Chained Compressed Sensing: A Blockchain-Inspired Approach for Low-Cost Security in IoT Sensing

Chained Compressed Sensing: A Block-Chain-inspired Approach for Low-cost Security in IoT Sensing

Mauro Mangia, *Member, IEEE,*, Alex Marchioni, *Student Member, IEEE,*, Fabio Pareschi, *Member, IEEE,*, Riccardo Rovatti, *Fellow, IEEE,*, Gianluca Setti, *Fellow, IEEE,*

***Abstract*—Chaining, i.e., the mode of operation in which each message is encrypted considering a digital summary of previous ones, is here applied to block-cipher stages based on compressed sensing. We show that this simple and parsimonious technique may significantly harden the resulting system with respect to common threats such that ciphertext-only, known-plaintext, and man-in-the-middle attacks. Non-negligible robustness comes at the price of not more than a 2% of energy overhead with respect to the pure compression stage which represents a 24× reduction with respect to straightforward implementation of a traditional cryptography primitive like AES.**

I. Introduction

The spreading adoption of the Internet of Things (IoT) paradigm in areas like smart cities [1], healthcare systems [2] and smart homes [3], and the consequent need to provide network access to a vast multitude of sensing nodes with minimal energy footprint, has resulted in an increasing attention to the need of guaranteeing the privacy of data gathered and distributed by networked devices [4]–[9]. Security is even of greater concern when the sensor nodes acquire sensitive biometric information or biomedical signals, e.g. for authentication purposes or remote health monitoring applications. The issue is also greatly complicated by the fact that IoT nodes are of low-complexity and extremely low-power by design and every resource, including those spent for security purposes, must by carefully tailored to the actual requirements of each application.

At the current state of the art, IoT security is also challenging due to lack of standardization in addressing the problem to the point that a large part of off-the-shelf IoT nodes lack fundamental considerations in terms of privacy [10]. Furthermore, whenever addressed, security is commonly granted via dedicated encryption stages with varying levels of complexity. These stages protect information after analog-to-digital conversion of the signal of interest, and need considerable resources, especially in terms of power consumption and implementation costs. Given the resource optimization needed in IoT nodes, methods to match this expense with the actual amount of security that is needed in each case are therefore desirable.

Recent surveys [11]–[13] associate security problems to different abstraction layers and, at each layer, address the challenge of devising new cryptographic primitives able to effectively address the security-energy trade-off.

Compressed Sensing (CS) [14], [15] is a signal acquisition technique embedding implicit compression in a so-called analog-to-information conversion [16] [17], and has been proposed as a method to also introduce security directly into the acquisition process at the analog-to-information interface or jointly with digital signal compression [18]–[20].

In rough terms, what happens in CS is that chunks of an input waveform are represented with fewer scalars than the number of samples indicated by the Nyquist-Shannon theorem, which makes CS very appealing for low-resources IoT nodes. Such a lower-resource acquisition is possible assuming that the signal to process is *sparse*, i.e. a proper basis exists such that the projection of any input waveform over that basis has only few terms significantly different from zero. Acquisition (encoding) is practically achieved via a random linear projection by means of a suitable sensing matrix, whose perfect knowledge is fundamental to reconstruct the original signal via a non-linear decoding algorithm [15]. Such a matrix can therefore be considered a sort of key [18], [20], which, once shared between the IoT node and the corresponding gateway, guarantees a certain degree of secrecy *without the need of any additional cryptographic stage*, as well as robustness again Known-Plaintext Attacks (KPAs) [19]. Unfortunately, due to its linearity, CS cannot provide *perfect secrecy* in the Shannon sense, since the information about the power of the acquired signal leaks into the compressed measurements that are transmitted to the receiver.

Another recent direction in IoT security is the exploitation of BlockChain (BC) techniques [21]–[25]. BC, which is, for example, the cornerstone of Bitcoin (the first cryptocurrency system launched in 2008 [26]), involves the creation of a public and distributed ledger by appending (mining) blocks to it when a transaction is performed. Each block contains a digital summary of the previous block and, once linked to

M. Mangia, A. Marchioni and R. Rovatti are with the Department of Electrical, Electronic and Information Engineering (DEI), University of Bologna, 40125 Bologna, Italy (e-mail: {mauro.mangia2, alex.marchioni, riccardo.rovatti@unibo.it}). R. Rovatti is also with the Advanced Research Center on Electronic Systems (ARCES), University of Bologna, 40125 Bologna, Italy.

F. Pareschi and G. Setti are with the Department of Electronics and Telecommunication (DET), Politecnico di Torino, 10129 Torino, Italy, and also with the Advanced Research Center on Electronic Systems (ARCES), University of Bologna, 40125 Bologna, Italy (e-mail: {fabio.pareschi, gianluca.setti}@polito.it).

the chain, makes impossible to change the payload of any single block without recomputing all subsequent blocks. This chaining is one of the features that contributed to the overall security of BC.

Yet, computing a block and linking it to the chain requires the solution of a cryptographic challenge called Proof-Of-Work (POW). As such, despite promising, a straightforward application of BC to distributed IoT nodes is challenging for the large resources and long delay associated to POW operations. Solutions to this problem have been proposed in [23], [24] which rely on the elimination of the concept of POW by creating a local private BC between the IoT nodes and the corresponding gateway.

The main contribution of this paper is to take inspiration from the BC technology to improve the privacy level of CS acquisition and therefore to create a solution for secure data transmission between an IoT node and the corresponding gateway. To do so, we introduce a suitable *chainer* block, pair it with CS and exploit its properties to mask the power signature of the signals, to increase resistance to KPAs, and to grant robustness to Man-in-the-Middle Attacks (MiTMs), that in some critical applications may be a key issue.

The aim of this work is to propose chained CS as a new cryptography primitive that provides data privacy coping with the strict energy constraints typical of IoT applications. Note that CS processes successive signal chunks so we will compare our CS based methods with standard block cipher approaches.

More specifically, the paper is organized as follows. In Section II we briefly review the essentials of CS and its applicability as a block cipher stage. In Section III we identify the actors in our secure transmission and the potential attacks we deal with: Ciphertext-Only Attacks (COAs), KPAs and MiTMs. In Section IV we define a *chainer* as a block that acts on subsequent vectors of digital words with a mixture of hashing and arithmetic processing. We also state two properties that we leverage on in our manuscript, whose proof is sketched in the Appendix. In Section V we apply a chainer at the output of a CS stage and derive the security properties of the resulting ensemble. In Section VI we apply a chainer before a CS stage and derive the security properties of the resulting ensemble. In Section VII we discuss the approaches introduced in the previous two sections in terms of energy overhead with respect to the standard CS and comparisons with standard block cipher has been also provided. In Section VIII we show an example of how a chainer before CS can be used to improve robustness against MiTMs in a potentially life-critical application in which Electro Cardio Graphics signals (ECGs) are acquired and transmitted. Conclusions are drawn at the end.

II. Compressed Sensing and Block Ciphers

We work in a discrete-time setting in which the signal waveform is acquired as a sequence of time windows. We assume that the t-th of them starts at discrete instant t, and that in each window the signal is sampled n times, and we also collect in a vector $\xi[t] = (\xi[t]_0, \ldots, \xi[t]_{n-1})$ the corresponding digital entries.

TABLE I
LIST OF ACRONYMS

Basis Pursuit	BP
BlockChain	BC
Ciphertext-Only Attack	COA
Compressed Sensing	CS
Electro Cardio Graphic Signal	ECG
Internet of Things	IoT
Known-Plaintext Attack	KPA
Man-in-the-Middle Attack	MiTM
Pseudo-Random Number Generator	PRNG
Signal-to-Noise Ratio	SNR
Subset Sum Problem	SSP

To indicate the ranges of digital quantities, given a number B of bits, it is useful to define the sets of integers $\mathbb{Z}(B) = \left\{-2^{B-1}, \ldots, 2^{B-1}-1\right\}$ and $\mathbb{N}(B) = \left\{0, \ldots, 2^B-1\right\}$. With this, if B_ξ bits are used to encode each signed sample, we have $\xi[t] \in \mathbb{Z}\left(B_\xi\right)^n$.

CS works on sparse signals, i.e., it assumes that among the n entries of $\xi[t]$ at most $\kappa \ll n$ are non-zero [14]. Under this assumption, CS multiplies $\xi[t]$ by an $m \times n$ matrix $A[t]$ to obtain a so-called measurement vector $y[t] = A[t]\xi[t]$. A typical, very hardware-friendly option for the elements of $A[t]$ is a random choice of independent antipodal ± 1 symbols generated by a Pseudo-Random Number Generator (PRNG) with a known seed [27].

