Algorithm 1.

3 Cramér-Rao bounds of the model parameters

The Cramér-Rao bound (CRB) provides a lower bound on the variance of any unbiased estimator. We show in Section 4.1.1 that this unbiasedness condition is approximately verified (at least for moderate and high SNRs) for our estimated parameters, and hence, we can use the CRB to assess the performance of our estimation.

Evaluating the performance of the estimation consists of comparing the estimated parameters' variances with their CRB. Taking into account the dependence of $s(t_k)$ on the parameter vector $\boldsymbol{\theta}$ and N samples of $x(t_k)$, (4) can be written using vector notation as

$$\boldsymbol{x} = \boldsymbol{s}(\boldsymbol{\theta}) + \boldsymbol{w}, \tag{15}$$

where $\boldsymbol{x}$ is normally distributed with mean $\boldsymbol{\mu}(\boldsymbol{\theta}) = \boldsymbol{s}(\boldsymbol{\theta}) = [s(t_0, \boldsymbol{\theta}), \ldots, s(t_{N-1}, \boldsymbol{\theta})]^T$ and a covariance matrix $\boldsymbol{C}(\boldsymbol{\theta}) = \boldsymbol{C} = \sigma^2 \boldsymbol{I}$.

Algorithm 1: Transient Current Estimation (TCE) algorithm

Pre-specified quantities: $f_i \ (i = 1, \ldots, d), n, d, t_{ss1}, t_{ss2}$.
Initialization

   1. Compute $\hat{\mathbf{a}}$ and $\hat{\boldsymbol{\phi}}$ such that

$$\hat{\mathbf{a}} = \sqrt{\hat{\mathbf{cs}} \odot \hat{\mathbf{cs}} + \hat{\mathbf{sn}} \odot \hat{\mathbf{sn}}} \quad (11)$$

$$\hat{\boldsymbol{\phi}} = \arctan\left(\hat{\mathbf{sn}} \oslash \hat{\mathbf{cs}}\right) \quad (12)$$

   using the least squares (LS) criterion where
   $\hat{\mathbf{cs}} = \left[\hat{a}_1 \cos\hat{\phi}_1, \ldots, \hat{a}_d \cos\hat{\phi}_d\right]^T$ and
   $\hat{\mathbf{sn}} = \left[\hat{a}_1 \sin\hat{\phi}_1, \ldots, \hat{a}_d \sin\hat{\phi}_d\right]^T$ extracted from

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \frac{1}{2} \|\mathbf{x_{ss}} - \mathbf{Mc}\|_2^2 = \mathbf{M}^+\mathbf{x_{ss}} (10)$$

   where $\mathbf{M}$ is computed using (8) and $\mathbf{c} = [a_1 \cos\phi_1, a_1 \sin\phi_1, \ldots, a_d \cos\phi_d, a_d \sin\phi_d]^T$.

   2. Compute

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} \frac{1}{2} \|\mathbf{x} - (\mathbf{e} \odot \mathbf{M}_{ov}\hat{\mathbf{c}})\|_2^2,$$

$$\text{s.t. } p_n < 0 \quad (13)$$

   using the TRR algorithm [29, 30] initialized by
   $\mathbf{p_0} = \mathbf{0}$. $\mathbf{M}_{ov}$ is the matrix $\mathbf{M}$ (8) with
   $t \in [t_0, \ldots, t_{N-1}]$ instead of $t \in [t_{ss1}, \ldots, t_{ss2}]$.

Parameter estimation

   1. Compute $\hat{\boldsymbol{\theta}}$ such that

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2} \|\mathbf{x} - (\mathbf{e} \odot \mathbf{M}_{ov}\mathbf{c})\|_2^2 \quad (14)$$

   using the TRR algorithm [29, 30] with
   $\hat{\boldsymbol{\theta}}_0 = \left[\hat{\mathbf{p}}^T, \hat{\mathbf{a}}^T, \hat{\boldsymbol{\phi}}^T\right]^T$ as initial value.

The CRB is defined as the inverse of the Fisher information matrix (FIM). If we assume $\boldsymbol{\theta} = \left[\mathbf{p}^T, \mathbf{a}^T, \boldsymbol{\phi}^T\right]^T = [\theta_1, \ldots, \theta_K]^T$, $K = n + 1 + 2d$ (the noise power is assumed known here) and $\mathbf{x} = [x(t_0), \ldots, x(t_{N-1})]^T \in \mathbb{R}^N$, then for the general Gaussian case where $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{C}(\boldsymbol{\theta}))$, the FIM is given elementwise by [26]

$$[\mathbf{F}(\boldsymbol{\theta})]_{ij} = \left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i}\right)^T \mathbf{C}^{-1}(\boldsymbol{\theta}) \left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j}\right) + \frac{1}{2}\text{tr}\left[\mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_i} \mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_j}\right], \tag{16}$$

where $\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i} = \left[\frac{\partial[\boldsymbol{\mu}(\boldsymbol{\theta})]_1}{\partial \theta_i}, \ldots, \frac{\partial[\boldsymbol{\mu}(\boldsymbol{\theta})]_N}{\partial \theta_i}\right]^T$, and

$$\frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_i} = \begin{bmatrix} \frac{\partial[\mathbf{C}(\boldsymbol{\theta})]_{11}}{\partial \theta_i} & \cdots & \frac{\partial[\mathbf{C}(\boldsymbol{\theta})]_{1N}}{\partial \theta_i} \\ \vdots & \ddots & \vdots \\ \frac{\partial[\mathbf{C}(\boldsymbol{\theta})]_{N1}}{\partial \theta_i} & \cdots & \frac{\partial[\mathbf{C}(\boldsymbol{\theta})]_{NN}}{\partial \theta_i} \end{bmatrix}.$$

The symbol $[\cdot]_i$ denotes element of index i of the corresponding vector, $[\cdot]_{ij}$ denotes the element of index ij of the corresponding matrix, and tr $[\cdot]$ denotes the trace operator.

For our model where $\boldsymbol{\mu}\left(\boldsymbol{\theta}\right) = \mathbf{s}\left(\boldsymbol{\theta}\right)$ and $\mathbf{C}\left(\boldsymbol{\theta}\right) = \sigma^2\mathbf{I}$ (covariance matrix independent of $\boldsymbol{\theta}$), the (elementwise) FIM becomes

$$[\mathbf{F}\left(\boldsymbol{\theta}\right)]_{ij} = \frac{1}{\sigma^2}\left(\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \theta_i}\right)^T \left(\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \theta_j}\right), \quad i,j = 1,\ldots,K.$$

(17)

Taking into account (17) and the structure of $\boldsymbol{\theta}$, the FIM can be written using matrix notation in the form of a nineblock matrix representing the partial derivatives with respect to the elements of $\boldsymbol{\theta}$ as

$$\mathbf{F}\left(\boldsymbol{\theta}\right) = \frac{1}{\sigma^2}\begin{bmatrix} \left(\frac{\partial \mathbf{s}^T}{\partial p_l}\frac{\partial \mathbf{s}}{\partial p_m}\right) & \left(\frac{\partial \mathbf{s}^T}{\partial p_l}\frac{\partial \mathbf{s}}{\partial a_m}\right) & \left(\frac{\partial \mathbf{s}^T}{\partial p_l}\frac{\partial \mathbf{s}}{\partial \phi_m}\right) \\ \left(\frac{\partial \mathbf{s}^T}{\partial a_l}\frac{\partial \mathbf{s}}{\partial p_m}\right) & \left(\frac{\partial \mathbf{s}^T}{\partial a_l}\frac{\partial \mathbf{s}}{\partial a_m}\right) & \left(\frac{\partial \mathbf{s}^T}{\partial a_l}\frac{\partial \mathbf{s}}{\partial \phi_m}\right) \\ \left(\frac{\partial \mathbf{s}^T}{\partial \phi_l}\frac{\partial \mathbf{s}}{\partial p_m}\right) & \left(\frac{\partial \mathbf{s}^T}{\partial \phi_l}\frac{\partial \mathbf{s}}{\partial a_m}\right) & \left(\frac{\partial \mathbf{s}^T}{\partial \phi_l}\frac{\partial \mathbf{s}}{\partial \phi_m}\right) \end{bmatrix},$$

(18)

where $l, m \in \{1,\ldots,n+1\}$ for the elements of $\boldsymbol{p}$ and $l, m \in \{1,\ldots,d\}$ for the elements of $\boldsymbol{a}$ and $\boldsymbol{\phi}$.

