Putting the PRNU Model in Reverse Gear: Findings with Synthetic Signals

Miguel Masciopinto, Fernando Pérez-González
Signal Theory and Communications Department, University of Vigo
E. E. Telecomunicación, Campus-Lagoas Marcosende, Vigo 36310, Spain
Email: {mmasciopinto, fperez}@gts.uvigo.es

***Abstract*—The prevalent model for the photoresponse non-uniformity (PRNU) of digital cameras is (negatively) tested. The PRNU serves as a fingerprint of the underlying device that has proven its usefulness in image and video forensics. When such model is applied to derive analytical expressions characterizing the PRNU detection statistic, the predictions do not conform with reality. However, those expressions are thoroughly validated here through extensive experimentation using synthetic signals with parameters taken from real images. As a consequence, the current multiplicative PRNU model must be revised. This opens new venues for performance improvements in both the extraction and detection of device fingerprints.**

I. Introduction

The photoresponse non-uniformity, or PRNU for short, is an intrinsic property of digital imaging sensors due to imperfections in the manufacturing process and different sensitivity to light of individual detectors. Thanks to its random nature, the PRNU acts as a fingerprint of the underlying device that has proven its usefulness in image and video forensics. The PRNU can be used for source attribution, i.e., deciding which camera among a group has taken a certain picture; fingerprint detection, i.e., deciding whether a given image was captured by a certain device; image clustering, i.e., grouping images taken with the same devices; and tampering detection, i.e., deciding whether an image was manipulated by looking for inconsistencies in its PRNU.

Being a very weak signal, reliable extraction of the PRNU from a set of images has paramount importance, but this can only be done with a proper model of the residuals left by a denoising process aimed at reducing the interference of the image that contains the fingerprint. The simplified multiplicative model, prevalent in all the recent literature (e.g., [1]–[3]), was settled in [4], [5], and leads to nearly optimal PRNU estimators and detectors. Minor variations [6] of the model only produce very similar detection and estimation statistics. Surprisingly, mathematical analyses of performance are very scarce, and mostly focus on the null hypothesis (i.e. testing against the wrong source). In [4], [5], the null hypothesis (denoted by H_0) is modeled by a generalized Gaussian distribution whose parameters are estimated by correlating a set of images with the fingerprints extracted from different cameras. When the normalized crosscorrelation detection statistic for H_0 is considered, a zero-mean Gaussian distribution with variance inversely proportional to the number of pixels is proposed in [7]. In [8], all possible shifts excluding a small neighborhood surrounding the aligned position are used to individually estimate the variance of H_0 for each test image; resulting in the Peak to Correlation Energy (PCE) as a better detection statistic than the normalized correlation, as shown in [9].

On the other hand, the statistical modeling of the alternative (i.e., H_1) hypothesis (i.e., testing an image against a matching camera) is more complex. The analysis made in [4], [5] tries to infer the expected value of the normalized correlation for each test image, based on the image intensity, texture and flattening, but it needs to be built in advance from a set of training images. In [10], for the problem of matching two extracted fingerprints, H_1 is modeled by a Gaussian with theoretical values for its mean and variance; unfortunately, this analysis cannot be directly extended to the more common fingerprint detection problem.

Recent investigations [11] have been able to produce analytical expressions for the two hypotheses of the fingerprint detection problem. These analytical expressions allow one to compute theoretically the probabilities of detection and false alarm and assess the expected performance *for a given* test image. Paradoxically, these expressions do not conform with reality with the accuracy that would be expected if the model were correct. Thus, in this paper we take a reverse approach towards validation of the PRNU model: we generate synthetic PRNUs with the assumed characteristics and "plant" them on both synthetic and real images. We increase the level of complexity of the synthetic signals to account for cross-dependencies and local correlations. We find that, while the theory accurately matches the experimental results obtained with synthetic PRNUs, such is not the case when dealing with real images. This indicates that the long-standing model for the PRNU falls short of holding in practice, and calls for a revision.

Notation: Vectors are represented in boldface. The ith component of the vector $\mathbf{x}$ is represented by $x[i]$. $\mathbf{I}$, $\mathbf{1}$ and $\mathbf{0}$ denote the identity matrix, the all-ones and the all-zeros vectors, respectively. The sample-wise (Hadamard) product between $\mathbf{x}$ and $\mathbf{y}$ is denoted by $\mathbf{x} \circ \mathbf{y}$, while the scalar product is $\langle \mathbf{x}, \mathbf{y} \rangle$. Any other operations among vectors or matrices, such as addition, ratio, or raising to a power, are element-wise.