Despite the fact the $y[t]$ is only m-dimensional, a decoder is able to recover the original signal $\xi[t]$. In fact, the recovery procedure exploits the sparsity prior to select, among the infinite number of n-dimensional vectors $\xi[t]$ compatible with the values in $y[t]$, the sparsest one $\hat{\xi}[t]$. More precisely, it can be proved [14] that under suitable assumptions this may be done by setting $\hat{\xi}[t]$ to the solution of the convex optimization problem

$$\hat{\xi}[t] = \arg\min_{\xi} \|\xi\|_1 \ \text{ s.t. } A[t]\xi = y[t] \tag{1}$$

where $\|\cdot\|_1$ is the ℓ_1-norm, i.e., $\|\xi\|_1 = \sum_{j=0}^{n-1} |\xi_j|$. This method is called Basis Pursuit (BP) and, though it is not by far the unique proposed in the literature, it is the prototype of most non-greedy recovery algorithms commonly employed in CS based acquisition schemes (see f.i. [28] and references therein).

It is worth noting that, (1) being convex (to the point that it can be translated into a purely linear optimization), it admits a unique solution so that, since $\xi[t]$ has integer entries, then $y[t] = A[t]\xi[t]$ also has integer entries and theoretically allows errorless recovery, that is $\hat{\xi}[t] = \xi[t]$.

Theory ensures that this happens when m is $O\left(\kappa \log\left(n/\kappa\right)\right)$ [14], [15]. In any practical case one usually has $m \ll n$ and thus, assuming that the data needed to recover the signal must be transmitted or stored, the obvious advantage of CS is that *a simple linear transformation yields acquisition with*

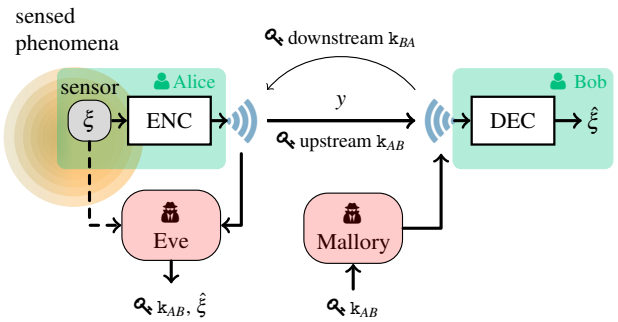

Fig. 1. Links and some of the possible attacks in a distributed sensing environment.

compression. Actually, literature abounds of methods [29]–[33] to lower m and thus increase the achievable compression and some of them [34], [35] are compatible with the antipodal structure of $A[t]$.

Further to that, one may note that, to be able to recover the original signal, the decoder must know $A[t]$, i.e., the seed from which the PRNG generating it is run. Hence, that seed plays the role of the private key of a block cipher that uses it to encrypt the *plaintext* $\xi[t]$ into the *ciphertext* $y[t]$. Such a ciphertext is then passed to a decoder that uses the same key to solve (1) and recover the plaintext.

From this point of view, CS has been proposed and analyzed as a way of simultaneously providing compression and some form of security [18]–[20], [36] at an extremely low-cost, thus perfectly fitting in the design of low-resources subsystems dedicated to sensing.

Characterization of the trade-off between security and resources is classically carried out by analyzing cryptographic attacks and quantifying the effort needed by the attacker to have a non-negligible probability of success, as it will be detailed in the following sections.

III. A PROTECTED COMMUNICATION SCHEME BETWEEN SENSORS AND GATEWAYS

The conventional setting sees a legitimate transmitter (Alice) who encodes a plaintext into a ciphertext and sends it to a legitimate receiver (Bob). Various attackers try to exploit the information that leaks from the encryption, possibly leveraging additional side information that depends on the kind of attack. In our setting, Alice may represent the sensing subsystem of any IoT device and Bob is the gateway collecting readings and dispatching them for further processing. An intuitive scheme of the links and the possible attacks is reported in Figure 1.

By means, for example, of a Diffie-Hellman-Merkle protocol [37], Alice and Bob agree on two keys $\mathtt{k}_{AB}$ and $\mathtt{k}_{BA}$ that control, respectively, the *upstream* communication between Alice and Bob and the *downstream* from Bob to Alice. The two links are used in a different way since Alice sends sensor readings to Bob while Bob sends control words to Alice. We analyze attacks on the upstream link that carry sensor readings. In particular, we focus on three different classes of attacks:

- **Ciphertext-Only Attack (COA)**: this is the most straightforward attack to the data sent by Alice to Bob. Here an eavesdropper (Eve) collects successive ciphertexts to evaluate the statistics of transmitted symbols and tries to guess pieces of information about the plaintext. In this case, Eve embodies the gateway (Bob), i.e, she is not able to access to the sensor readings in any case.
- **Known-Plaintext Attack (KPA)**: this class of attacks is more insidious with respect to COAs. In KPA the attacker, Eve, possesses the ability to capture a certain number of plaintext-ciphertext pairs and tries to use them to compute the $\mathtt{k}_{AB}$ which will allow her to decode the following ciphertexts. In our IoT inspired scenario performs this attack is easy. It is reasonable to suppose that Eve may be able to temporarily deploy an identical node close to the attacked one, i.e, she is able to acquire the signal $\xi[t]$ starting from a certain time $t > 0$. She is also eavesdropping the transmitted ciphertexts, therefore Eve embodies both Bob and she partially embodies Alice[1].
- **Man-in-the-Middle Attack (MiTM)**: A different attacker (Mallory), interposes between Alice and Bob and sends messages to the latter pretending to be the former. This is called Man-in-the-Middle Attack. To be able to communicate to Bob, Mallory knows the upstream key $\mathtt{k}_{AB}$. The most threatening aspect of MiTM is that Bob may receive counterfeited information and this can be critical when dealing with sensors that produce sensitive information.

Though in this paper we focus on the uplink, note that MiTM attacks are critical also on the downlink from Bob to Alice as this would allow Mallory either to set the internal parameter of the encryption to values known to attackers, or misconfigure the sensor nodes, or even give forged instructions to the node that may, for example, host also actuators along with sensors.

The most classical way of mapping CS onto a block cipher is to use $\mathtt{k}_{AB}$ as a seed for the PRNG that generates $A[t]$ both at the encoder and at the decoder. This is the scheme that most of the current literature discusses [18]–[20], [36].

From those contributions we know that the trade-off between resources and security has consequences on the robustness against COAs. In particular, in the large n limit, the statistical distribution of the entries of $y[t]$ becomes Gaussian with a zero average and a variance that depends on the energy of the encoded signal $\sum_{j=0}^{n-1} \xi[t]_j^2$ [18], [20]. This means that Eve may observe the measurement vector that is the ciphertext and extract a small but non-negligible amount of information on the signal that is the plaintext.

Robustness with respect to KPAs is also not complete. In particular, given $\xi[t]$ and $y[t]$, the task of finding a matrix $A[t]$ made of antipodal symbols $A[t]_{j,k} = \pm 1$ that is compatible with $y[t] = A[t]\xi[t]$ is quite easy. In this case, security stems from the fact that the number of candidate solutions $A[t]$ can be made so large that pinpointing the true matrix is practically impossible [19].

[1] In KPA the attacker accesses to the information to be encrypted but he does not posses the ability to impose the current value of the data to be encrypted. The case where Eve is also able to chose the $\xi[t]$ values correspond to another class of attacks, the Chosen Plaintext Attack

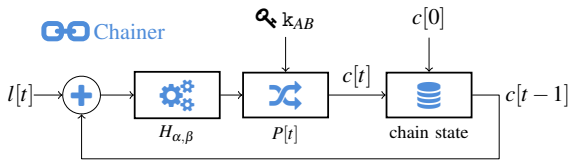

Fig. 2. Scheme of the chainer block.

The main contribution of this work is to introduce a chaining technique that, by means of a small additional complexity, allows to increase robustness against COAs and KPAs and helps countering MiTM attempts.

These attacks are selected as they present themselves as obvious threats in the IoT framework when nodes cannot be tampered with but their immediate surrounding can be physically accessed thus making them the weakest links in the upstream chain. Clearly, other attacks like Chosen Plaintext and Chosen Ciphertext attacks are possible and will be the topic of future investigations.

IV. Window Chaining

Chaining is a classical mode of operation of block ciphers. To apply it in a framework in which CS acts as the block cipher primitive, it is convenient to define chaining on vectors of digital words instead of blocks of bits.