Since this matrix is symmetric (because of the symmetry of the second order partial derivatives), we only have to compute the following terms to find all the elements of the matrix: $\frac{\partial \mathbf{s}^T}{\partial p_l}\frac{\partial \mathbf{s}}{\partial p_m}, \frac{\partial \mathbf{s}^T}{\partial p_l}\frac{\partial \mathbf{s}}{\partial a_m}, \frac{\partial \mathbf{s}^T}{\partial p_l}\frac{\partial \mathbf{s}}{\partial \phi_m}, \frac{\partial \mathbf{s}^T}{\partial a_l}\frac{\partial \mathbf{s}}{\partial a_m}, \frac{\partial \mathbf{s}^T}{\partial a_l}\frac{\partial \mathbf{s}}{\partial \phi_m}, \frac{\partial \mathbf{s}^T}{\partial \phi_l}\frac{\partial \mathbf{s}}{\partial \phi_m}$. After straightforward derivations, we get

$$\frac{\partial \mathbf{s}^T}{\partial p_l}\frac{\partial \mathbf{s}}{\partial p_m} = \sum_{k=0}^{N-1}\left(s_s(t_k)e^{\mathbf{p}^T\mathbf{t}_k}\right)^2 t_k^{l+m}$$

$$\frac{\partial \mathbf{s}^T}{\partial p_l}\frac{\partial \mathbf{s}}{\partial a_m} = \sum_{k=0}^{N-1} s(t_k)e^{\mathbf{p}^T\mathbf{t}_k} t_k^l \cos\left(2\pi f_m t_k + \phi_m\right)$$

$$\frac{\partial \mathbf{s}^T}{\partial p_l}\frac{\partial \mathbf{s}}{\partial \phi_m} = -\sum_{k=0}^{N-1} s(t_k)e^{\mathbf{p}^T\mathbf{t}_k} t_k^l a_m \sin\left(2\pi f_m t_k + \phi_m\right)$$

$$\frac{\partial \mathbf{s}^T}{\partial a_l}\frac{\partial \mathbf{s}}{\partial a_m} = \sum_{k=0}^{N-1} e(t_k)^2 \cos\left(2\pi f_l t_k + \phi_l\right)\cos\left(2\pi f_m t_k + \phi_m\right)$$

(19)

$$\frac{\partial \mathbf{s}^T}{\partial a_l}\frac{\partial \mathbf{s}}{\partial \phi_m} = -\sum_{k=0}^{N-1} e(t_k)^2 a_m \cos\left(2\pi f_l t_k + \phi_l\right)\sin\left(2\pi f_m t_k + \phi_m\right)$$

$$\frac{\partial \mathbf{s}^T}{\partial \phi_l}\frac{\partial \mathbf{s}}{\partial \phi_m} = \sum_{k=0}^{N-1} e(t_k)^2 a_l a_m \sin\left(2\pi f_l t_k + \phi_l\right)\sin\left(2\pi f_m t_k + \phi_m\right),$$

where $\mathbf{t}_k = \left[1, t_k,\ldots,t_k^n\right]^T$, and $s(t_k), s_s(t_k), e(t_k)$ are defined in (1), (2), and (3), respectively. Finally, the CRB is equal to $\mathbf{F}^{-1}\left(\boldsymbol{\theta}\right)$ obtained after inserting expressions (19) in (18) and inverting $\mathbf{F}\left(\boldsymbol{\theta}\right)$.

In the previous CRB derivation, we assumed the noise variance σ^2 known so that $K = n + 1 + 2d$. If we assume that σ^2 is also an unknown parameter to be estimated such that $\boldsymbol{\theta}' = \left[\mathbf{p}^T, \mathbf{a}^T, \boldsymbol{\phi}^T, \sigma^2\right]^T$, then the FIM $\mathbf{F}\left(\boldsymbol{\theta}'\right)$ is

equal to the FIM $\mathbf{F}\left(\boldsymbol{\theta}\right)$ augmented with one row and one column corresponding to partial derivatives with respect to σ^2. Using (16), we get

$$\mathbf{F}\left(\boldsymbol{\theta}'\right) = \begin{bmatrix} \mathbf{F}\left(\boldsymbol{\theta}\right) & \mathbf{0} \\ \mathbf{0} & \frac{N}{2\sigma^4} \end{bmatrix}.$$

(20)

The CRB is then given by

$$\text{CRB}(\boldsymbol{\theta}') = \mathbf{F}^{-1}\left(\boldsymbol{\theta}'\right) = \begin{bmatrix} \mathbf{F}^{-1}\left(\boldsymbol{\theta}\right) & \mathbf{0} \\ \mathbf{0} & \frac{2\sigma^4}{N} \end{bmatrix}.$$

(21)

This indicates that it is sufficient to independently compute $\mathbf{F}^{-1}\left(\boldsymbol{\theta}\right)$ and $\frac{2\sigma^4}{N}$ to find $\mathbf{F}^{-1}\left(\boldsymbol{\theta}'\right)$. It also means that the existence or lack of information about σ^2 does not affect the performance bound (CRB) of the other desired parameters.

4 Results and discussion
4.1 Estimation performance assessment
4.1.1 Assessment on simulated data
In this section, we present the results of the estimation performance evaluation on simulated data. Hereafter, we present (i) the simulated signal and its parameters, (ii) the bias of the estimated parameters, (iii) the estimated parameters variance and its comparison to the CRB, (iv) the CRB variation with respect to the sampling frequency, and (v) the convergence of the TCE algorithm.

Simulated signal and its parameters Taking the considered setup ($n = 3$ and $d = 5$) in this section, we end up with 14 parameters for the simulated signal. So, with such large number of degrees of freedom, we decided to choose the set of parameters such that the simulated signal will resemble as much as possible real signals, and without a priori knowledge on what parameter values are appropriate, we decided to tweak the model parameters and choose the ones that gave a simulated signal "resembling" (similar waveform) typical real current waveforms from our dataset. The noiseless signal model is

$$s(t_k) = e(t_k)s_s(t_k) = \left(e^{\mathbf{p}^T t_k} + 1\right)\sum_{i=1}^{d} a_i\cos\left(2\pi f_i t_k + \phi_i\right),$$

where $\mathbf{p} = \left[p_0, p_1,\ldots,p_n\right]^T$, $\boldsymbol{t}_k = \left[1, t_k,\ldots,t_k^n\right]^T$, $t_k \in \left[t_0, t_{N-1}\right]$, $a_i \ (\geq 0)$, $\phi_i \in \left[-\pi, \pi\right]$ and $f_i = (2i - 1)f_0$, $i = 1,\ldots,d$ with a fixed fundamental frequency $f_0 = 50$ Hz. The chosen model parameter values are
$F_s = 30$ kHz: sampling frequency
$t_0 = 0$ s, $t_{N-1} = 3$ s: specify signal duration
$n = 3$: polynomial degree
$d = 5$: number of harmonics
$\mathbf{p} = [\, 1.9, -9, 8.5, -4\,]^T$
$\mathbf{a} = [\, 1.8, 0.5, 0.2, 0.1, 0.05\,]^T$
$\boldsymbol{\phi} = [\, -3, 3, 2.5, 1.5, 1\,]^T$
$t_{ss1} = 2.5$ s, $t_{ss2} = 3$ s: define steady state portion.

The polynomial degree is chosen relatively small (we found that $n = 3$ is sufficient to characterize the tested signals, and hence, this value will be used in the remainder of the paper). The previous transient signal is then corrupted by an additive white Gaussian noise of zero mean and variance σ^2 (varied such that the signal-to-noise ratio (SNR) $= \frac{1/N \sum_{k=0}^{N-1} s(t_k)^2}{\sigma^2}$ varies in the range $[0 - 50]$ dB). The obtained simulated signal is shown in Fig. 5.

Bias of the estimated parameters As mentioned before, the CRB applies for unbiased estimators. Our maximum likelihood estimation method based on criterion (14) is known to be asymptotically, i.e., for high SNR, unbiased [26]. Here, we evaluate our estimator bias numerically using (1000) Monte-Carlo runs.

Figure 6 gives the bias computed for the parameters estimated using the proposed algorithm versus the signal-to-noise ratio (SNR). We can see that all estimated parameters have negligible bias for SNR values greater than or equal to 30 dB. Between 10 dB and 30 dB, the estimated parameters have very small biases, and below 10 dB, we start getting some bias, nonetheless with values that are still small compared to the true parameter values.

Estimated parameters' variance and its comparison to the CRB Similarly to the bias computation, the variance is also computed numerically. Figure 7 shows the different parameter variances compared with their respective CRBs. We note that all the parameter variances coincide with their respective CRBs almost perfectly. Hence, our estimation is efficient (unbiased and the variance reaches the CRB).