The rest of the paper is organized as follows: Sect. II summarizes the existing model and the test statistic; Sect. III presents the main theoretical findings in [11]; Sect. IV des-

cribes the experimental setup followed to produce the results given in Sect. V. Finally, Sect. VI discusses the potential consequences of our work.

II. Mathematical background on the PRNU model

We start by adopting the sensor output model proposed in [12] which has served as the basis for most of the subsequent literature

$$\mathbf{y} = g^{\gamma}[(\mathbf{1} + \mathbf{k}_0) \circ \mathbf{x}_0 + \mathbf{r}]^{\gamma} + \mathbf{n}_q, \tag{1}$$

where $\mathbf{y}$ is the output of the imaging sensor rearranged in one-dimensional form, $\mathbf{x}_0$ is the incident light intensity, $\mathbf{k}_0$ is the PRNU term, g and γ are the color channel gain and the gamma correction respectively, $\mathbf{r}$ accounts for other noise sources including dark currents, and $\mathbf{n}_q$ represents the quantization noise. This model can be expanded in Taylor series around $\mathbf{k}_0 = \mathbf{0}$, and the constants g and γ absorbed as $\mathbf{x} = (g\mathbf{x}_0)^{\gamma}$ and $\mathbf{k} = \gamma\mathbf{k}_0$, to write

$$\mathbf{y} = (\mathbf{1} + \mathbf{k}) \circ \mathbf{x} + \mathbf{n}, \tag{2}$$

where $\mathbf{n} \doteq \gamma(g\mathbf{x}_0)^{\gamma}\mathbf{r}/\mathbf{x}_0 + \mathbf{n}_q$ collects the remaining noise sources.

In practice, it is necessary to obtain an estimate $\hat{\mathbf{x}}$ of the signal $\mathbf{x}$ by applying an image denoising algorithm to $\mathbf{y}$. The denoising residue for a single image is defined as $\mathbf{w} \doteq \mathbf{y} - \hat{\mathbf{x}}$. In this work, the denoising algorithm described in [13] is considered without loss of generality.

A. PRNU fingerprint estimation

The PRNU term of any camera device can be estimated from a set $\{\mathbf{y}^l\}_{l=1}^{L}$ of images taken with the same device. After obtaining the residue $\mathbf{w}^l = \mathbf{y}^l - \hat{\mathbf{x}}^l$, and assuming i.i.d. Gaussian noise, the maximum likelihood estimator (MLE) $\hat{\mathbf{k}} = (\sum_{l=1}^{L} \mathbf{w}^l \circ \hat{\mathbf{x}}^l)/(\sum_{l=1}^{L} \hat{\mathbf{x}}^l \circ \hat{\mathbf{x}}^l)$ derived in [6] can be used. Next, following Chen et al. [4], mean subtraction and Wiener filtering in the Discrete Fourier Transform (DFT) domain are applied to $\hat{\mathbf{k}}$, in order to remove different artifacts that are not the PRNU itself. The MLE of the PRNU implicitly assumes the model

$$\hat{\mathbf{k}} = \mathbf{k} + \mathbf{n}_e, \tag{3}$$

where the samples of the extraction noise $\mathbf{n}_e$ are i.i.d. Gaussian with zero-mean and variance σ_e^2.

B. PRNU detection problem formulation

Let $\mathbf{y}_t$ be a test image of size $N_t \times N_t$ taken with a target camera device of native resolution $N \times N$ ($N_t \leq N$), and $\mathbf{w}_t = \mathbf{y}_t - \hat{\mathbf{x}}_t$ the corresponding residue after denoising. In order to attribute $\mathbf{y}_t$ to one specific target camera device, the following hypothesis testing problem is formulated:

H_0 : $\mathbf{w}_t$ and $\hat{\mathbf{k}}$ correspond to different PRNUs;
H_1 : $\mathbf{w}_t$ and $\hat{\mathbf{k}}$ correspond to the same PRNU;

where $\mathbf{y}_t$ is attributed to the target camera if H_1 holds.

From (2), we can write

$$\mathbf{w}_t = \mathbf{k} \circ \mathbf{x}_t + \mathbf{n}_t, \tag{4}$$

where $\mathbf{n}_t$ is the estimation error with variance σ_t^2.