Formally speaking, we assume that $c[t] \in \mathbb{N}(B_c)^{n_c}$ is the n_c-dimensional chain state vector at time t, and that $l[t] \in \mathbb{Z}(B_l)^{n_c}$ is the vector representing the link to be added to the current chain with $B_l \leq B_c$. The next chain state is computed as

$$c[t] = P[\mathtt{k}_{AB}]\left(H_{\alpha,\beta}\left(c[t-1] + l[t]\right)\right)$$

In that formula, $H_{\alpha,\beta} : \mathbb{N}(B_c)^{n_c} \mapsto \mathbb{N}(B_c)^{n_c}$ is a component-wise linear congruential hashing function such that, if $c'' = H_{\alpha,\beta}(c')$ then

$$c''_j = \left(\alpha c'_j + \beta\right) \mod 2^{B_c}$$

for $j = 0, \ldots, n_c - 1$, where the modulus is defined to always yield values in $\mathbb{N}(B_c)$ and we set $\beta = 1$ and $\alpha = 2^{B_c} - 3$ to keep implementation as simple as possible but still guarantee that the congruential mapping preserves maximum length cycles [38].

Moreover, $P[\mathtt{k}_{AB}] : \mathbb{Z}(B_c)^{n_c} \mapsto \mathbb{Z}(B_c)^{n_c}$ is a permutation network that swaps the entries of its argument according to a permutation that is pseudo-randomly generated spreading from $\mathtt{k}_{AB}$.

Figure 2 reports a scheme of a chainer block. Such a block features two properties that we will exploit in the following, i.e.,

1) If $P[\mathtt{k}_{AB}]$ and the sequence and $c[t]$ are known then one can infer the sequence $l[t]$;
2) Under mild conditions on the statistics of $l[t]$, the entries of $c[t]$ tend to be uniformly distributed in $\mathbb{N}(B_c)$ for $t \to \infty$.

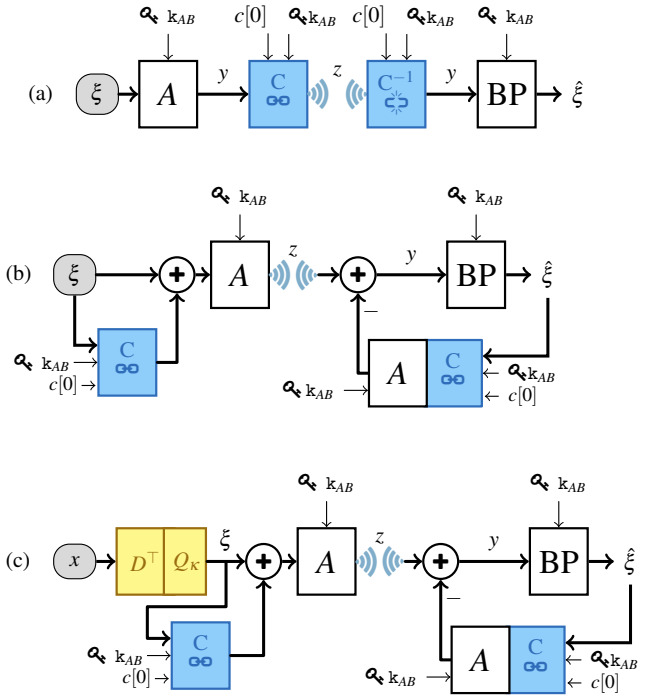

Fig. 3. Using a chainer to improve CS security: (a) chaining measurements; (b) chaining sparse inputs; (c) chaining after sparsification.

A sketch of the proof of the two properties, along with the required conditions, is reported in the Appendix.

By exploiting the chaining block, we extend the use of CS as a block cipher in two ways that are sketched in Figure 3-(a) and -(b) (Figure 3-(c) is a variation on Figure 3-(b) needed to cope with some real-world signal that will be explained and exploited in Section VIII).

As a final comment, it is interesting to evaluate what is the overhead in terms of complexity due to the introduction of a window chaining in a encryption/compression scheme based on the CS paradigm. According to the applications proposed in the next two sections, the n_c value is at maximum equal to n. Given that, the computational cost for the chainer block can be approximated with a number of arithmetic operations running in $\mathcal{O}(n)$ and $P[\mathtt{k}_{AB}]$.

Conversely, the computational of a CS encoder is proportional to nm. The evaluation of the measurement vector requires nm sums as well as the number of random bits to be generated for one antipodal sensing matrix is nm. Knowing that $m \approx \mathcal{O}(\alpha n \log(1/\alpha))$ with $\alpha = \kappa/n$, the CS encoder cost is $\mathcal{O}(\alpha n^2 \log(1/\alpha))$.

Thus, it is reasonable to affirm that the introduced overhead is limited, i.e., the whole computational cost for the proposed encryption/compression scheme does not drastically increase with respect to the case where only a CS encoder is devoted to such task.

V. Chaining of Measurements

As a first option, we may apply chaining to the sequence of measurement vectors by setting $l[t] = y[t]$ and using the

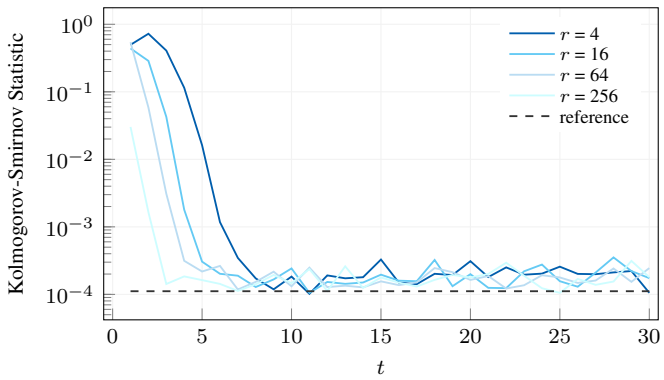

Fig. 4. Convergence of ciphertext statistics to the uniform distribution.

chain state as the ciphertext sent to Bob by setting transmitted signal $z[t]$ equal to the chainer output $c[t]$. This is sketched in Figure 3-(a).

The initial condition $c[0]$ is set by Bob and sent to Alice through the downstream link encrypted with $\mathtt{k}_{BA}$. Initialization can be a regular process associated to all other forms of periodic synchronization that normally take place between gateways and sensor nodes.

The transfer of information from Alice to Bob depends on property 1 in Section IV. In fact, at each $t > 0$, Bob observes $c[t]$ and remembers $c[t-1]$ from the previous time step. Knowing $P[\mathtt{k}_{AB}]$, it is possible to compute $y[t]$ that is fed into a BP reconstruction block along with $A[t]$ that is also known starting from $\mathtt{k}_{AB}$. Providing m is properly decided, this allows errorless recovery of $\hat{\xi[t]} = \xi[t]$.

Eve may also observe $z[t] = c[t]$ and remember $z[t-1] = c[t-1]$, but cannot reconstruct $P[\mathtt{k}_{AB}]$ and thus cannot decide which entry in $c[t]$ must be matched with which entry in $c[t-1]$ to compute the corresponding entry of $l[t]$. With no other information, Eve would need to guess the permutation $P[\mathtt{k}_{AB}]$, i.e., choose one out of $m!$ equally probable options. Since m is commonly in the order of many tens if not hundreds, this is feasible only with a substantial computational effort. Moreover, whenever $l[t] = y[t]$ could be retrieved, Eve would still need to break CS-based encoding to gain information on $\xi[t]$.

Even more interesting than adding a further layer of complexity to the encryption, chaining of measurements prevents the leakage of information about the energy of the signal $\xi[t]$ that affects classical CS-based schemes, i.e., COA immunity is strongly reinforced. In fact, thanks to property 2 in Section IV, the distribution of $z[t] = c[t]$ is asymptotically uniform.

Though the proof of that property does not give hints on the speed of convergence to the uniform behavior (technically speaking, it ensures convergence by proving that the trend is faster than the slowest possible one), we may obtain an idea of the typical speed by performing some simulation. Our test system encodes each measurement $y[t]_j$ with 10 bits, thus setting $B_c = 10$. We know that $y[t]_j$ is distributed as discrete version of a Gaussian distribution and to consider different levels of energy we scale it by constraining $y[t]_j \in \{-r, \ldots, r\}$.

A thorough Montecarlo simulation allows to accumulate samples of the chained measurements form which we calculate empirical distributions at subsequent time steps t. Each empirical distribution is matched against the uniform one by computing its Kolmogorov-Smirnov statistic [39, Chapter 15]. The result is reported in Figure 4 along with the value of the Kolmogorov-Smirnov statistic of a set of samples generated by a truly uniform distribution (dashed line). Notice how convergence is so fast that after 10 time steps the ciphertext distribution is practically indistinguishable from a uniform one.