CRB variation with respect to the sampling frequency Due to the transient behavior of the observed phenomena, a good choice for the sampling frequency F_s of measurements is mandatory. We seek a sufficiently high sampling frequency to catch the transient behavior but not too high to avoid heavy computational load. The CRB allows us to evaluate the impact of the sampling frequency on the parameter variance lower bounds and therefore decide on the desired performance taking into account computational complexity.
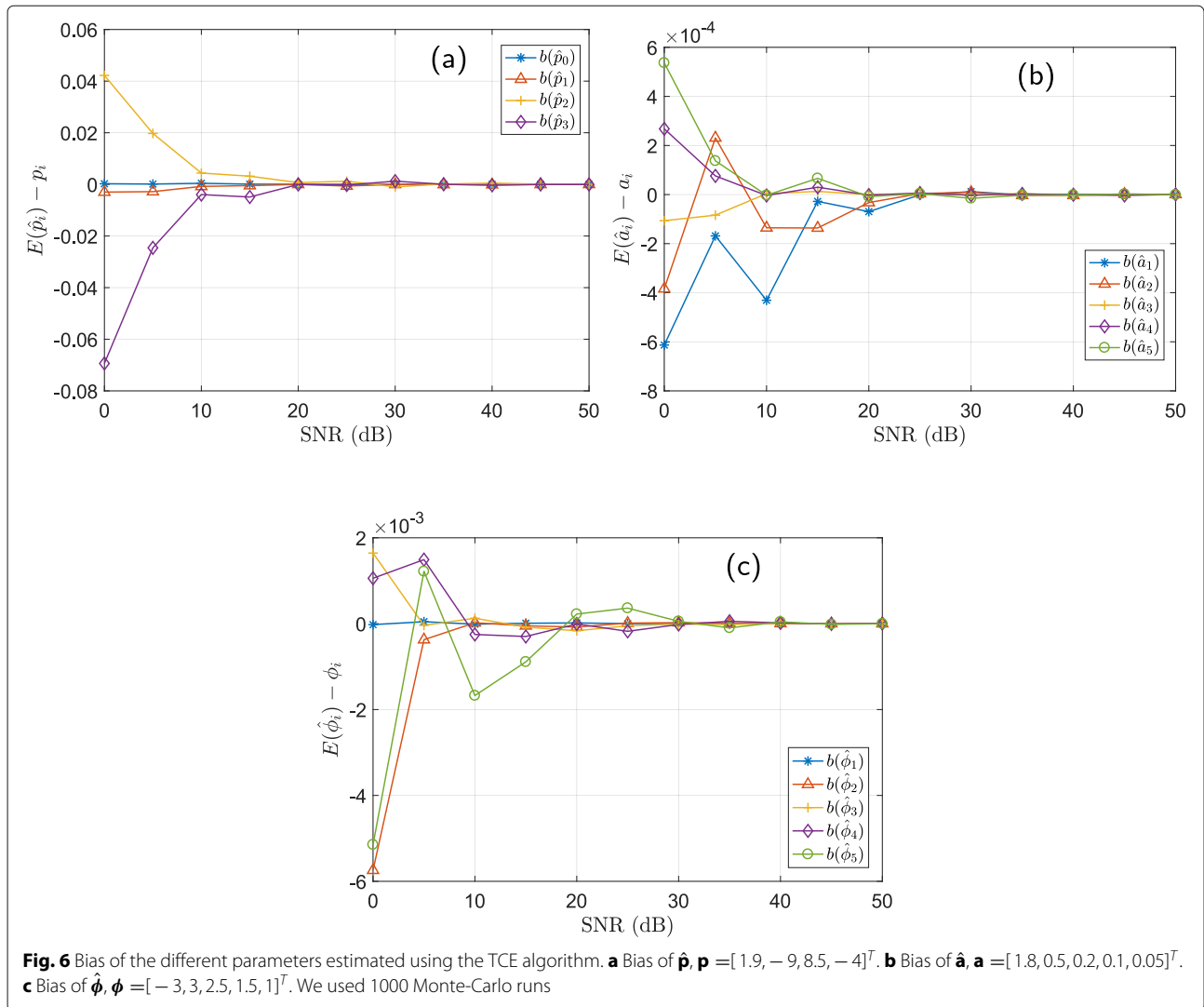
Figure 8 gives the variation of the parameters' CRB as a function of F_s (1 to 100 kHz) on a logarithmic scale. The results show that an increase in F_s results in a better estimation performance (linearly decreasing variance w.r.t. to the sample size) for all the parameters. This is expected, since a higher F_s means more data samples (on a fixed time period) and hence better performance. When considering real signals, however, this is not necessarily true. The white noise (independence) assumption initially verified for relatively low F_s (still high frequency though) might not be verified in practice for higher frequency values. At higher frequencies, the data samples become closer and might become correlated, when the time duration between two samples is too small to assume independence. In that case, the computed CRB assuming a white noise can no longer be used to evaluate the estimation performance.

Practically, finding the adequate sampling frequency is not easy since it depends on different parameters: transient waveforms of interest and their frequency contents, computational complexity, desired performance,
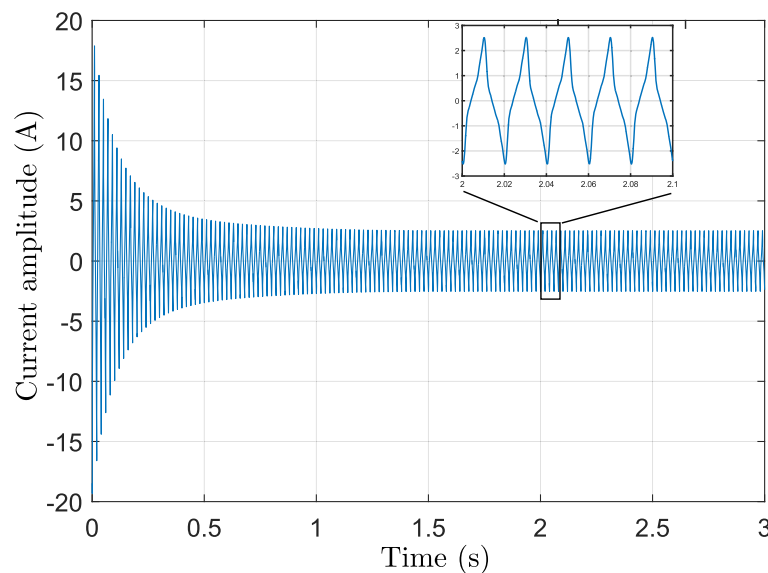
Fig. 5 Simulated turn-on transient current used for performance assessment

Fig. 6 Bias of the different parameters estimated using the TCE algorithm. **a** Bias of $\hat{\mathbf{p}}$, $\mathbf{p} = [1.9, -9, 8.5, -4]^T$. **b** Bias of $\hat{\mathbf{a}}$, $\mathbf{a} = [1.8, 0.5, 0.2, 0.1, 0.05]^T$. **c** Bias of $\hat{\boldsymbol{\phi}}$, $\boldsymbol{\phi} = [-3, 3, 2.5, 1.5, 1]^T$. We used 1000 Monte-Carlo runs

etc. According to our experiments, and for the study of turn-on transient signals, a sampling frequency at least equal to 5 kHz is recommended (captures around 50 harmonics) whereas going beyond 100 kHz starts generating heavy data processing. A sampling frequency of 30 kHz seems to be a good compromise, since it captures around 300 harmonics and is less computationally heavy. Hence, our choice was $F_s = 30$ kHz for simulations.
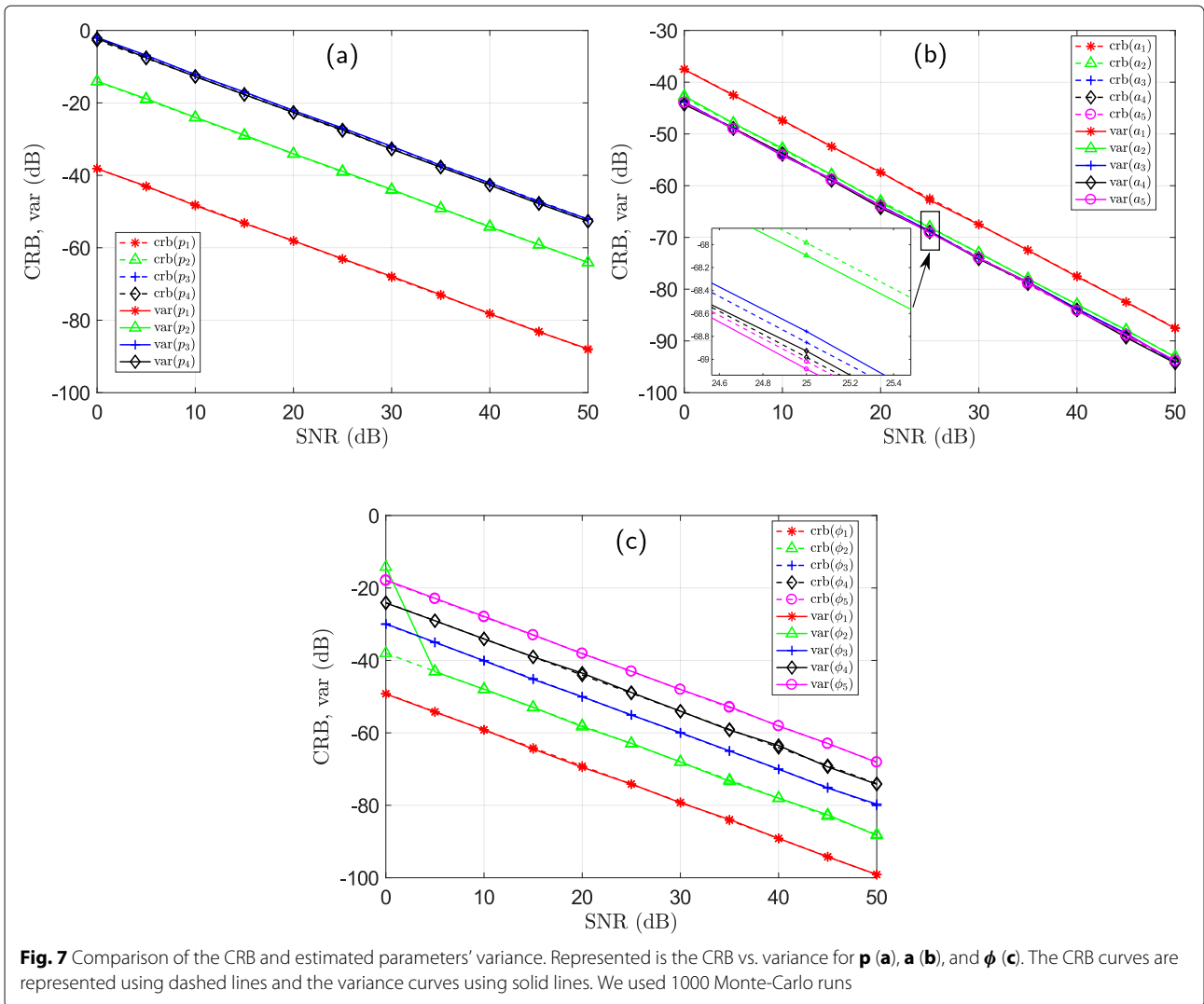
Convergence of the TCE algorithm Since the TCE algorithm uses an optimization algorithm in both its estimation and refinement phases, it is important to check its convergence, especially if there is a need for real-time processing. Hereafter, we check the convergence of the nonlinear optimization algorithm, trust-region-reflective (TRR), for both phases.

