scch
software competence center
hagenberg

FUZZY.LOGIC.LAB.LINZ

Advances in Knowledge-Based Technologies

Proceedings of the
Master and PhD Seminar
Summer term 2016

Softwarepark Hagenberg
SCCH, Room 0/2
29 June 2016

Software Competence Center Hagenberg
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 800
Fax +43 7236 3343 888
www.scch.at

Fuzzy Logic Laboratorium Linz
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 431
www.flll.jku.at

Program

Session 1. Chair: Bernhard Moser

9:00 A. Kogler:
Efficient and Robust Median-of-Means Algorithms for Location and Regression
9:30 J. Hernández:
Nature Inspired Optimization Algorithms

Session 2. Chair: Susanne Saminger-Platz

10:15 W. Zellinger:
Transfer Learning with Deep Neural Networks for Regression Problems
10:45 B. Moser:
On Preserving Metric Properties of Integrate-and-Fire Sampling

Efficient and Robust Median-of-Means Algorithms for Location and Regression

Alexander Kogler, Patrick Traxler[1]

Abstract. We consider the computational problem to learn models from data that is possible contaminated with outliers. We design and analyze algorithms for robust location and robust linear regression. We show that our algorithms, which are based on a novel extension of the Median-of-Means method by employing the discrete geometric median, are efficient and robust against many outliers in the data. We present theoretical and experimental results.

[1] Software Competence Center Hagenberg, Austria, email: patrick.traxler@scch.at. This research was supported by the FFG Comet Research Program.

Nature Inspired Optimization Algorithms

Jenny Hernández

Abstract - Managing great amounts of information and studying this data has become of great importance in the past years. Data mining and discretization methods have been developed as a solution to this problem. In this talk we will discuss some nature inspired algorithms regarding this matter.

Transfer Learning with Deep Neural Networks for Regression Problems

Werner Zellinger

Knowledge-Based Mathematical Systems (KBMS) - Johannes Kepler University Linz
Software Competence Center Hagenberg (SCCH)

Abstract - Transfer Learning (TL) in the field of time-series prediction with neural networks is considered. In particular, the research domain around an industrial project (TRUMPF) guided by the Software Competence Center Hagenberg (SCCH) is examined and possible improvements of the current approaches are considered. This work aims at providing evidence regarding the following aspects: (a) On the goals and directions of the next years and (b) on some experiments of the last three months. Concerning (a), the review of the state-of-the-art for TL with neural networks shows, that the primary key findings of these works are measures for the similarity between the neural networks hidden activation distributions w.r.t different learning tasks. This fact together with the time-series aspect motivates possible future directions. Concerning (b), different regression algorithms are tested on the TRUMPF data, including elastic net regression, support vector machines, long-short-term memory networks, gated recurrent neural networks, auto encoders and some combinations. In particular, a new L2 regularizer-technique for multi-task neural networks is proposed and tested on the data.

On Preserving Metric Properties of Integrate-and-Fire Sampling

Bernhard A. Moser
Software Competence Center Hagenberg, Austria
Email: bernhard.moser@scch.at

***Abstract*—The leaky integrate-and-fire model (LIF), which consists of a leaky integrator followed by a threshold-based comparator, is analyzed from a mathematical metric analysis point of view. The question is addressed whether metric properties are preserved under this non-linear operator that maps input signals to spike trains, or, synonymously, event sequences. By measuring the distance between input signals by means of Hermann Weyl's discrepancy norm and applying its discrete counterpart to measure the distance between event sequences, it is proven that LIF approximately preserves the metric. It turns out that in this setting, for arbitrarily small thresholds, LIF is an asymptotic isometry.**

***Keywords*—Integrate-and-fire sampling, discrepancy norm, isometry**

I. INTRODUCTION

Integrate-and-fire sampling is well known in computational neuroscience as a simplified model of the input-output behavior of a neuron [1]–[4]. Its time encoding principle is also encountered in neuromorphic engineering [5]–[9] and in the mathematical literature of signal recovery from non-uniform samples [10]–[12].