VI. Chaining of Sparse Signals

As a second option, we may apply chaining to the sequence of sparse signals by setting $l[t] = \xi[t]$. Then, we may use the chain state to perturb $\xi[t]$ before it enters the CS stage. The ciphertext is then $z[t] = A[t](\xi[t]+c[t-1]) = y[t]+A[t]c[t-1]$. This is sketched in Figure 3-(b).

The transfer of information from Alice to Bob relies on the fact that Bob is continuously receiving data from Alice and thus is able to update its local copy of the chain state, starting from the initial condition $c[0]$. As in the previous case, the value of $c[0]$ is to be considered a message sent to Alice through the downstream link encrypted with $\mathtt{k}_{BA}$.

Chaining of sparse signals increases the robustness of CS-based encryption with respect to KPAs and grants some immunity to MiTM attacks.

In KPAs, Eve knows both $\xi[t]$ and $z[t]$ at T time steps $t_0, \ldots, t_{T-1}$. To compute the corresponding $A[t_0], \ldots, A[t_{T-1}]$ and hope to identify the $\mathtt{k}_{AB}$ that regulates their generation, Eve should solve

$$z[t] = A[t]\left(\xi[t] + c[t-1]\right) \tag{2}$$

for $t = t_0, \ldots, t_{T-1}$.

If Eve knew $c[t_j]$, then the result in [19], [40] would imply that solving (2) is extremely easy but generates such a deluge of indistinguishable candidates $A[t_j]$ to completely spoil the effectiveness of the attack. When chaining on the sparse signal enters into play, $c[t_j - 1]$ is hidden, and the right-hand side of (2) contains two unknowns, $A[t_j]$ and $c[t_j - 1]$. Since by property 2 in Section IV, we have no prior on $c[t_j - 1]$, all possible 2^{nB_c} candidates are feasible and for each of them, the resulting equation in $A[t_j]$ would yield an enormous number of indistinguishable solutions.

This considered, and using the result in [19] to estimate the number of equivalent solutions to (2) when $c[t_j - 1]$ is known, we obtain that on the average, the number of equivalent solutions to the complete version of (2) is $2^{n-\max\{B_\xi, B_c\}-1+nB_c}\sqrt{3/\pi n}$. The security level offered against KPA can, thus, be roughly estimated as the minimum between 2-logarithm of that number, and the number of bits used to encode $\mathtt{k}_{AB}$ and $\mathtt{k}_{BA}$ (that protects $c[0]$), i.e., $\min\{(n-1)(B_c+1), B_{\mathtt{k}_{AB}} + B_{\mathtt{k}_{BA}}\}$-bits, which is equal to $B_{\mathtt{k}_{AB}} + B_{\mathtt{k}_{BA}}$ for n large enough.

In MiTM attacks, Mallory is interested in sending messages to Bob pretending to be Alice, exploiting the fact that she knows the upstream key $\mathtt{k}_{AB}$. At any time step t, Bob expects

a ciphertext that depends not only on the plaintext (that Mallory wants to counterfeit) but also on a chain state that Mallory does not know.

Hence, to fool Bob into accepting her message at time t, Mallory must first reconstruct the chain state $c[t-1]$ and thus must solve (2) where $z[t]$ is observed and $A[t]$ is computed from $\mathtt{k}_{AB}$. The attack must go through two steps. It must first compute $c[t-1]+\xi[t]$ starting from (2), and then separate $c[t-1]+\xi[t]$ into $c[t-1]$ and $\xi[t]$. Both steps generate a lot of indistinguishable candidates whose number is a quantitative evaluation of the offered security.

By dropping the fixed t we may concentrate on solving $q = Ap$ with $q = y[t]+2^{B_\xi}A[t]u$ and $p = c[t-1]+\xi[t]+2^{B_\xi}u$, where u is the n-dimensional unit vector. The unknown vector p is made of non-negative entries and thus there is a set of $n \times B_p$ binary variables $b_{k,l}$ such that $p_k = \sum_{l=0}^{B_p-1} b_{k,l}2^l$ with $B_p = \max\{B_c, B_\xi\} + 1$. By substituting this into the expression for the j-th component of q we get

$$q_j = \sum_{k=0}^{n-1} A_{j,k}p_k = \sum_{k=0}^{n-1}\sum_{l=0}^{B_p-1} A_{j,k}2^k b_{k,l} \qquad (3)$$

for $j = 0, \ldots, m-1$.

This shows that the first step of MiTM attacks is equivalent to a m simultaneous Subset Sum Problems (SSPs) with variables $b_{k,l}$ and coefficients $A_{j,k}2^k$.

SSP is known to be an NP-complete class of problems potentially very hard to solve [41]. Yet, the reason why Mallory has little chance of being successful in the first step lies in the same phenomenon that regulates KPAs [19], [40]. SSPs are characterized by a parameter δ called density that is the ratio between the number of variables and the number of bits needed to encode the corresponding coefficients. In our case

$$\delta = \frac{nB_p}{\log_2(2^{B_p})} = n \gg 1$$

so that (3) is a so-called high-density SSP. It is known that high-density SSPs are likely to admit many solutions [42]. Regrettably, due to the fact that the coefficients $A_{j,k}2^k$ do not obey the statistical assumptions on which classical theoretical quantification of the number of solutions is based, we may only resort to simulation.

Since counting solutions is a potentially hard problem by itself, we test only very small instances and observe how the number of solutions behaves when n, m and $B_c = B_\xi$ change. Our reference configuration features $n = 8$, $m = 6$, $\kappa = 3$ and $B_c = B_\xi = 5$. Starting from that we consider configurations in which either n, m, or $B_c = B_\xi$ decreases or increases. For each configuration we generate 1000 instances of an MiTM attack and enumerate all possible solutions to the first step needed by Mallory. The results are reported in Table II.

As expected, the number of solutions of an MiTM attack increases when either n or B_c and B_ξ increase. This is because in both cases the overall number of binary variables in the equivalent simultaneous SSP increases. On the contrary, increasing m reduces the number of solutions as the matrix $A[t]$ is $m \times n$ and increasing m makes it closer to a square one

TABLE II
AVERAGE NUMBER OF SOLUTIONS TO THE FIRST STEP OF AN MITM ATTACK FOR SOME SMALL-SIZE SYSTEMS TESTED TO ASSESS THE SENSITIVITY TO EACH PARAMETER. THE FIRST BOLD ROW REFERS TO A REFERENCE CASE WHOSE NEIGHBORHOOD IS EXPLORED BY THE OTHER ROWS. ARROWS HIGHLIGHT THE TREND WITH RESPECT TO THE REFERENCE CASE.

n	m	B_c=B_ξ	average # of solutions
8	**6**	**5**	**65584**
7 ↓	6	5	↓ 6198
9 ↑	6	5	↑ 408117
8	5 ↓	5	↑ 630472
8	7 ↑	5	↓ 7536
8	6	4 ↓	↓ 4992
8	6	6 ↑	↑ 150041

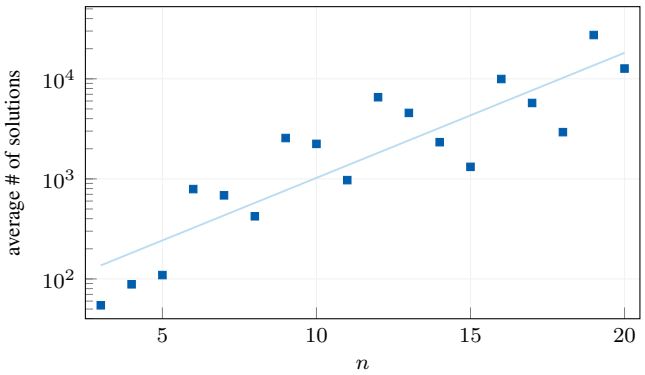

Fig. 5. Average of the number of solutions of an MiTM attack for small-size problems in which $m \simeq 70\%n$ and $\kappa \simeq 30\%n$ (integer approximations apply). The very small number of bits $B_c = B_\xi = 3$ allows exhaustive enumeration of the solutions for each Montecarlo trial. The line is the least-square fitting of the trend.

for which $z = Ap$ would be a one-to-one relationship. Note how, in every case, sensitivity to configuration parameters is extremely high.