Figure 9 gives, for different SNR values, the mean-square-error (MSE) as a function of the number of iterations in the estimation phase of the TCE algorithm.

Independently of the SNR, the algorithm converges after ten iterations. The convergence of the TRR algorithm in the refinement phase is even faster. Indeed, the initialization point being better defined, it converges at most after three iterations.

4.1.2 Assessment on real data
Real data considerations Until now, we have implicitly assumed, for simulated data, some simplifying assumptions to test the parameter estimation independently of the performance of other blocs that may condition the estimation. These assumptions are as follows: (**A1**) a well-defined portion of the steady state, (**A2**) transient starting from a maximum, and (**A3**) known polynomial degree n and number of harmonics d. In real situations, however, these assumptions are not necessarily verified for the following reasons: (i) the definition of the steady state portion is affected by the precision of turn-on transient (end) detection and is never perfect (depends on the detector

Fig. 7 Comparison of the CRB and estimated parameters' variance. Represented is the CRB vs. variance for **p** (**a**), **a** (**b**), and **ϕ** (**c**). The CRB curves are represented using dashed lines and the variance curves using solid lines. We used 1000 Monte-Carlo runs

accuracy); (ii) physically, there will always be a latency in the appliance response before the current signal reaches its maximum amplitude; and (iii) the polynomial degree as well as the harmonic numbers are only chosen parameters used to trade-off between complexity and modeling efficiency.

The easiest assumption to get around in a real situation is **A2** since we only need to detect the signal maximum amplitude and model the damped part, starting from this maximum (so the portion of transient signal preceding the peak value will be disregarded). For assumption **A1**, we use the HAND detector [28] built specifically to allow high accuracy detection of turn-on transients. For **A3**, we relied on our dataset of real life signals to get our ad hoc choice of the "effective" polynomial degree $n = 3$ and "effective" number of harmonics $d = 5$ that have been experimentally shown to be suitable for a good modeling of the considered transient signals. As an example,

Fig. 10 shows different plots comparing the real signal $x(t)$ to its estimate $\hat{x}(t)$ for different values of the polynomial degree n (1, 3, 5 and 7). We note the improvement of the root-mean-square error (RMSE) between $n = 1$ and $n = 3$, hence a better estimation using $n = 3$, and a slight improvement of the RMSE between $n = 3$, $n = 5$, and $n = 7$. We consider $n = 3$ to be a good trade-off between model complexity and the estimation performance (less than 10% of relative RMSE difference, e.g., between $n = 3$ and $n = 5$, we have a relative RMSE difference of $\frac{0.2473 - 0.2433}{0.2473} \approx 1.6\%$).
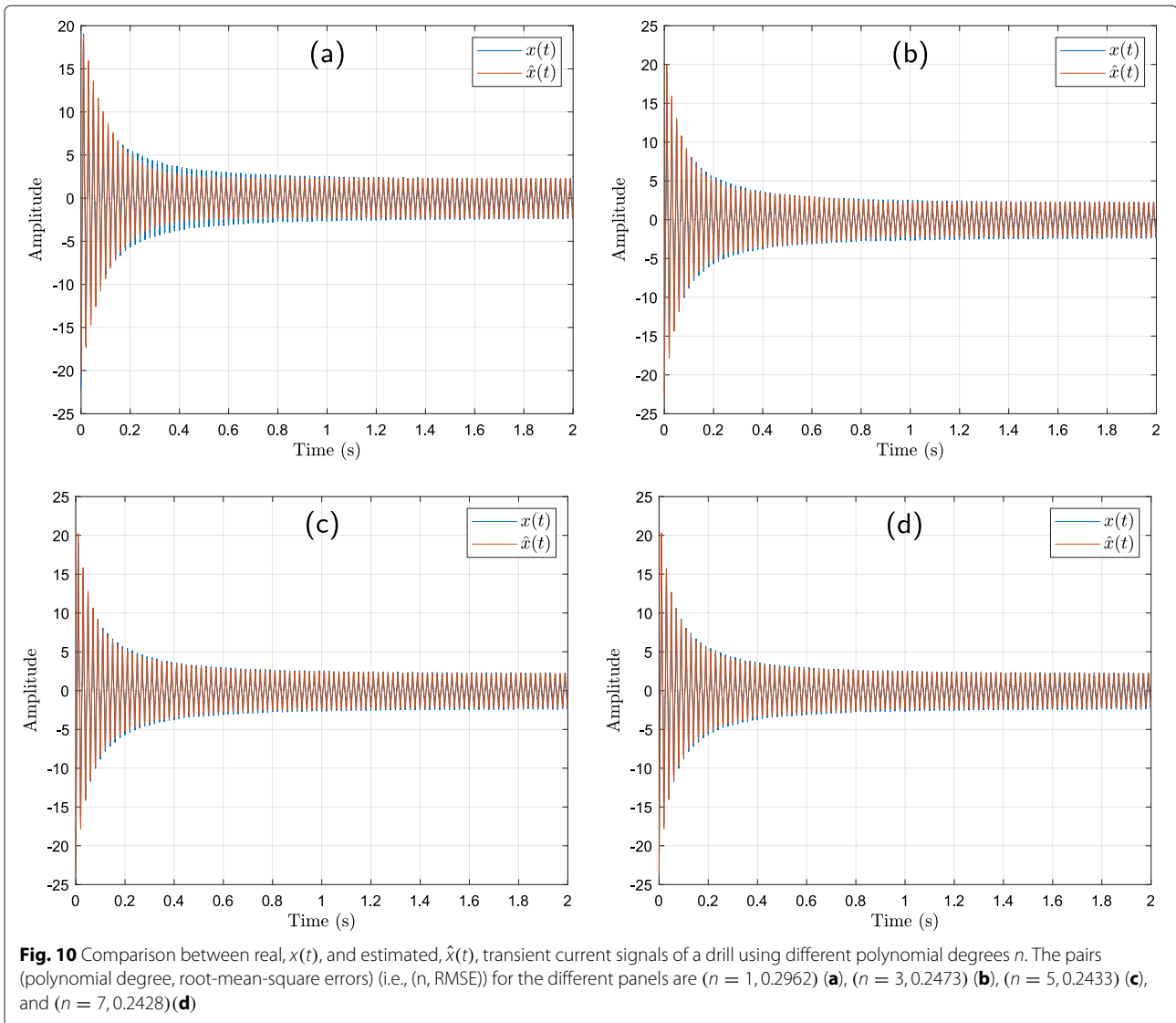
As an aside, note that the particular "two-steps" structure of the proposed parameter estimation algorithm (Section 2.2) is motivated by the highly nonlinear optimization problem involved. Such a two-step approach helps to avoid local minima by providing a good initialization point in the first step, then refining the obtained estimate in the second step. For completeness, we have
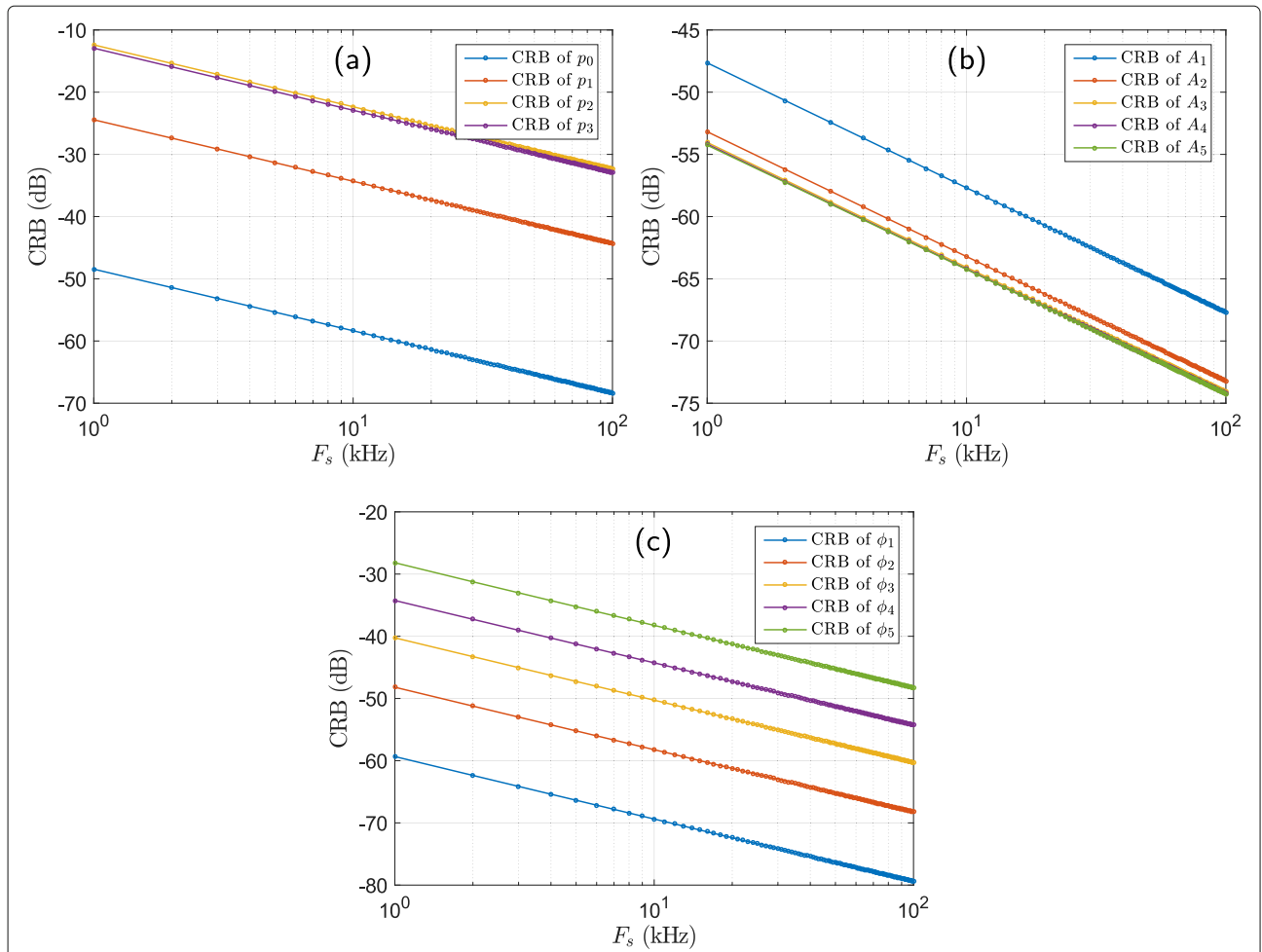
Fig. 8 CRB of the model parameters as a function of the sampling frequency F_s (from 1 kHz to 100 kHz) at an SNR of 25 dB. Represented is the CRB of **p** (**a**), **a** (**b**), and **φ** (**c**)
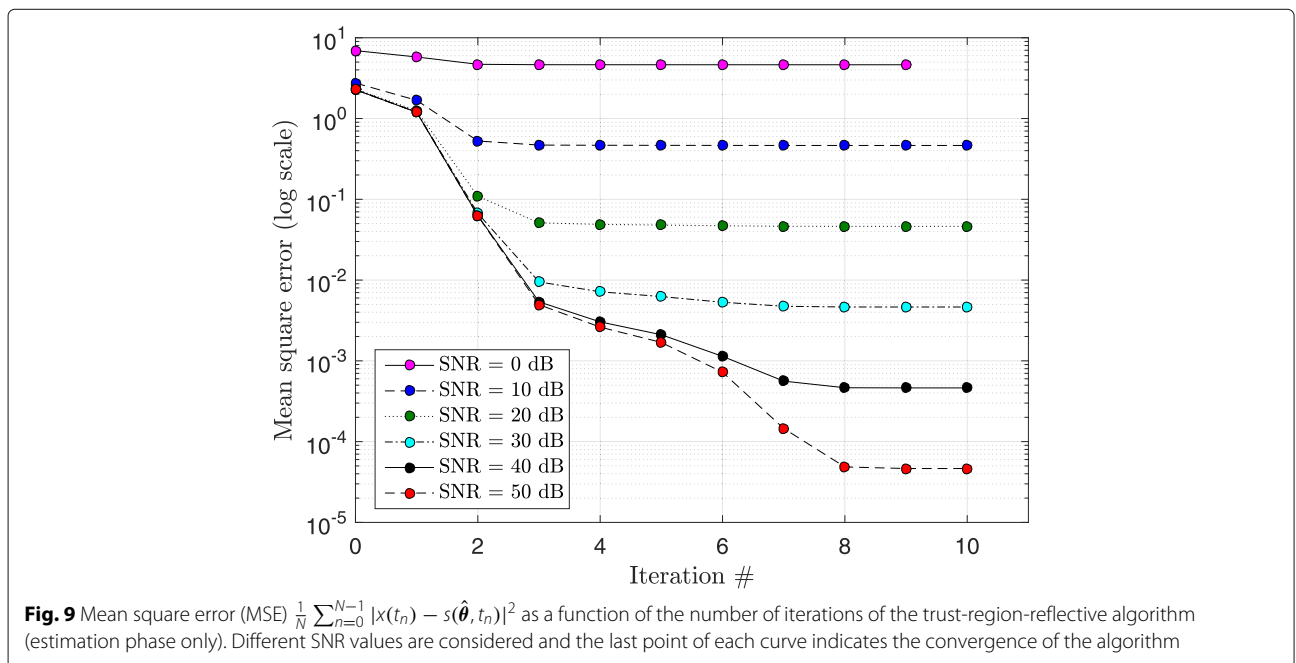
Fig. 9 Mean square error (MSE) $\frac{1}{N} \sum_{n=0}^{N-1} |x(t_n) - s(\hat{\boldsymbol{\theta}}, t_n)|^2$ as a function of the number of iterations of the trust-region-reflective algorithm (estimation phase only). Different SNR values are considered and the last point of each curve indicates the convergence of the algorithm