The leaky integrate-and-fire model (LIF) is probably the most prominent example of a formal spiking neuron model. Its model relies on a resistor-capacitor circuit representation [13] which consists of a capacitor C in parallel with a resistor R driven by a current $I(t)$. This RC circuit yields the standard differential equation $RC\frac{dU}{dt} = -U(t) + RI(t)$ which relates the induced voltage $U(t)$ to the driving current $I(t)$ at time t. The time constant $\tau = RC$ refers to the the leakage of the integrator. Finally, by solving this first-order linear differential equation we obtain the leaky integrator with leakage parameter τ, i.e.,

$$U(t) = \frac{1}{C}\int_{t_k}^{t} I(s)e^{\frac{s-t}{\tau}}ds, \tag{1}$$

where, in the neuronal context, $I(t)$ models the stimulus, U the membrane potential and τ the membrane time constant of the neuron. The firing time instants t_k are defined by a threshold criterion. Synonymously, the firing time instants are referred to as time events. The event that $U(t)$ reaches a positive threshold $\theta_p > 0$ at time t_k is encoded by the pair (t_k, η_k), where $\eta_k = \eta(t_k) = \theta_p$ which represents the value of the threshold. Analogously, if $U(t)$ reaches a negative threshold $-\theta_n < 0$ at time t_k then this event is represented by (t_k, η_k) where $\eta_k = -\theta_n < 0$. Immediately after a firing event the value of the integrator is reset and the process repeats. Figure 1 illustrates the block diagram of the LIF sampler. Note that the sequence of pairs (t_k, η_k) defines a function in time

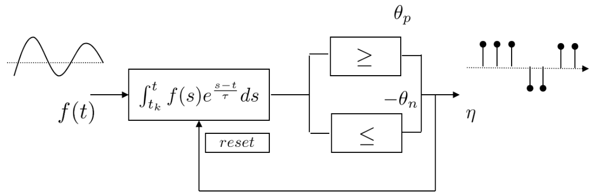

Fig. 1. Block diagram of leaky integrate-and-fire sampling with time constant $\tau > 0$

which assumes the value η_k at time point t_k and vanishes elsewhere. The corresponding sequence is referred to as *event sequence*. Consequently, the notion of a union or intersection of event sequences is defined by means of the corresponding set operations on the underlying sets of events.

For sake of convenience of notation, in the following we restrict to equal positive and negative thresholds, i.e., $\theta_p = \theta_n = \theta > 0$. The results of this paper can proven also for $\theta_p \neq \theta_n$ by only slight modifications. For the purpose of our mathematical analysis, therefore, we have the LIF model

$$\pm\theta = \int_{t_k}^{t_{k+1}} f(s)e^{\frac{s-t_{k+1}}{\tau}}ds, \tag{2}$$

where $f : [t_0, t_E] \to \mathbb{R}$ is supposed to be Lebesgue integrable, i.e., $f \in L[t_0, t_E]$. Let us introduce the set of event sequences $\mathbb{E}[t_0, t_E]$ on the time interval $[t_0, t_E]$ as the set of all functions $\eta : [t_0, t_E] \to \{-\theta_n, 0, \theta_p\}$ where the set of time events, i.e., $\{t \in [t_0, t_E] | \eta(t) \neq 0\}$ has no accumulation point. Note that on a compact domain $[t_0, t_E]$ the assumption of being Lebesgue integrable, $f \in L[t_0, t_E]$, implies (essential) boundedness of f, which in turn guarantees that (2) induces a sequence of events $(t_k, \eta(t_k))_k$ where $(t_k)_k$ has no accumulation point. This means that the resulting sequence of events is an event sequence. In particular, the notion of *next event* is well defined.

This paper is motivated by the question whether it is necessary to (approximately) reconstruct the stimuli from the spike trains when asking whether two spike trains refer to similar stimuli. Therefore, we are interested in the question whether these input and output spaces can be enriched by a topological (metric) structure that remains invariant under LIF.