When test images may have been cropped, it is not advisable to use the detector based on the statistic $\langle \mathbf{w}_t, \hat{\mathbf{k}} \circ \hat{\mathbf{x}}_t \rangle$ due to computational reasons [8]. Thus, we consider instead the cross-correlation test proposed in [8]:

$$r \doteq \langle \mathbf{w}_t, \hat{\mathbf{k}} \rangle \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \tag{5}$$

for a fixed threshold λ, which is usually chosen to bound the false alarm probability.

In practice one should take the maximum of the cross-correlation for every possible lag when alignment is unknown. The formal analysis of such statistic is a bit more complicated, but once the probabilities of false positive and misdetection are correctly predicted for every single lag, the corresponding probabilities for the maximum will also match the empirical ones. Thus, for simplicity, here we will assume perfect alignment between $\mathbf{w}_t$ and $\hat{\mathbf{k}}$.

III. Statistical modeling of the test statistic

Finding the distribution of r under H_0 and H_1 is crucial not only for establishing some design parameters (prominently, threshold λ), but also to assess the expected performance of the detector. Unsurprisingly, the desired distributions strongly depend on the underlying statistical assumptions for all the signals involved in the models above. We illustrate this next by considering the cases of white and non-white signals.

A. White signals

If the samples of the signals $\mathbf{n}_t$, $\mathbf{k}$, and $\mathbf{n}_e$ are mutually uncorrelated, zero-mean i.i.d. (i.e., white), and N_t^2 is large, it can be shown [11] that r under H_0 is $\mathcal{N}(0, \sigma_{u,0}^2)$ where

$$\sigma_{u,0}^2 = N_t^2\left[\sigma_k^4(\mu_x^2+\sigma_x^2)+\sigma_k^2\left(\sigma_e^2(\mu_x^2+\sigma_x^2)+\sigma_t^2\right)+\sigma_e^2\sigma_t^2\right], \tag{6}$$

with $\mu_x \doteq \frac{1}{N_t^2}\sum_{i=1}^{N_t^2} x[i]$ and $\sigma_x^2 \doteq \frac{1}{N_t^2-1}\sum_{i=1}^{N_t^2}(x[i]-\mu_x)^2$.

When H_1 holds, the distribution of r can be also modeled as a normal random variable, with mean and variance:

$$\mu_{H_1} \doteq \mathrm{E}(r|H_1) = N_t^2\sigma_k^2\mu_x, \tag{7}$$

$$\sigma_{u,1}^2 = N_t^2\left[\sigma_k^4(2\mu_x^2+3\sigma_x^2)+\sigma_k^2\left(\sigma_e^2(\mu_x^2+\sigma_x^2)+\sigma_t^2\right)+\sigma_e^2\sigma_t^2\right]. \tag{8}$$

B. Non-white signals

The i.i.d. assumption is not entirely realistic, because signals may exhibit dependencies between neighboring pixels due to imaging sensor operations, demosaicing, compression, and denoising. Even though signals are not even stationary, the long averages implicit in (5) justify the use for analysis purposes of stationary signals with average first and second-order moments. In particular, we can improve the i.i.d. model by allowing for the existence of local autocorrelations in $\mathbf{k}, \mathbf{n}_t, \hat{\mathbf{k}}, \mathbf{n}_e$. The autocorrelation sequence of $\mathbf{k}$ will be denoted by $\mathbf{R}_k[i]$, with a direct extension to the autocorrelations of the remaining signals. Notice that if $\mathbf{u}, \mathbf{v}$ are uncorrelated, and $\mathbf{s} = \mathbf{u} \circ \mathbf{v}$, then $\mathbf{R}_s[i] = \mathbf{R}_u[i] \cdot \mathbf{R}_v[i]$, which becomes useful when deriving the pdf of r for this case.

From (6) and (8), the expressions for $\sigma^2_{u,0}$ and $\sigma^2_{u,1}$ for colored signals can be shown to be [11]

$$\sigma^2_{u,0} = N_t^2 \Bigg[(\mu_x^2+\sigma_x^2) \sum_{\Delta_q} \mathbf{R}_k^2 + (\mu_x^2+\sigma_x^2) \sum_{\Delta_q} \mathbf{R}_k \circ \mathbf{R}_{n_e} + \sum_{\Delta_q} \mathbf{R}_k \circ \mathbf{R}_{n_t} + \sum_{\Delta_q} \mathbf{R}_{n_e} \circ \mathbf{R}_{n_t} \Bigg], \tag{9}$$