Fig. 10 Comparison between real, $x(t)$, and estimated, $\hat{x}(t)$, transient current signals of a drill using different polynomial degrees n. The pairs (polynomial degree, root-mean-square errors) (i.e., (n, RMSE)) for the different panels are ($n = 1, 0.2962$) (**a**), ($n = 3, 0.2473$) (**b**), ($n = 5, 0.2433$) (**c**), and ($n = 7, 0.2428$)(**d**)

also tried to improve the estimate of f_0 by jointly estimating it when (i) estimating $\hat{p}$ (13), (ii) when refining the estimation of all parameters (14), and (iii) in both (i) and (ii). This, however, did not improve the results, indicating that the proposed approach already leads to near optimal values (due in part to the highly precise estimate of f_0 obtained using [27]). As an example, we have conducted the joint estimations described above considering the real signal used in Fig. 10 with $n = 3$ that gave initially RMSE = 0.2473. The newly obtained results, in terms of RMSE, were 0.2473, 0.2479, and 0.2479, respectively for (i), (ii), and (iii). The joint estimation of f_0 was not considered further as it would generate more computational load without performance gain.

Estimation with TCE on a real signal of the COOLL dataset The real signal is taken from a turn-on transient dataset we built especially for transients analysis.

The dataset is called Controlled On/Off Loads Library (COOLL) [32] and is freely available on the internet (https://coolldataset.github.io/). Since the measurement system [33] (Fig. 11) used to collect the dataset's signals allows the control over the turn-on/off, we know exactly the turn-on/off time instants and assumptions **A1** and **A2** hold then true. Moreover, we consider the signal starting from its maximum in order to verify **A3**.

The COOLL dataset signals (Table 1) are sampled at $F_s = 100$ kHz[3]. The dataset consists of turn-on transient current and voltage signals of 12 different electrical appliances and each appliance has 20 signal examples. Figure 12 shows a typical histogram of the noise on a measured current signal taken from its pre-turn-on part (noise only).

---

[3]After measurements were done we found that $F_s = 30$ kHz would have been enough to capture the transient behavior of interest.

Fig. 11 Photograph of the measurement system

This shows that the noise distribution for the COOLL current signals is Gaussian with zero mean and a standard deviation of 2.2 mA (equivalent to an approximate power consumption of 0.5 W).