Further to that, to be able to reasonably estimate the number of solutions for values of the system parameters in typical ranges, we simulate configurations with $4 \leq n \leq 20$ and $m \simeq 70\%n$ and $\kappa \simeq 30\%n$ that are ratios close to what commonly appears in real-world applications. Exhaustive enumeration of all solutions is possible if we adopt $B_c = B_\xi = 3$, that keeps the number of binary variables in the corresponding SSP reasonably low. A semilogarithmic plot of the results is in Figure 5, that also reports a linear interpolation of the increasing trend. The least-square fitting of the logarithm of the data yields that the average number of solutions is approximately $2^{0.416n+5.847}$.

Hence, assuming that the trends we observe for these small-size cases are maintained, when Mallory attacks a typical system in which n is of the order of few hundreds, the first step of the attack generates a number of indistinguishable candidates p and thus $c[t-1]+\xi[t]$ whose order of magnitude is of few tens.

Once the first step is over, splitting each of these candidates into a guess for $c[t-1]$ and the corresponding guess for $\xi[t]$ entails deciding which of the n components of $\xi[t]$ are non-

null. This can be done in $\binom{n}{\kappa} \simeq 2^{nG(\kappa/n)}$ ways, where $G(\epsilon) = -\epsilon \log_2(\epsilon) - (1-\epsilon)\log_2(\epsilon)$. With n in the hundreds, this adds another several tens to the order of magnitude of the number of equally plausible decisions.

Overall, in this particular case ($m = 70\%n$ and $\kappa = 30\%n$) and very small ($B_c = B_\xi = 3$) case, the security level against an MiTM attack can be estimated as the minimum of the 2-logarithm of the number of equivalent candidates and the number of bits used to encode $\mathtt{k}_{AB}$ and $\mathtt{k}_{BA}$ (that protects $c[0]$), i.e., $\min\{(0.416 + G(0.3))n, B_{\mathtt{k}_{AB}} + B_{\mathtt{k}_{BA}}\}$ which is equal to $B_{\mathtt{k}_{AB}} + B_{\mathtt{k}_{BA}}$ for n large enough.

Note that, though these estimations may be very rough. The linear trend in n of the number of equivalent bits quantifying the security level is a structural property of the encryption scheme that, for sufficiently large n, should be enough to discourage any low-resource MiTM attack.

VII. Energy overheads in node-oriented implementations

With the aim of providing an example of the possible energy costs, we report here some results on the implementation of the proposed approaches on a commercial microcontroller. A standard CS stage followed by a $128\,\mathrm{bit}$ Advanced Encryption Standard block (AES) [43] has been considered for comparison. We refer to the implementation proposed in [44].

Our setting uses sparse signals ξ with $n = 128$, $\kappa = 8$ and $B_\xi = 9$. The antipodal matrices A devoted to the compression task feature $m = 64$ and entries generated by a Linear Feedback Shift Register (LFSR). As a result, each element of y requires $B_y = B_\xi + \log_2(n) = 16\mathrm{bit}$. To be comparable with the AES that uses a $128\,\mathrm{bit}$ block size with a $128\,\mathrm{bit}$ key length, we implement a $128\,\mathrm{bit}$ LFSR in Galois configuration, i.e., $\mathtt{k}_{AB}$ represents both the initial state of the LFSR and the AES key.

To emulate the node implementation we use a resource-constrained TI EK-TM4C123GXL evaluation board [45], embedding a low-power low-cost ARM Cortex-M4F TM4C123GH6PMI microcontroller. With no peripherals enabled and for a fixed working frequency f_{clk}, the energy required to execute a task on this device can be estimated as $E = V_{\mathrm{dd}} I_{\mathrm{avg}} N_{\mathrm{clk}} / f_{\mathrm{clk}}$ where $V_{\mathrm{dd}} = 3.3\,\mathrm{V}$, $f_{\mathrm{clk}} = 16\,\mathrm{MHz}$, N_{clk} is the number of clock cycles, while, experimentally, we measure $I_{\mathrm{avg}} = 11.3\,\mathrm{mA}$. Even if this approach has some drawbacks (i.e., we are measuring the whole ARM consumption, including that for the clock generation and distribution), it is a good proxy of the energy required by different tasks.

As already discussed, the introduction of a window chaining increases the ability of a CS block to prevent KPA in an IoT framework. Nevertheless, the main contribution of this work is twofold: the hardening with respect to COA (see Section V) and the resistance to MiTM attacks (see Section VI). Energy requirements along with the corresponding encryption schemes using standard AES blocks are discussed in the following two subsections.

A. Energy requirements of COA-hardened schemes

In the proposed framework, COA is countered by adopting the configuration in Figure 3-(a) (indicated as CS + Chaining$^{(\mathrm{a})}$). The same capability can be obtained by applying an AES block in Electronic Codebook mode (ECB) to the measurement vector y, i.e., the CS stage output is divided into $128\,\mathrm{bit}$ length blocks, and each of them is encrypted separately using the same encryption key (indicated as CS + AES$^{(\mathrm{ECB})}$).

For each plaintext (the vector ξ), the two approaches share the cost associated to the measurement computation, i.e., $m\,n$ multiply-and-accumulate operations, and the generation of $m\,n$ random bits. The overhead for Chaining$^{(\mathrm{a})}$ consists in $2m$ sums, m multiplications[2] and m assignments due to the permutation.

For AES$^{(\mathrm{ECB})}$, the associated overhead is $mB_y/128 = 8$ times the energy required for the encryption of a single block.

Results are shown in the first two rows of Table III. Note first that the energy required for the computation of measurements, and thus for compression, dominates the chaining overhead that is just about 2% of the total, while AES in ECB mode is not negligible as it needs 37% of the compression energy. The improvement ratio between AES$^{(\mathrm{ECB})}$ overhead and Chaining$^{(\mathrm{a})}$ overhead is $63\times$.

B. Energy requirements of MiTM-hardened schemes

The robustness against MiTM attacks is given by the configuration in Figure 3-(b) (indicated as CS + Chaining$^{(\mathrm{b})}$) while for the AES block that encrypts the CS measurements we consider the Cipher Block Chaining (addressed as CS + AES$^{(\mathrm{CBC})}$). As discussed in Section VI, for each plaintext, before the CS stage, we evaluate an intermediate vector as the sum of the plaintext itself and the output of the window chaining. In a similar fashion, with the CBC mode, incoming data is XORed with the previous ciphertext before the AES block. The main difference between these approaches is that CS + Chaining$^{(\mathrm{b})}$ computes the matrix multiplication after the chaining stage, while CS + AES$^{(\mathrm{CBC})}$ applies the chaining to the measurement vector y. This case compared to ECB mode requires an additional $128\,\mathrm{bit}$ initialization vector to compute the XOR with the first vector y.

These approaches share the energy cost for matrix multiplication and thus compression, while the overheads differ from the previous case. For CS + Chaining$^{(\mathrm{b})}$ n additional sums are needed with respect to the previous case. Similarly the energy required by CS + AES$^{(\mathrm{CBC})}$ differs form CS + AES$^{(\mathrm{ECB})}$ only in the additional XOR operations.

In the last two rows of Table III we report the measured energy costs for both CS + Chaining$^{(\mathrm{b})}$ and CS + AES$^{(\mathrm{CBC})}$. As in the previous case the CS stage dominates the overall energy consumption. Observed overheads are 1.58% for the former and 37.8% for the latter. The improvement ratio is $24\times$.

C. Qualitative energy-security trade-off

As anticipated at the end of Section IV, energy overheads for both Chaining$^{(\mathrm{a})}$ and Chaining$^{(\mathrm{b})}$ are almost negligible

[2]energy cost for the module operation is negligible in the proposed setting.

TABLE III
ENERGY CONSUMPTION [μJ] FOR CS ONLY AND FOR THE OVERHEAD DUE TO CHAINING OR TO AES.

	CS	Overhead		Improvement w.r.t. AES
	[μJ]	[μJ]	percent	
CS + Chaining$^{(a)}$	345	2.01	0.58%	63×
CS + AES$^{(ECB)}$		128	37.0%	
CS + Chaining$^{(b)}$	345	5.44	1.58%	24×
CS + AES$^{(CBC)}$		131	37.8%	

TABLE IV
A QUALITATIVE VIEW OF THE SECURITY-ENERGY TRADE-OFF (THE LARGER THE NUMBER OF FILLED CIRCLES, THE BETTER)

	Immunity to COA	Immunity to KPA	Immunity to MiTM	Energy saving
CS + Chaining$^{(a)}$	●●○	●●○	○○○	●●●
CS + AES$^{(ECB)}$	●●●	●●●	○○○	●○○
CS + Chaining$^{(b)}$	●○○	●●○	●●○	●●●
CS + AES$^{(CBC)}$	●●●	●●●	●●●	●○○

with respect to cost for the standard CS encoders. This advantage comes at the price of a potentially reduced security with respect to the perfect secrecy obtained with AES blocks. Qualitative comparisons are in Table IV which proposes a high-level outline of the corner cases we analyzed.