$$\sigma^2_{u,1} = N_t^2 \Bigg[(2\mu_x^2+3\sigma_x^2) \sum_{\Delta_q} \mathbf{R}_k^2 + (\mu_x^2+\sigma_x^2) \sum_{\Delta_q} \mathbf{R}_k \circ \mathbf{R}_{n_e} + \sum_{\Delta_q} \mathbf{R}_k \circ \mathbf{R}_{n_t} + \sum_{\Delta_q} \mathbf{R}_{n_e} \circ \mathbf{R}_{n_t} \Bigg], \tag{10}$$

where Δ_q denotes a square neighborhood of radius q, i.e., $(2 \cdot q + 1)^2$ pixels around $(0, 0)$.

To use these formulas in practice, one must estimate the autocorrelation matrices from the available images. $\mathbf{R}_{n_t}$ can be estimated for the target image by assuming that $\mathbf{n}_t \approx \mathbf{y}_t - \hat{\mathbf{x}}_t$ (due to the weakness of the PRNU) as its sample autocorrelation. $\mathbf{R}_{\hat{k}}$ can be estimated also from the sample autocorrelation of $\hat{\mathbf{k}}$. The estimation of $\mathbf{R}_k$ is more difficult, because even for a large number of images, $\mathbf{k}$ will be eclipsed by $\mathbf{n}_e$. We propose the following method to mitigate this influence: split the available images from the target camera into two disjoint sets, and estimate the PRNU for each set, say $\hat{\mathbf{k}}_1$ and $\hat{\mathbf{k}}_2$, as in Sect. II-A. Then, $\mathbf{R}_k$ can be estimated as the sample cross-correlation between $\hat{\mathbf{k}}_1$ and $\hat{\mathbf{k}}_2$, where we benefit from the fact that $\mathbf{n}_{e_1}$ (resp. $\mathbf{n}_{e_2}$) does not contain traces of $\hat{\mathbf{k}}_2$ (resp. $\hat{\mathbf{k}}_1$) and both estimation errors are uncorrelated. Finally, $\mathbf{R}_{n_e}$ is estimated as $\mathbf{R}_{\hat{k}} - \mathbf{R}_k$, hence assuming uncorrelation between $\mathbf{n}_e$ and $\mathbf{k}$, which is reasonable if a moderate number of images is used in the estimation of the PRNU.

After rearranging the signals in their original imaging sensor two-dimensional form, the empirical two-dimensional autocorrelations of $\mathbf{k}$, $\mathbf{n}_e$ and $\mathbf{n}_t$ are shown in Figs. 1a-1c. Notice that the signals $\mathbf{k}$, $\mathbf{n}_e$ and $\mathbf{n}_t$ are not white in general.

C. Cross-correlations

As stated above $\mathbf{k}$ and $\mathbf{n}_e$ can be safely considered mutually uncorrelated, as it is also the case with $\mathbf{n}_e$ and $\mathbf{n}_t$. However, due to imperfect denoising, $\mathbf{n}_t$ will in general contain traces of $\mathbf{k}$ (see Fig. 1d). If such cross-correlation is to be accounted for, then the mean of r for H_1 must be corrected as follows:

$$\mathrm{E}(r|H_1) = N_t^2 \sigma_k^2 \mu_x + \langle \mathbf{k}, \mathbf{n}_t \rangle, \tag{11}$$

Unfortunately, in a real scenario this correction cannot be directly applied, because the last term is unknown.

IV. Experimental set-up

In principle, the formulas in Sect. III should serve to predict the performance of the cross-correlation detector; after all, they are derived directly from the PRNU model in Sect. II by making assumptions of increasing complexity (i.e., from white, uncorrelated to non-white, correlated signals). However, as we will later confirm, these theoretical predictions ostensibly fail with real signals. In order to cast some light on the reasons why this happens, we have conducted a series of experiments with synthetic and real images. These experiments validate our derivations to the extent that it is the PRNU model itself what should be questioned.