Next, we provide an illustrative example corresponding to a test signal of a fan (Fig. 13). The total duration of the measurement is 6 s with a 0.5 s of pre-turn-on. The estimation results of TCE on the fan signal (Fig. 13) are as follows:

$$
l\hat{\mathbf{p}} = \begin{bmatrix} -1.15 \\ -0.19 \\ 0.20 \\ -0.32 \end{bmatrix}, \hat{\mathbf{a}} = \begin{bmatrix} 0.21 \\ 7.3 \times 10^{-3} \\ 6.4 \times 10^{-3} \\ 4.7 \times 10^{-3} \\ 1.0 \times 10^{-3} \end{bmatrix}, \hat{\boldsymbol{\phi}} = \begin{bmatrix} -3.10 \\ 2.94 \\ 2.72 \\ 0.22 \\ 1.27 \end{bmatrix},
$$

$$
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} |x(t_n) - s(\hat{\boldsymbol{\theta}}, t_n)|^2} = 7.1 \times 10^{-3}\,A,
$$

Table 1 COOLL dataset summary

N°	Appliance type	No. of appliances	No. of current signals (20 per appliance)
1	Drill	6	120
2	Fan	2	40
3	Grinder	2	40
4	Hair dryer	4	80
5	Hedge trimmer	3	60
6	Lamp	4	80
7	Paint stripper	1	20
8	Planer	1	20
9	Router[a]	1	20
10	Sander	3	60
11	Saw	8	160
12	Vacuum cleaner	7	140
	Total	42	840

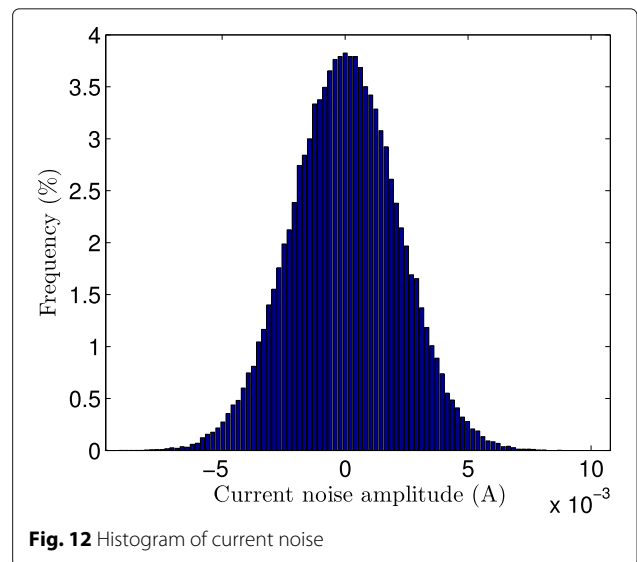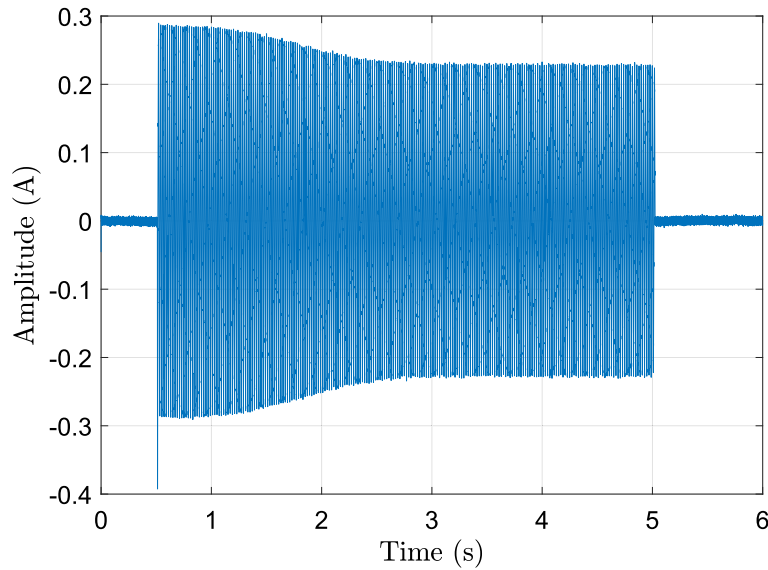[a]This is an electrical router for woodworking not a network router

Fig. 12 Histogram of current noise

Fig. 13 Turn-on transient current of a fan from the COOLL dataset

where RMSE is the root-mean-square-error. The above estimation results indicate little information on the estimation quality, especially that we are applying the algorithm on a real signal. Nonetheless, the RMSE gives an idea about the estimation quality but is still without much meaning if not considered relative to some reference value. Here, we propose to compare it to the average maximum value of the steady state amplitude (around 0.2 A). We get a relative RMSE of 3.6%. Note that we got an average relative RMSE of around 8% for the whole dataset, which is acceptable considering the variability of real signals.

Figure 14 allows to get a visual feel for the estimation quality. Figure 14a and b show a good fit between the reconstructed signal and the original zone.

4.2 Classification of cOOLL dataset's appliances using the model parameters

Here, we propose to classify the appliances of the COOLL dataset using the model parameters. We use the classical supervised k-nearest neighbors algorithm (k-NN) [34, Chap. 13], which proceeds by taking the test example (here the vector of parameters representing the test signal) and classifying it according to a majority vote of the k-nearest examples (of the training dataset). We used the Euclidean distance as a distance metric. We assess the result using K-fold cross-validation with $K = 10$. This validation works by first partitioning the dataset to K equal partitions (in our case each partition contains 84 example), then take one partition for testing and keep the other nine partitions for training, and we assess the performance using for example the classification accuracy (*CA*). This process is repeated K times, taking at each time a different partition for testing and the remaining nine for training. The final result is the average of the K accuracy results.
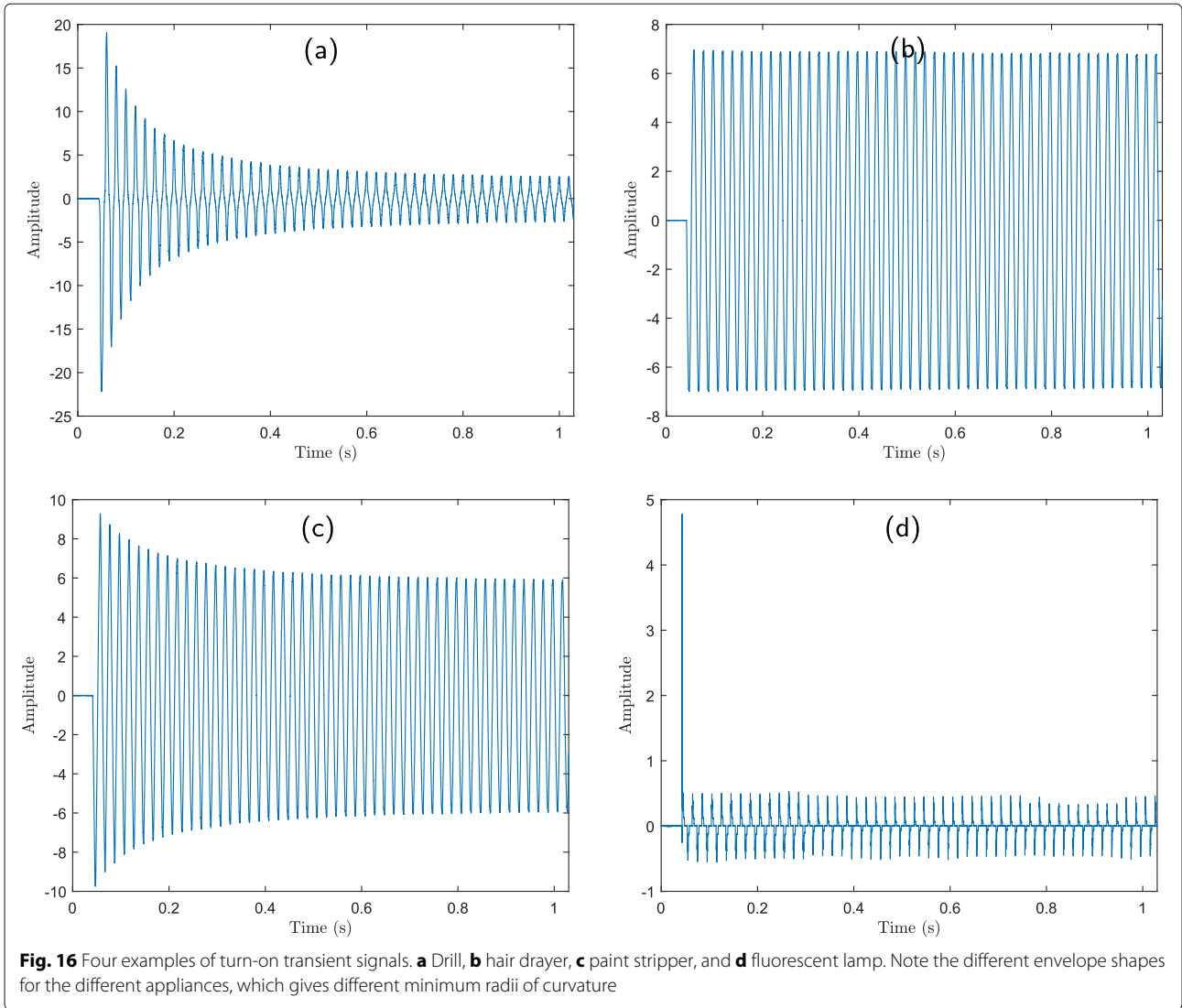
Note that the estimated values of the phase parameters ϕ_i are too random to be considered as features for the classification and, hence, are discarded hereafter. We apply the k-NN on the data using the estimated $\hat{p}_j, j = 0, \ldots, 3$ and $\hat{a}_i, i = 1, \ldots, 5$. The results are presented as a confusion matrix (Fig. 15).