- Chaining$^{(a)}$ in Section V increases COA robustness with respect to straightforward CS, making it closer to standard AES. Conversely, Chaining$^{(b)}$ in Section VI possesses the same robustness to COA of standard CS.
- For both Chaining$^{(a)}$ and Chaining$^{(b)}$, immunity to KPA, discussed in Section VI, is based on the same working principle of standard CS that grants non-negligible robustness though less than that of AES.
- Chaining$^{(b)}$ in Section VI grants MiTM immunity that standard CS does not provide.

VIII. MITM-RESISTANT COMPRESSION OF ECGS

Real-world signals posses a sparse representation when expressed through a proper sparsity basis. In this case, the block scheme of Figure 3-(b) should be replaced with that in Figure 3-(c). The input signal, split into a sequence of time windows $x[t]$, needs to be processed by a *sparsification* block in order to generate a sequence of sparse vectors $\xi[t]$ to be processed as described in Section VI.

Being $x[t]$ composed of N samples, the sparsification block makes a first compression of $x[t]$ by projecting it on its N-size sparsity basis D' and considering only $n < N$ elements. This can be simply achieved by statistically estimating the energy associated, for the considered class of input signal, to the different projections along the columns of D'. By collecting the associated columns of D' in a matrix D, this operation can be simply described as multiplying $x[t]$ by the $n \times N$ matrix $D^\top$ as shown Figure 3-(c). A further block Q_κ provides

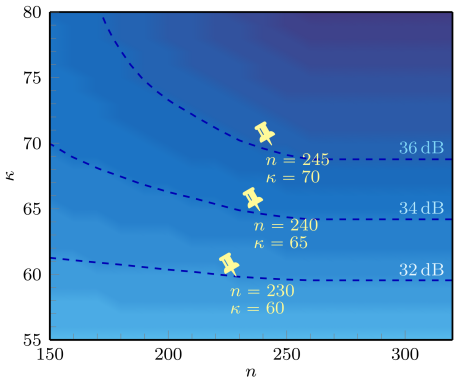

Fig. 6. Performance of the sparsification block in the ECG example as a function of n and κ.

TABLE V
PERFORMANCE OF THE PROPOSED SYSTEM FOR ECG COMPRESSION IN DIFFERENT CONFIGURATIONS. FOR ALL CASES, D' IS THE SYMLET-6 BASIS WITH $N = 512$.

Target SNR	n	κ	m	Compression ratio		
				Sparsification block	CS	overall
36 dB	245	70	182	2.09	1.35	2.81
34 dB	240	65	175	2.13	1.37	2.93
32 dB	230	60	164	2.23	1.40	3.12

quantization and allows to ensure that the sequence of vectors $\xi[t]$ has digital entries.

The Q_κ also forces sparsity in $\xi[t]$ by zeroing all elements except the κ with the largest magnitude. This step is necessary since real world signals commonly present at most $\kappa \ll n$ elements that are significant while all others are negligible but non-zero, thus not matching the exact mathematical definition of sparsity we leverage on. By forcing sparsity, the considered Q_κ ensures that the above schemes work properly.

Removal of the less energetic components in the D' expansion, sparsification, and quantization of $\xi[t]$, however, introduce an error in the signal chain, as $D\xi[t]$ only approximates $x[t]$. Performance in terms of signal-to-noise ratio (SNR) when considering $x[t]$ as composed by windows of a synthetic ECG generated accordingly to [46] has been plotted in Figure 6 as a function of n and κ. We set $N = 512$ and D' to the orthonormal Symlet-6 wavelet basis [47], with Q_κ a 10-bit quantization function. The sampling rate of the ECG is $256\,\mathrm{Sa/s}$. Figure 6 focuses on values around $34\,\mathrm{dB}$ that are commonly considered enough for a medical-grade ECG signal [48]. The compression ratio introduced can be quantified by N/n, since the number of coefficients required to represent the signal decreases from N to n.

The $\xi[t]$ obtained are further processed according to the block scheme in Figure 3-(c). The CS chain introduces a further signal compression quantified by n/m. Differently from the case of the sparsification, this is a lossless compression, being the $\xi[t]$ composed by digital quantities that are exactly recovered by the reconstruction algorithm, i.e., $\hat{\xi}[t] = \xi[t]$.

Three cases are pinpointed in Figure 6, corresponding

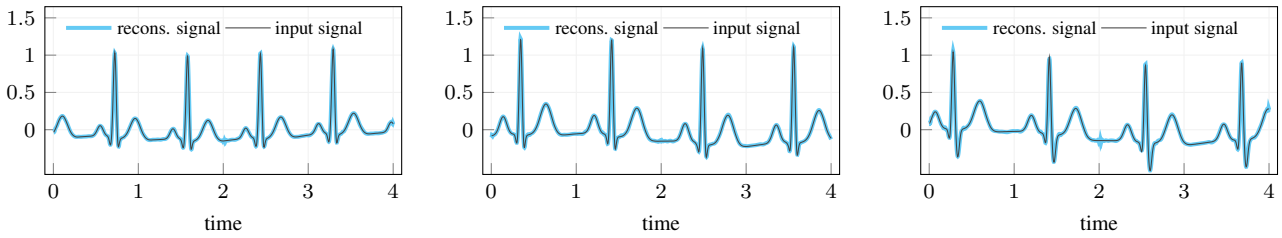

Fig. 7. Reconstruction of ECG signals in the examples of Section VIII. The obtained SNR is (a): 36.2 dB; (b): 34.7 dB; (b): 32.6 dB.

to three target qualities of 36 dB, 34 dB and 32 dB. The corresponding parameters are summarized in Table V, along with the value of m ensuring that the CS reconstruction stage is always capable of perfectly recovering $\xi[t]$, that has been experimentally identified by looking for the minimum value ensuring perfect reconstruction in 10000 subsequent time windows in a Montecarlo simulation. An example of the waveforms obtained by the reconstruction algorithm is depicted in Figure 7.

Note that in an application like this, rejection of MiTM attacks is a major issue. In fact, were she successful, Mallory could alter what is received by the health monitoring system Bob, causing it to believe that a patient Alice is in normal conditions while she is experiencing, for example, severe arrhythmia. Luckily enough, following the estimations of Section VI we may anticipate that any MiTM attack will produce a number of indistinguishable candidate chain states $c[t-1]$ in excess of $2^{1.3n}$, that is 1.02×10^{90} for the less accurate system in Table V and 7.56×10^{95} for the most accurate one.

IX. Conclusion

The adoption of a chainer as defined in Section IV after or before a CS stage increases security of the resulting two-stages system. In fact, further to low-resources compression, the resulting ensemble is able to secure transmitted against COAs, KPAs and MiTM attacks. As far as COAs are concerned, chaining is able to mask the power signature of the signals that is known to leak when only CS is employed as a block cipher. KPAs also become more difficult as a chainer increases the number of equally probable candidates for the key that the attacker wants to infer from each plaintext-ciphertext pair. This may be important as sensing systems are intrinsically prone to KPAs, being the plaintext potentially available to anyone deploying a sensor on the same physical phenomenon. With respect to a standard AES block the proposed approach guarantees a $63\times$ overhead improvement with an energy increase of just 0.58% of the total for the compression/encryption block. A chainer before CS is also able to harden the system against MiTM attacks in which an attacker knows the key used to secure communications from the transmitter to the receiver and wants to send fake messages to the latter. This may be a key issue in systems dealing with critical data like the one acquiring ECGs that is used as a practical example. In this case, we observe a $24\times$ reduction of energy overhead that is limited to 1.58% of the total energy cost.

Appendix - Properties of the Chaining Block

As far as property 1 in Section IV is concerned, note that we may exploit the fact that $P[\mathtt{k}_{AB}]$ is known to write

$$P[\mathtt{k}_{AB}]^{-1}\left(c[t]\right) = \alpha\left(c[t-1] + l[t]\right) + \beta \mod 2^{B_c}$$

where the whole left-hand side is known. Referring to the generic j-th component of the above vector equality, assume that $c^* \in \mathbb{N}(B_c)$ is the solution to the linear congruential equation $\left[P[\mathtt{k}_{AB}]^{-1}\left(c[t]\right)\right]_j = \alpha c^* + \beta \mod 2^{B_c}$. Then, one may set

$$l[t]_j = c^* + k2^{B_c} - c[t-1]_j$$

for the value of the integer k that makes $l[t]_j \in \mathbb{Z}(B_l)$, which is unique since $B_l \leq B_c$.