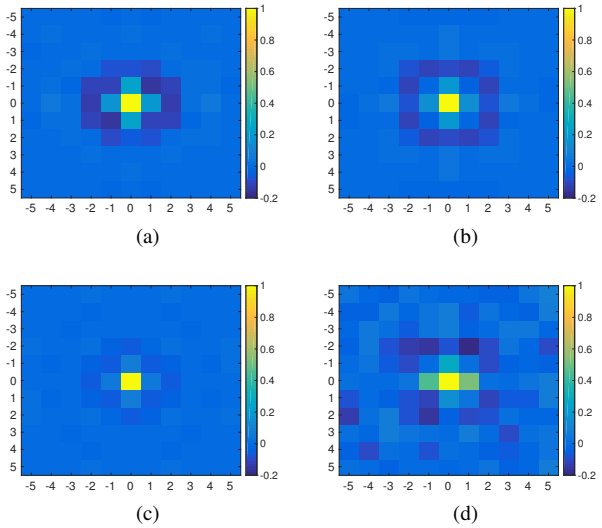


Fig. 1: Two-dimensional auto- and cross-correlations from images of the Nikon D60 device: (a) $\mathbf{R}_k$, (b) $\mathbf{R}_{n_e}$, (c) $\mathbf{R}_{n_t}$; 97 images to obtain $\hat{\mathbf{k}}_1$ and $\hat{\mathbf{k}}_2$; $L = 50$ images to obtain $\hat{\mathbf{k}}$; $\mathbf{R}_{n_t}$ obtained by averaging over 147 images of the Nikon D60; (d) average cross-correlation between $\mathbf{k}$ and $\mathbf{n}_t$ over 197 images. All plots are normalized by their value at the origin.

One of the advantages of using synthetic signals is that the denoising process, which is hard to model accurately, can be bypassed. This is carried out by directly sythesizing the denoising residuals; for instance, the signals $\mathbf{k}$ and $\mathbf{n}_t$ in (4) can be synthesized as samples of the corresponding random variables, either assuming whiteness and uncorrelatedness or complicating the model to account for auto- and cross-correlations. Some of the parameters/values (e.g. variances) needed in the synthetic experiments are estimated from real images. The image database used is a collection of pictures from different camera devices containing TIFF images taken with our own cameras, from the Dresden image database [14], and from the RAISE database [15] (see Table I). The same database is also used to perform the experiments on real images. Next, we describe how each experiment was set up.

Experiment E1. Signals $\mathbf{k}, \mathbf{n}_t, \mathbf{n}_e$ and $\mathbf{x}_t$ are synthesized as white and mutually uncorrelated. $\hat{\mathbf{k}}$ and $\mathbf{w}_t$ are generated according to (3) and (4), respectively.

Experiment E2. Synthetic signals are colored by filtering a white signal through a filter with frequency response $\sqrt{|\mathrm{DFT}(\mathbf{R}_a)|}$, where $\mathbf{R}_a$ is the desired autocorrelation. The images from the database are used to estimate the autocorrela-

TABLE I: Database of real images from different cameras.

Camera Model	Sensor	Native resolution	Devices	Number of images	Database
Canon 600D	22.3x14.9mm CMOS	5184x3456	1	241	Own
Canon1100D	23.2x14.7mm CMOS	4272x2848	3	316/122/216	Own
Nikon D60	23.6x15.8mm CCD	3872x2592	1	197	Own
Nikon D70	23.7x15.6mm CCD	3008x2000	2	43/43	Dresden
Nikon D70S	23.7x15.6mm CCD	3008x2000	2	43/47	Dresden
Nikon D90	23.6x15.8mm CMOS	4288x2848	1	250	RAISE
Nikon D200	23.6x15.8mm CCD	3872x2592	2	48/43	Dresden
Nikon D3000	23.6x15.8mm CCD	3872x2592	1	230	Own
Nikon D3200	23.2x15.4mm CMOS	6016x4000	1	250	Own
Nikon D5100	23.6x15.6mm CMOS	4928x3264	1	250	Own
Nikon D7000	23.6x15.6mm CMOS	4928x3264	1	250	RAISE

tions $\mathbf{R}_{x_t}$ and $\mathbf{R}_{n_t}$.[1] Autocorrelations $\mathbf{R}_k$ and $\mathbf{R}_{n_e}$ are estimated as described in Sect. III-B, with $\mathbf{R}_{\hat{k}}$ the autocorrelation of $\hat{\mathbf{k}}$ averaged over $L = 50$ images. $\hat{\mathbf{k}}$ and $\mathbf{w}_t$ are constructed as in E1.

Experiment E3. The images from the database are sequentially used as $\mathbf{y}_t$, from which $\hat{\mathbf{x}}_t$ are obtained after denoising them with the filter in [13]. $\mathbf{n}_t$ is simply $\mathbf{y}_t - \hat{\mathbf{x}}_t$. All other signals are synthesized as in E2. This experiment aims at determining whether denoising artifacts not captured by the autocorrelations may influence the results.