The classification accuracy (*CA*) is given in the bottom rightmost corner. It is defined as $CA = \frac{TP}{\text{Tot}}$ where *TP* is the number of true positives (i.e., examples correctly classified) and Tot is the total number of considered examples. Figure 15 also gives the values of two largely used performance metrics for classification known as recall (rightmost column) and precision (bottom row). These are defined as recall $= \frac{TP}{RP}$ and precision $= \frac{TP}{CP}$ where *RP* is the number of relevant positives (i.e., examples belonging to the true class), and *CP* is the number of examples classified as positives. Note that these metrics depend on the relevant (considered) appliance class and are, hence, recomputed for each class. To illustrate this, consider the first row of the confusion matrix (Fig. 15) corresponding to the true class "drill." Here, recall $= \frac{103}{120} = 85.8\%$ (i.e., 103 examples are correctly classified among a total of 120 relevant positives (row sum)). Similarly, for the first column corresponding to the predicted (classified) class "drill," we have precision $= \frac{103}{121} = 85.1\%$ (i.e., 103

Fig. 14 Turn-on transient current of a fan from the COOLL dataset (in blue) with the reconstructed signal (in red) generated using the estimated parameters with the TCE algorithm. **a** Current signals, **b** zoom on the interval $[0.50 - 0.60]$ s, and **c** zoom on the interval $[3.50 - 3.60]$ s

Fig. 15 Classification result as a confusion matrix using model parameters $\hat{p}_i$ and $\hat{a}_i$. The bottom rightmost cell gives the classification rate, the rightmost column contains the recall values, and the bottom row the precision values

Fig. 16 Four examples of turn-on transient signals. **a** Drill, **b** hair drayer, **c** paint stripper, and **d** fluorescent lamp. Note the different envelope shapes for the different appliances, which gives different minimum radii of curvature

examples are correctly classified among 121 classified as a drill (column sum)).

We obtain a CA of 92.4%. Although the CA is higher than 92%, we expect a less variable characteristic feature capturing the envelope shape to be more relevant for the classification. In fact, the $\hat{p}_j$ are sensitive to the chosen origin of time (except the last parameter $\hat{p}_n$) and their estimated values are less stable due to the difficulty of precisely defining the origin of time for transient signals. To remedy this, we need to construct a new feature that is independent of the time origin and that still is characteristic of the envelope shape. We propose to use the *minimum radius of curvature* of the estimated envelope signal $\hat{e}(t)$ constructed using the $\hat{p}_j$ parameters. For a function $f(t)$, the radius of curvature at point t_0 is defined as [35]

$$R(t_0) = \left| \frac{(1 + f'(t_0)^2)^{3/2}}{f''(t_0)} \right| \qquad (22)$$

where $f'(t_0)$ and $f''(t_0)$ are the first and second derivatives of $f(t)$ at point t_0, respectively. Practically, we compute this value for each sample point of $\hat{e}(t_k)$ and take the minimum value $R_{\min}$. This minimum value is inversely proportional to the maximum curvature, which is a distinctive feature of the turn-on transient signals as can be seen in Fig. 16.

It is important to notice that the phase of the grid (referred to as action delay) when switching-on the appliance might affect the shape of the transient signal envelope. To investigate the influence of the phase of the grid on the chosen minimum radius of curvature, we evaluate the variation of the latter parameter for different delays as illustrated in Fig. 17 for a drill and a vacuum cleaner. As can be seen from these plots, the minimum radius of curvature remains relatively stable w.r.t. action delay parameter values. This observation is valid for most of the electrical appliances.
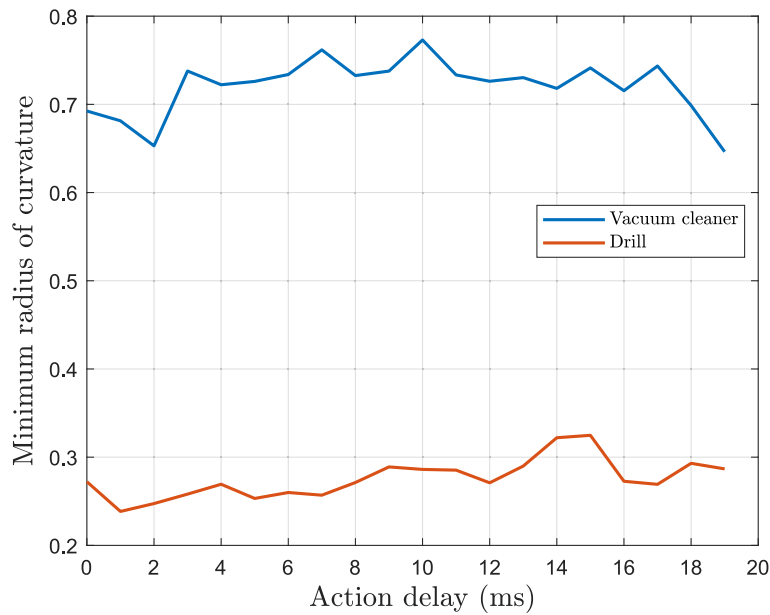
Fig. 17 Stability of the minimum radius of curvature versus various values of the grid phase (action delay) for two appliances (a drill and a vacuum cleaner). Action delay is the delay w.r.t. the (positive-to-negative) zero-crossing of the voltage signal before switching-on an appliance, e.g., an action delay of 4 ms means that the appliance is switched-on after 4 ms of voltage zero-crossing

The classification results using $R_{\min}$ and $\hat{a}_i$ are shown in Fig. 18. Compared to the previous result, we obtain an improvement of 5.6%, with a *CA* of 98.0%. Another performance metric used often in assessing classification performance is the F1 score defined as $2\frac{\text{precision}\times\text{recall}}{\text{precision}+\text{recall}}$. This metric can be computed for each appliance and gives a single number assessing the performance, which is especially helpful when comparing different classifiers or, as is the case here, the result of two different sets of features.

The F1 score results for our classification are given in Table 2. These results show an improvement in the F1 score for almost all the appliances when using the set of features $\{R_{\min}, \hat{a}_i\}$ compared to the set of features $\{\hat{p}_j, \hat{a}_i\}$.

5 Conclusion

We proposed in this paper a new mathematical representation suitable for modeling turn-on transient current signals and proposed an algorithm for the model parameter estimation. The efficiency of the algorithm is assessed

Predicted classes

True classes		Drill	Fan	Grinder	Hair drayer	Hedge trimmer	Lamp	Paint stripper	Planer	Router	Sander	Saw	Vacuum cleaner	
	Drill	119	0	0	0	1	0	0	0	0	0	0	0	99.2%
	Fan	0	40	0	0	0	0	0	0	0	0	0	0	100.0%
	Grinder	0	0	40	0	0	0	0	0	0	0	0	0	100.0%
	Hair drayer	0	0	0	80	0	0	0	0	0	0	0	0	100.0%
	Hedge trimmer	2	0	0	0	58	0	0	0	0	0	0	0	96.7%
	Lamp	0	4	0	0	0	76	0	0	0	0	0	0	95.0%
	Paint stripper	0	0	0	0	0	0	20	0	0	0	0	0	100.0%
	Planer	0	0	0	0	0	0	0	20	0	0	0	0	100.0%
	Router	0	0	0	0	0	0	0	0	20	0	0	0	100.0%
	Sander	0	0	0	0	0	0	0	0	0	55	5	0	91.7%
	Saw	0	0	0	0	1	0	0	0	0	0	156	3	97.5%
	Vacuum cleaner	0	0	0	0	0	1	0	0	0	0	0	139	99.3%
		98.3%	90.9%	100.0%	100.0%	96.7%	98.7%	100.0%	100.0%	100.0%	100.0%	96.9%	97.9%	**98.0%**

Fig. 18 Classification result as a confusion matrix using the minimum radius of curvature $R_{\min}$ and model parameters $\hat{a}_i$. The bottom rightmost cell gives the classification rate, the rightmost column contains the recall values, and the bottom row the precision values
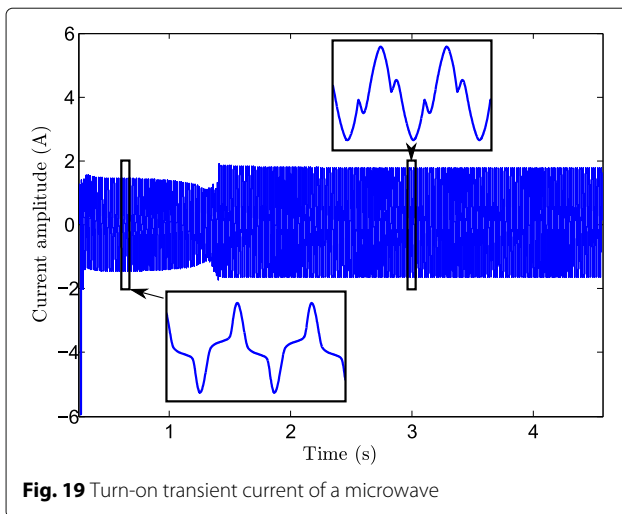
Fig. 19 Turn-on transient current of a microwave

theoretically via benchmarking its estimation error variances with respect to the CRB derived in Section 3 of this paper. Later on, the proposed parametric model is validated using real data from the COOLL dataset that we developed specifically for this research work. A "good" fitting between the proposed signal model and the real-life signals has been observed with, in particular, an average relative mean-square-error of about 8%. Note also that our experimental tests showed the need for estimating the fundamental frequency due to its deviation from the nominal value (i.e., 50 Hz). A classification method using the model parameters has been proposed. The obtained

results show the usefulness of the transient signal parameters as relevant features for the characterization of electrical appliances with a correct classification accuracy of 98% in the considered context.

The proposed model is valid for a lot of electrical appliances that show a single-phase behavior during the turn-on such as incandescent light bulbs, compact fluorescent lamps, heaters, vacuum cleaners, and hairdryers. However, some electrical appliances may have a turn-on transient current signal consisting of different phases each with a distinct signal content (harmonics with different amplitudes and phases) and a distinct envelope shape corresponding to the different regimes that the appliance goes through during turn-on. For instance, the microwave turn-on transient shown in Fig. 19 has two phases (some microwaves may have more than two) each with its specific characteristics. As a perspective work, one can consider using our model to characterize each phase independently (as a single-phase appliance) and then devise some rule to identify the corresponding multi-phase appliance (e.g., considering the occurrence of the different phases in a time series).