For what concern property 2 in Section IV, notice first that, despite the permutation applied at each step, one may choose any entry of $c[0]$ and follow its evolution as it is processed by $H_{\alpha,\beta}$ and then mapped into a possibly different entry of $c[1]$ and to a different entry of $c[2]$ and so on. From this point of view, the evolution of $c[t]$ is the parallel of n_c evolutions of single entries, intertwined by the subsequent permutations.

If we indicate with $s[t]$ the evolving scalar (contained in a different entry of $c[t]$ at each time) and with $w[t]$ the value of the entry of $l[t]$ that is combined with $s[t]$, each of the intertwined trajectories is modelled by the discrete-time dynamical system

$$s[t+1] = H_{\alpha,\beta}\left(s[t] + w[t]\right) \tag{4}$$

in which we may interpret $s[t] \in \mathbb{N}(B_c)$ as the system state and $w[t] \in \mathbb{Z}(B_l)$ as a perturbation applied at each time step.

We know from [38] that α and β make the unperturbed system periodic and we indicate the corresponding cycle with $\sigma_0, \ldots, \sigma_{2^{B_c}-1}$, that is nothing but a specified ordering of the digital B_c-bits words.

To analyze the statistical properties of (4), we assume that at each instant, the state $s[t]$ can be any of the possible values in $\mathbb{N}(B_c)$, each with probability $p[t]_j = \Pr\{s[t] = \sigma_j\}$. State probabilities $p[t]$ evolve according to a matrix of transition probabilities $K_{j,k} = \Pr\{s[t+1] = \sigma_j | s[t] = \sigma_k\}$ that defines a finite Markov chain $p[t+1] = Kp[t]$.

Such a Markov chain is ergodic if the transition matrix K is *irreducible*, i.e., if for any pair j, k there is an integer τ such that $K^{\tau}_{j,k} > 0$. If there is an integer τ such that $K^{\tau} > 0$ simultaneously in all its entries, then the matrix is called *primitive* and the Markov chain is mixing. We may prove that the Markov chain corresponding to (4) is mixing. To do so, note first that with no perturbation the system would follow its cycle so that K would mainly have null entries with the exception of $K_{(j+1) \mod 2^{B_c}, j} = 1$ for all j to model the fact that $\sigma_{j+1} = H_{\alpha,\beta}(\sigma_j)$. Such a matrix is irreducible but not primitive.

We will assume that our perturbation is such that the probability $\Pr\{\sigma_{j''} = H_{\alpha,\beta}(\sigma_{j'} + w[t])\}$ is non-null in at least two cases. The first is for all $j'' = (j'+1) \mod 2^{B_c}$, that together with the definition of the unperturbed period $\sigma_{(j+1) \mod 2^{B_c}} = H_{\alpha,\beta}(\sigma_j)$, is equivalent to ask that $\Pr\{w[t] = 0\} > 0$ for any t, i.e., that there is always a non-negligible chance that the input signal $w[t]$ is null. The second is that for at least one pair of indexes one has that $j' - j''$ is even.

From the first assumption, we have that, in the true transition matrix we have $K_{(j+1) \mod 2^{B_c}, j} > 0$ for all j. From the second assumption, we have that $\Pr\{s[t+1] = \sigma_{j''} | s[t] = \sigma_{j'}\} > 0$, for at least one pair of indexes with an even difference and thus that, for those indexes, $K_{j'',j'} > 0$.

Hence, we may build a matrix J that is almost everywhere null with the exception of the indexes j', j'' implied by the above assumptions in which $J_{j'',j'} = K_{j'',j'}$. With this we know that $K \geq J$ entry by entry.

Yet, from $K \geq J \geq 0$ we get that $K^{\tau} \geq J^{\tau}$ for any integer τ. Moreover, due to the very simple structure of J it is easy to verify that J^{τ} tends to be a full matrix when $\tau \to \infty$ and thus that also K^{τ} tends to be a full matrix. This guarantees that (4) is a mixing dynamical system for which it is known that the evolution of state probabilities $p[t+1] = Kp[t]$ tends towards the unique invariant probability assignment such that $\tilde{p} = K\tilde{p}$.

Yet, it is easy to verify that such an invariant probability is the uniform one. In fact, note first that $s[t+1] = H_{\alpha,\beta}(s[t] + w[t]) = H_{\alpha,\beta}\left((s[t] + w[t]) \mod 2^{B_c}\right)$ and that, if $s[t]$ is distributed according to the uniform $\tilde{p}$, then the probability distribution of $(s[t] + w[t]) \mod 2^{B_c}$ is also distributed according to the uniform $\tilde{p}$. From the fact that $H_{\alpha,\beta}(\cdot)$ produces a maximum-length cycle and thus is a bijection, we finally know that when its input is uniformly distributed, its output is also uniformly distributed, confirming that $\tilde{p}$ is invariant.

All together we know that, providing that $l[t]$ and this $w[t]$ satisfy some mild conditions, (4) is a mixing dynamical systems that tends to output uniformly distributed digital words.

REFERENCES

[1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22–32, Feb 2014.

[2] P. A. Catherwood, D. Steele, M. Little, S. Mccomb, and J. Mclaughlin, "A community-based iot personalized wireless healthcare solution trial," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 6, pp. 1–13, 2018.

[3] S. Feng, P. Setoodeh, and S. Haykin, "Smart home: Cognitive interactive people-centric internet of things," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 34–39, February 2017.

[4] M. N. Aman, B. Sikdar, K. C. Chua, and A. Ali, "Low power data integrity in iot systems," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 3102–3113, Aug 2018.

[5] V. Sai and M. H. Mickle, "Exploring energy efficient architectures in passive wireless nodes for iot applications," *IEEE Circuits and Systems Magazine*, vol. 14, no. 2, pp. 48–54, Secondquarter 2014.

[6] V. Sai and M. H. Mickle, "Low power 8051-misa-based remote execution unit architecture for iot and rfid applications," *International Journal of Circuits and Architecture Design*, vol. 1, no. 1, pp. 4–19, 2013.

[7] V. Sai and M. H. Mickle, "Low-power smart passive reu for industrial iot applications," *International Journal of Circuits and Architecture Design*, vol. 2, no. 2, pp. 105–117, 2016.

[8] M. Gao, Q. Wang, M. T. Arafin, Y. Lyu, and G. Qu, "Approximate computing for low power and security in the internet of things," *Computer*, vol. 50, no. 6, pp. 27–34, 2017.

[9] R. Es-sadaoui, L. Azergui, Y. Ghanam, and J. Khallaayoune, "Design and experimentation of a low-power iot embedded system for wireless underwater sensing," in *2017 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Nov 2017, pp. 1–6.

[10] V. Sivaraman, H. H. Gharakheili, A. Vishwanath, R. Boreli, and O. Mehani, "Network-level security and privacy control for smart-home iot devices," in *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct 2015, pp. 163–167.

[11] Y. Yang, L. Wu, G. Yin, L. Li, and H. Zhao, "A survey on security and privacy issues in internet-of-things," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1250–1258, Oct 2017.

[12] J. Deogirikar and A. Vidhate, "Security attacks in iot: A survey," in *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Feb 2017, pp. 32–37.

[13] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1125–1142, Oct 2017.

[14] D. L. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[15] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.

[16] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *2007 IEEE International Symposium on Circuits and Systems*, May 2007, pp. 1959–1962.

[17] D. Bellasi, R. Rovatti, L. Benini, and G. Setti, "A Low-Power Architecture for Punctured Compressed Sensing and Estimation in Wireless Sensor-Nodes," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 5, pp. 1296–1305, May 2015.

[18] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Low-complexity multiclass encryption by compressed sensing," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2183–2195, May 2015.

[19] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "On known-plaintext attacks to a compressed sensing-based encryption: A quantitative analysis," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2182–2195, Oct. 2015.

[20] T. Bianchi, V. Bioglio, and E. Magli, "Analysis of one-time random projections for privacy preserving compressed sensing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 313–327, Feb. 2016.

[21] S. Huh, S. Cho, and S. Kim, "Managing iot devices using blockchain platform," in *2017 19th International Conference on Advanced Communication Technology (ICACT)*, Feb 2017, pp. 464–467.

[22] M. Conoscenti, A. Vetrò, and J. C. D. Martin, "Blockchain for the internet of things: A systematic literature review," in *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, Nov 2016, pp. 1–6.