Experiment E4. A synthetic PRNU $\mathbf{k}_s \sim \mathcal{N}(\mathbf{0}, \sigma_k^2\mathbf{I})$ is embedded in database images $\mathbf{y}$ using $\mathbf{y}_t = \mathbf{y} \circ (\mathbf{1} + \mathbf{k}_s)$. Thus, the test images $\mathbf{y}_t$ contain a known PRNU following the model in (1). This signal is denoised to produce the residual $\mathbf{w}_t$, and $\hat{\mathbf{k}}_s$ is generated following (3), where $\mathbf{n}_e$ is synthesized as in E2.

Experiment E5. Same as E4, but now $\mathbf{k}_s$ is an actual PRNU extracted from real images.

Experiment E6. Database images are subjected to the estimation and detection procedures described in Sects.II-A and II-B.

V. Experiments

In all synthetic experiments reported in this section, the following values are considered: $N = 4096$, $N_t = 2048$, $\sigma_e^2 = 10^{-6}$, $\sigma_k^2 = 10^{-7}$. For the experiments on real images, σ_k^2 is estimated following the procedure that splits the available images into two disjoint sets (see Sect. III-B) as $\sigma_k^2 = \mathbf{R}_k[0,0]$. Then, $\hat{\sigma}_k^2 \approx 10^{-7}$ for all cameras except for Canon 600D, where $\hat{\sigma}_k^2 \approx 10^{-6}$. Images taken with the reflex camera Nikon D60 are used for the H_1 hypothesis; while the other cameras of Table I are used to evaluate the H_0 hypothesis. Moreover, $q = 10$ in (9) and (10). The results are reported as histograms for the H_0 hypothesis, and as plots of the experimental samples of $r|H_1$ vs. the theoretical values μ_{H_1}.

A. Synthetic signals

Figures 2a and 2c show the distributions for Experiment E1. The theoretical ones are obtained through (6)-(8). As we can see, for both hypotheses the theory matches perfectly the empirical results.[2] Figures 2b and 2d show the corresponding distributions for Experiment E2, while Fig. 2e depicts the distribution of H_1 for Experiment E3. Here, (9) and (10) are used to evaluate the theoretical variances $\sigma_{u,0}^2$ and $\sigma_{u,1}^2$. Again, the matching between the theoretical and the empirical results is remarkable.

For the remaining experiments, we only report the results on H_1, since hypothesis H_0 does not show any mismodeling even for real images.

Experiments E4 and E5 embed a known PRNU $\mathbf{k}_s$ into the images in order to measure the extent to which incomplete knowledge of the PRNU affects the results in a real detection scenario. In E4, a synthetic PRNU is used; Figure 3a shows that the values of $r|H_1$ are lower than predicted by (7). In Experiment E5, the PRNU is extracted using $L = 100$ images of the cloudy sky taken with the same Nikon D60 that is used for hypothesis H_1. Figure 3b now shows values of $r|H_1$ that are higher than expected. This is likely due to the fact that we are embedding a similar PRNU than is already contained in the images used for H_1. This is overcame by embedding a PRNU extracted from a different camera. Figure 3c shows the result, where a very similar behavior to Fig. 3a is evident.

The discrepancies observed in Figs. 3a and 3c are explained by the inability of the denoising process to remove all of the embedded PRNU. This leaves some correlation between $\mathbf{k}_s$ and $\mathbf{n}_t$ which can be corrected as in (11) with $\mathbf{n}_t = \mathbf{y}_t - \hat{\mathbf{x}}_t - \mathbf{y} \circ \mathbf{k}_s$ (see Fig. 1d and Experiment E4). Figure 3d shows the result after applying the correction in (11) when a synthetic PRNU is embedded, while Fig. 3e reports the performance of the correction when the embedded PRNU comes from a different camera. With this correction, the theoretical predictions match again the empirical results.