Authors' contributions
MNM, PR, KAM, and GL conceived and designed the experiments, analyzed the data, and interpreted the results. MNM performed the experiments and wrote the manuscript. MNM, PR, and KAM contributed in developing the model and parameter estimation algorithm. JCLB and YR provided their expertise for the power grid aspects of the experiments. All authors read and approved the final manuscript.

Availability of data and materials
The dataset created and used during the current study is freely available at https://coolldataset.github.io/.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] PRISME Laboratory, University of Orléans, 12 rue de Blois, 45067 Orléans, France. [2] GREMAN Laboratory, UMR 7347 CNRS–University of Tours, 20 avenue Monge, 37200 Tours, France.

Table 2 Comparative F1 scores of the different COOLL appliances for the classification result using the two sets of features $\{\hat{p}_j, \hat{a}_i\}$ and $\{R_{min}, \hat{a}_i\}, j = 0, \ldots, 3$, and $i = 1, \ldots, 5$

	F1 score	
	$\{\hat{p}_j, \hat{a}_i\}$	$\{R_{min}, \hat{a}_i\}$
Drill	85.5	98.8
Fan	98.7	95.2
Grinder	82.4	100.0
Hair dryer	99.4	100.0
Hedge trimmer	81.8	96.7
Lamp	96.8	96.8
Paint stripper	100.0	100.0
Planer	95.2	100.0
Router	78.8	100.0
Sander	99.2	95.7
Saw	87.4	97.2
Vacuum cleaner	100.0	98.6

References

1. C. A. García, A. Otero, X. Vila, D. G. Márquez, A new algorithm for wavelet-based heart rate variability analysis. Biomed. Signal Process. Control. **8**(6), 542–550 (2013). https://doi.org/10.1016/j.bspc.2013.05.006
2. X. Chen, H. Wen, Q. Li, T. Wang, S. Chen, Y.-P. Zheng, Z. Zhang, Identifying transient patterns of in vivo muscle behaviors during isometric contraction by local polynomial regression. Biomed. Signal Process. Control. **24**, 93–102 (2016). https://doi.org/10.1016/j.bspc.2015.09.009
3. T. P. Exarchos, A. T. Tzallas, D. I. Fotiadis, S. Konitsiotis, S. Giannopoulos, EEG transient event detection and classification using association rules. IEEE Trans. Inf. Technol. Biomed. **10**(3), 451–457 (2006). https://doi.org/10.1109/TITB.2006.872067
4. A. Belsak, J. Flasker, Adaptive wavelet transform method to identify cracks in gears. EURASIP J. Adv. Signal Proc. **2010**(1), 879875 (2010). https://doi.org/10.1155/2010/879875
5. C. Capilla, Application of the Haar wavelet transform to detect microseismic signal arrivals. J. Appl. Geophys. **59**(1), 36–46 (2006). https://doi.org/10.1016/j.jappgeo.2005.07.005
6. X. Li, Z. Li, E. Wang, J. Feng, L. Chen, N. Li, X. Kong, Extraction of microseismic waveforms characteristics prior to rock burst using hilbert–huang transform. Measurement. **91**, 101–113 (2016). https://doi.org/10.1016/j.measurement.2016.05.045
7. J. Seymour, T. Horsley, The seven types of power problems. White paper. **18**, 1–21 (2005)
8. M. H. J. Bollen, E. Styvaktakis, I. Y.-H. Gu, Categorization and analysis of power system transients. IEEE Trans Power Deliv. **20**(3), 2298–2306 (2005). https://doi.org/10.1109/TPWRD.2004.843386
9. S. Wang, Z. K. Zhu, Y. He, W. Huang, Adaptive parameter identification based on Morlet wavelet and application in gearbox fault feature detection. EURASIP J. Adv. Signal Process. **2010**(1), 842879 (2010). https://doi.org/10.1155/2010/842879
10. W. Jiao, S. Qian, Y. Chang, S. Yang, Research on vibration response of a multi-faulted rotor system using LMD-based time-frequency representation. EURASIP J. Adv. Signal Process. **2012**(1), 73 (2012). https://doi.org/10.1186/1687-6180-2012-73
11. S. B. Leeb, S. R. Shaw, J. L. Kirtley Jr, Transient event detection in spectral envelope estimates for nonintrusive load monitoring. Power Deliv. IEEE Trans. **10**(3), 1200–1210 (1995)
12. C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, P. Armstrong, Power signature analysis. Power Energy Mag. IEEE. **1**(2), 56–63 (2003)
13. H.-H. Chang, H.-T. Yang, Applying a non-intrusive energy-management system to economic dispatch for a cogeneration system and power utility. Appl. Energy. **86**(11), 2335–2343 (2009)
14. R. Kumaresan, D. Tufts, Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise. IEEE Trans. Acoust. Speech. Signal Process. **30**(6), 833–840 (1982)
15. L. Lovisolo, M. P. Tcheou, E. A. B. da Silva, M. A. M. Rodrigues, P. S. R. Diniz, Modeling of electric disturbance signals using damped sinusoids via atomic decompositions and its applications. EURASIP J. Adv. Signal Process. **2007**(1), 029507 (2007). https://doi.org/10.1155/2007/29507
16. R. Boyer, K. Abed-Meraim, Audio modeling based on delayed sinusoids. IEEE Trans. Speech Audio Process. **12**(2), 110–120 (2004). https://doi.org/10.1109/TSA.2003.819953
17. D. V. Rubtsov, J. L. Griffin, Time-domain Bayesian detection and estimation of noisy damped sinusoidal signals applied to NMR spectroscopy. J Magn. Reson. **188**(2), 367–379 (2007). https://doi.org/10.1016/j.jmr.2007.08.008
18. M. A. Al-Radhawi, K. Abed-Meraim, Parameter estimation of superimposed damped sinusoids using exponential windows. Signal Process. **100**, 16–22 (2014). https://doi.org/10.1016/j.sigpro.2013.12.025
19. R. Prony, Essai expérimental et analytique : sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures. J. de l'École Polytechnique Floréal et Plairial. **1**(22), 24–76 (1795)
20. V. F. Pisarenko, The retrieval of harmonics from a covariance function. Geophys. J. Int. **33**(3), 347–366 (1973)
21. Y. Hua, T. K. Sarkar, Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. Acoust. Speech. Signal Process. IEEE Trans. **38**(5), 814–824 (1990)
22. R. Roy, T. Kailath, ESPRIT—estimation of signal parameters via rotational invariance techniques. Acoust. Speech. Signal Process. IEEE Trans. **37**(7), 984–995 (1989)
23. R. Schmidt, Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. **34**(3), 276–280 (1986)
24. E. K. Howell, How switches produce electrical noise. Electromagn. Compat. IEEE Trans. **EMC-21**(3), 162–170 (1979)
25. CENELEC, Voltage characteristics of electricity supplied by public electricity networks. European Standard EN 50160 (2010)
26. S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. (Prentice Hall PTR, Upper Saddle River, 1993)
27. E. Aboutanios, B. Mulgrew, Iterative frequency estimation by interpolation on fourier coefficients. IEEE Trans. Signal Process. **53**(4), 1237–1242 (2005)
28. M. Nait Meziane, P. Ravier, G. Lamarque, J.-C. Le Bunetel, Y. Raingeaud, in *42nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. High accuracy event detection for non-intrusive load monitoring (IEEE, 2017), pp. 2452–2456. https://doi.org/10.1109/ICASSP.2017.7952597
29. T. F. Coleman, Y. Li, On the convergence of interior-reflective newton methods for nonlinear minimization subject to bounds. Math. Program. **67**(1), 189–224 (1994). https://doi.org/10.1007/BF01582221
30. T. F. Coleman, Y. Li, An interior trust region approach for nonlinear minimization subject to bounds. SIAM J. Optim. **6**(2), 418–445 (1996). https://doi.org/10.1137/0806023
31. Y.-x. Yuan, in *Iciam*. A review of trust region algorithms for optimization, vol. 99 (Citeseer, 2000), pp. 271–282
32. T. Picon, M. Nait Meziane, P. Ravier, G. Lamarque, C. Novello, J.-C. Le Bunetel, Y. Raingeaud, COOLL: Controlled on/off loads library, a public dataset of high-sampled electrical signals for appliance identification. arXiv preprint arXiv:1611.05803 [cs.OH] (2016)
33. M. Nait Meziane, T. Picon, P. Ravier, G. Lamarque, J.-C. Le Bunetel, Y. Raingeaud, in *Conference on Environment and Electrical Engineering (EEEIC), 2016 Proceedings of the 16th IEEE International*. A measurement system for creating datasets of on/off-controlled electrical loads, (2016), pp. 2579–2583
34. J. Friedman, T. Hastie, R. Tibshirani, *The Elements of Statistical Learning. vol. 1*. (Springer, New York, 2001)
35. J. D. Lawrence, *A Catalog of Special Plane Curves*. (Courier Corporation, North Chelmsford, 1972)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.