[23] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram, "Blockchain for iot security and privacy: The case study of a smart home," in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, March 2017, pp. 618–623.

[24] A. Dorri, S. S. Kanhere, and R. Jurdak, "Towards an optimized blockchain for iot," in *Proceedings of the Second International Conference on Internet-of-Things Design and Implementation*, ser. IoTDI '17. New York, NY, USA: ACM, 2017, pp. 173–178.

[25] K. Biswas and V. Muthukkumarasamy, "Securing smart cities using blockchain technology," in *2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Dec 2016, pp. 1392–1393.
[26] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008, available at https://bitcoin.org/bitcoin.pdf - accessed 2018.
[27] J. Haboba, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A pragmatic look at some compressive sensing architectures with saturation and quantization," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 443–459, Sept 2012.
[28] F. Pareschi, M. Mangia, D. Bortolotti, A. Bartolini, L. Benini, R. Rovatti, and G. Setti, "Energy analysis of decoders for rakeness-based compressed sensing of ecg signals," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 11, no. 6, pp. 1278–1289, Dec 2017.
[29] N. Cleju, "Optimized projections for compressed sensing via rank-constrained nearest correlation matrix," *Applied and Computational Harmonic Analysis*, vol. 36, no. 3, pp. 495–507, May 2014.
[30] J. Xu, Y. Pi, and Z. Cao, "Optimized projection matrix for compressive sensing," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 1, p. 560349, 2010.
[31] J. Kovačević and A. Chebira, "Life beyond bases: the advent of frames - part i," *IEEE Signal processing magazine*, vol. 24, no. 4, pp. 86–104, Jul. 2007.
[32] J. Kovačević and A. Chebira, "Life beyond bases: the advent of frames - part ii," *IEEE Signal processing magazine*, vol. 24, no. 6, pp. 115–125, Sep. 2007.
[33] S. Pazos, M. Hurtado, C. H. Muravchik, and A. NEhorai, "Projection matrix optimization for sparse signals in structured noise," *IEEE Transactions on Signal Processing*, vol. 63, no. 15, pp. 3902–3913, Aug. 2015.
[34] M. Mangia, R. Rovatti, and G. Setti, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 5, pp. 1001–1014, May 2012.
[35] M. Mangia, F. Pareschi, V. Cambareri, R. Rovatti, and G. Setti, "Rakeness-based design of low-complexity compressed sensing," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 5, 2017.
[36] M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Low-cost security of iot sensor nodes with rakeness-based compressed sensing: Statistical and known-plaintext attacks," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 2, pp. 327 – 340, Sep. 2018.
[37] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, November 1976.
[38] T. Hull and A. Dobell, "Random number generators," *SIAM Review*, vol. 4, no. 3, pp. 230–254, Jul. 1962.
[39] A. N. Shiryayev, *Selected Works of A. N. Kolmogorov*, ser. Mathematics and Its Application (Soviet Series). Springer, 1992.
[40] M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Security analysis of rakeness-based compressed sensing," in *Circuits and Systems (ISCAS), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 241–244.
[41] R. M. Karp, *Reducibility among combinatorial problems*. Plenum, 1972, pp. 85—-103.
[42] T. Sasamoto, T. Toyoizumi, and H. Nishimori, "Statistical mechanics of an np-complete problem: subset sum," *Journal of Physics A: Mathematical and General*, vol. 34, no. 44, pp. 9555–9568, Nov. 2001.
[43] V. Rijmen and J. Daemen, "Advanced encryption standard," *Information Processing Standard. Gaithersburg, MD, USA: National Institute of Standards and Technology (NIST)*, pp. 19–22, 2001.
[44] kokke, "tiny-aes-c," https://github.com/kokke/tiny-AES-c, 2018.
[45] *Tiva™ C Series TM4C123G LaunchPad Evaluation Kit*, Texas Instruments Inc., Apr. 2013.
[46] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 3, pp. 289–294, March 2003.
[47] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Access Online via Elsevier, 2008.
[48] Y. Zigel, A. Cohen, and A. Katz, "The weighted diagnostic distortion (wdd) measure for ecg signal compression," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 11, pp. 1422–1430, Nov 2000.

Mauro Mangia (S'09-M'13) received the B.Sc. and M.Sc. degrees in electronic engineering and the Ph.D. degree in information technology from the University of Bologna, Bologna, Italy, in 2005, 2009, and 2013, respectively. He was a Visiting Ph.D. Student with the Ecole Polytechnique Federale de Lausanne in 2009 and 2012. He is currently a Post-Doctoral Researcher with ARCES, Statistical Signal Processing Group, University of Bologna. His research interests are in nonlinear systems, compressed sensing, Internet of Things, ultra-wideband systems, and systems biology. He was a recipient of the 2013 IEEE CAS Society Guillemin-Cauer Award. He received the Best Student Paper Award at ISCAS2011. He was the Web and Social Media Chair of ISCAS2018.

Alex Marchioni (S'18) received the B.S. and M.S. degree (with honors) in electronic engineering from the University of Bologna in 2011 and 2015, respectively. He is currently working towards the Ph.D. degree in the Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi" (DEI) of the University of Bologna. His research interests include compressed sensing, biomedical applications and signal processing for the Internet of Things and Big Data analytics.

Fabio Pareschi (S'05-M'08) received the Dr. Eng. degree (Hons.) in electronic engineering from the University of Ferrara, Italy, in 2001, and the Ph.D. degree in information technology from the University of Bologna, Italy, in 2007, under the European Doctorate Project (EDITH). He is currently an Assistant Professor with the Department of Electronic and Telecommunication, Politecnico di Torino. He is also a Faculty Member with ARCES, University of Bologna. His research activity focuses on analog and mixed-mode electronic circuit design, statistical signal processing, compressed sensing, random number generation and testing, and electromagnetic compatibility. He received the Best Paper Award at ECCTD 2005 and the Best Student Paper Award at EMC Zurich 2005. He served as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS - PART II from 2010 to 2013. He was the Local Arrangement Co-Chair of ISCAS 2018.

Riccardo Rovatti (M'99-SM'02-F'12) received the M.S. degree in electronic engineering and the Ph.D. degree in electronics, computer science, and telecommunications from the University of Bologna, Italy, in 1992 and 1996, respectively. He is currently a Full Professor of electronics with the University of Bologna. He has authored approximately 300 technical contributions to international conferences and journals and two volumes. His research focuses on mathematical and applicative aspects of statistical signal processing and on the application of statistics to nonlinear dynamical systems. He was a recipient of the 2004 IEEE CAS Society Darlington Award and the 2013 IEEE CAS Society Guillemin-Cauer Award. He received the Best Paper Award at ECCTD 2005 and the Best Student Paper Award at the EMC Zurich 2005 and ISCAS 2011. He was elected IEEE Fellow in 2012 for contributions to nonlinear and statistical signal processing applied to electronic systems.

Gianluca Setti (S'89-M'91-SM'02-F'06) received the Ph.D. degree in electronic engineering and computer science from the University of Bologna in 1997. From 1997 to 2017, he has been with the School of Engineering, University of Ferrara, Italy. Since 2017 is a Full Professor of analog electronics at the Politecnico of Torino, Italy. He is also a permanent Faculty Member of ARCES, University of Bologna. His research interests include nonlinear circuits, implementation and application of chaotic circuits and systems, electromagnetic compatibility, statistical signal processing, and biomedical circuits and systems. He was a recipient of the 2013 IEEE CAS Society Meritorious Service Award and a co-recipient of the 2004 IEEE CAS Society Darlington Award, the 2013 IEEE CAS Society Guillemin-Cauer Award, the Best Paper Award at ECCTD2005, and the Best Student Paper Award at EMCZurich2005 and ISCAS2011. He held several editorial positions and served, in particular, as the Editor-in-Chief for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS - PART II from 2006 to 2007 and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS - PART I from 2008 to 2009. He was the Technical Program Co-Chair of ISCAS2007, ISCAS2008, ICECS2012, and BioCAS2013 and the General Co-Chair of NOLTA2006 and ISCAS2018. He was a Distinguished Lecturer of the IEEE CAS Society from 2004 to 2005 and from 2014 to 2015 and a member of its Board of Governors from 2005 to 2008. He served as the 2010 President for CASS. He held several other volunteer positions for the IEEE. From 2013 to 2014, he was the First Non North-American Vice President of the IEEE for Publication Services and Products. Since 2019 he is the Editor-in-Chief of the PROCEEDINGS OF THE IEEE.