B. Real images

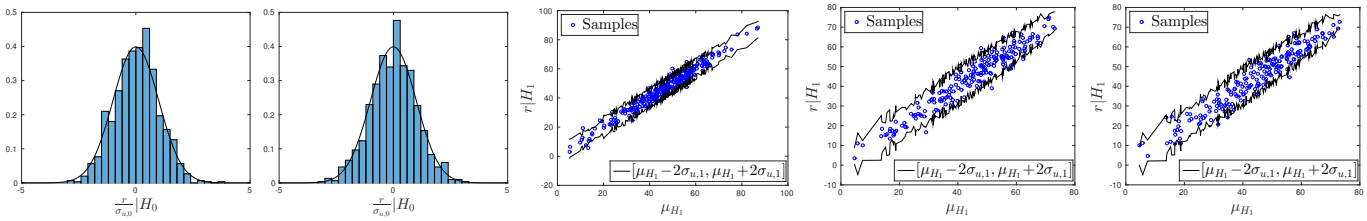
Figure 4 shows the empirical vs. the theoretical values of the detector in (5) when images from a target camera are directly used as H_1 hypothesis ($L = 50$ images are used to estimate $\hat{\mathbf{k}}$, and the remaining as test images $\mathbf{y}_t$). Although not shown here, the results for H_0 are well fit by the theory; in contrast, for H_1 hypothesis (see Fig. 4) they fall considerably out of the theoretical confidence band for all cameras tested (the plots correspond to four of them).

VI. Discussion

Our step-by-step generation of synthetic signals, entailing increasing levels of complexity, and the mirroring theoretical analyses, carried out systematically, point to one root cause: the PRNU model. Intriguinly enough, this model has remained unchallenged across practically all the research in PRNU-based forensics during the past decade. Validating the prevalent model is far from trivial, because the signals involved are extremely weak. However, we have already taken some steps in this direction which show promising results: it may well be that an improved model will yield better detection rates or, equivalently, reduce the size of areas under test when constructing tampering detectors that use the PRNU. Thus, although this paper reports "negative results", we expect it to open new venues of future "positive" research.

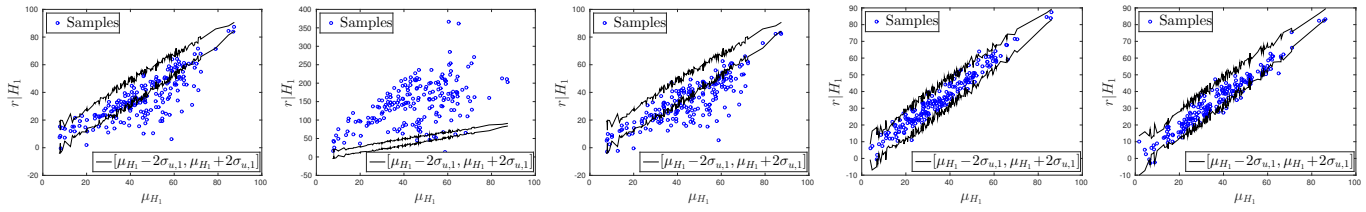
[1] We assume that $\mathbf{R}_{n_t} \approx \mathbf{R}_{w_t}$, since $\mathbf{k} \circ \mathbf{x}_t \ll \mathbf{n}_t$.

[2] Most of the samples of H_1 lie in the theoretical band $[\mu_{H_1} - 2\sigma_{u,1}, \mu_{H_1} + 2\sigma_{u,1}]$, which encloses $\sim 95\%$ of the values of a Gaussian distribution.



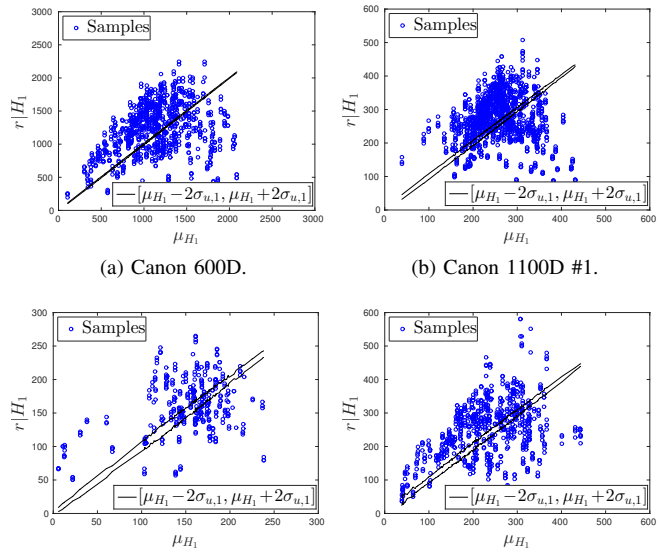
(a) (b) (c) (d) (e)

Fig. 2: Empirical and theoretical distributions. H_0 hypothesis: (a) white signals, (b) colored signals. H_1 hypothesis for synthetic signals: (c) white, (d) colored, (e) colored with $\mathbf{x}_t$ and $\mathbf{n}_t$ directly extracted from real images.



(a) (b) (c) (d) (e)

Fig. 3: Empirical vs. theoretical values of H_1 when a PRNU $\mathbf{k}_s$ is embedded (target camera: Nikon D60) with the following features: (a) Synthetic white Gaussian; (b) extracted from cloudy sky images taken with Nikon D60; (c) extracted from Nikon D90 images. Considering the correction in (11): (d) synthetic white Gaussian; (e) extracted from Nikon D90 images.



(a) Canon 600D. (b) Canon 1100D #1.

(c) Canon 1100D #2. (d) Nikon D60.

Fig. 4: Empirical vs. theoretical detection values on H_1 for real images from different camera devices.

ACKNOWLEDGMENT

GPSC is funded by the Agencia Estatal de Investigación (Spain) and the European Regional Development Fund (ERDF) under projects WINTER (TEC2016-76409-C2-2-R) and COMONSENS (TEC2015-69648-REDC). Also funded by the Xunta de Galicia and the European Union (European Regional Development Fund - ERDF) under projects Agrupación Estratéxica Consolidada de Galicia accreditation 2016-2019, Grupo de Referencia ED431C2017/53 and Red Temática RedTEIC 2017-2018.

REFERENCES

[1] P. Korus and J. Huang, "Multi-scale analysis strategies in PRNU-based tampering localization," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 4, pp. 809–824, Apr. 2017.

[2] F. Marra, G. Poggi, C. Sansone, and L. Verdoliva, "Blind PRNU-based image clustering for source identification," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 9, pp. 2197–2211, Sep. 2017.

[3] L. Bondi, F. Pérez-González, P. Bestagini, and S. Tubaro, "Design of projection matrices for PRNU compression," in *IEEE Workshop on Inf. Forensics and Security (WIFS)*, Dec. 2017, pp. 1–6.

[4] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Determining image origin and integrity using sensor noise," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 74–90, Mar. 2008.

[5] M. Chen, J. Fridrich, and M. Goljan, "Digital imaging sensor identification (further study)," in *Proc. SPIE, Security, Forensics, Steganography, and Watermarking of Multimedia Content X*, Feb. 2007, vol. 6505.

[6] F. Pérez-González, M. Masciopinto, I. González-Iglesias, and P. Comesaña, "Fast sequential forensic detection of camera fingerprint," in *IEEE Int. Conf. on Im. Proc. (ICIP)*, Sep. 2016, pp. 3902–3906.

[7] M. Goljan, J. Fridrich, and T. Filler, "Large scale test of sensor fingerprint camera identification," in *Proc. SPIE, Electronic Imaging, Media Forensics and Security XI*, Feb. 2009, vol. 7254, pp. 0I 1-0I 12.

[8] M. Goljan and J. Fridrich, "Camera identification from cropped and scaled images," in *Proc. SPIE, Security, Forensics, Steganography, and Watermarking of Multimedia Content X*, Mar. 2008, vol. 6819.

[9] Miroslav Goljan, "Digital camera identification from images–estimating false acceptance probability," in *International Workshop on Digital Watermarking, Busan, Korea*. Springer, Nov. 2008, pp. 454–468.

[10] J. Fridrich and M. Goljan, "Derivation of ROCs for composite fingerprints and sequential trimming," *Technical Report, Binghamton University*, Jan. 2010.

[11] M. Masciopinto and F. Pérez-González, "A mathematical analysis of PRNU detection," In Preparation.

[12] G. E. Healy and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 3, pp. 267–276, Mar. 1994.

[13] M. K. Mıhçak, I. Kozintsev, and K. Ramchandran, "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Sig. Proc.*, Mar. 1999, vol. 6, pp. 3253–3256.

[14] Thomas Gloe and Rainer Böhme, "The 'Dresden Image Database' for benchmarking digital image forensics," in *Proc. of the 25th Symp. On Applied Computing (ACM SAC 2010)*, Mar. 2010, vol. 2, pp. 1585–1591.

[15] D.T. Dang-Nguyen, C. Pasquini, V. Conotter and G. Boato, "RAISE – A raw images dataset for digital image forensics," in *Proc. 6th ACM Multimedia Systems Conference*, Mar. 2015, pp. 219–224.