Let us introduce the mapping

$$\Delta_{\text{LIF}}^{(\theta)} : L[t_0, t_E] \to \mathbb{E}[t_0, t_E], \tag{3}$$

which transfers an input signal f into an event sequence $\eta_f^{(\theta)} = \Delta_{\text{LIF}}^{(\theta)}(f)$ by applying the LIF sampler (2).

Equipping these spaces with a metric $d_F(.,.)$ and $d_E(.,.)$, respectively, the question reads whether

$$d_F(f_1, f_2) = \lim_{\theta \to 0} d_E(\Delta_{\text{LIF}}^{(\theta)}(f_1), \Delta_{\text{LIF}}^{(\theta)}(f_2)). \qquad (4)$$

As pointed out in [14], for the special case $\tau = \infty$ the asymptotic isometry property (4) can only be satisfied if the metric d_E is equivalent to that of Hermann Weyl's discrepancy measure, i.e., $d_E(\eta_1, \eta_2) = \sup_{a,b} |\int_a^b \eta_1 - \eta_2 dc|$, where dc denotes the counting measure. For functions f, g we have the analog definition of Weyl's discrepancy measure by integrating w.r.t the Lebesgue measure $d\lambda$ instead of the counting measure, i.e., $d_F(f_1, f_2) = \sup_{a,b} |\int_a^b f_1 - f_2 d\lambda|$.

In this paper, choosing Weyl's discrepancy measure for both, the input and the sequence space with respect to the Lebesgue and the counting measure, respectively, we will prove the asymptotic isometry property (4) for LIF for arbitrary time constants $\tau > 0$.

The paper is organized as follows. First of all, in Section II we recall some basic properties of Hermann Weyl's discrepancy norm. In particular, we consider MMD intervals which are intervals of minimal length of maximal partial sum (discrepancy). In the following, Section III contains the proof which consists of two steps, a denseness argument regarding the generated spikes of non-trivial signals in Subsection III-A and a convergence argument regarding MMD intervals in Subsection III-B.

II. Hermann Weyl's Discrepancy Measure

Let us consider a path $\gamma = ((0,0)^T, (1,x_1)^T, (1,x_1)^T + (1,x_2)^T, \ldots, \sum_{i=1}^n (1,x_i)^T)$ in $(\mathbb{N}_0 \times \mathbb{R})^{n+1}$. The diameter (range) of γ in the direction of $(0,1)^T$ is given by

$$\begin{aligned} & \max_{1 \le n_1, n_2 \le n} \left| \sum_{i=n_1}^{n_2} \left\langle (1, x_i)^T, (0,1)^T \right\rangle \right| \\ = & \max_{1 \le n_1, n_2 \le n} \left| \sum_{i=n_1}^{n_2} x_i \right| \\ = & \max_{1 \le i \le n} \sum_{j=0}^{i} x_j - \min_{1 \le i \le n} \sum_{j=0}^{i} x_j, \end{aligned}$$

where $\langle .,. \rangle$ denotes the usual inner product, and $x_0 = 0$. Hermann Weyl [15] introduced the measure

$$\mu_D : \mathbb{R}^n \to \mathbb{R}_0^+, \, (x_1, \ldots, x_n) \mapsto \max_{1 \le a \le b \le n} \left| \sum_{i=a}^{b} x_i \right|,$$

in the context of measuring irregularities of distributions. This measure was investigated in the context of numerical integration [16], computational geometry [17], [18], pattern recognition [19], image processing [20], [21] and random walks [22].

Analogously, this measure can be defined for an event sequence η given by its sequence of events $(t_k, \eta(t_k))_k$ and set of time events $\mathcal{T} = (t_k)_k$, i.e,

$$\|\eta\|_D = \sup_{a,b \ge 0} \left| \sum_{a \le t_k \le b} \eta(t_k) \right| \qquad (5)$$

and, respectively, for a Lebesgue integrable function f, i.e.,

$$\|f\|_D = \sup_{a,b \ge 0} \left| \int_a^b f d\lambda \right|. \qquad (6)$$

For convenience in (5) and (6) we use the same symbol $\|.\|_D$ as its definition becomes clear from the context. In both cases $\|.\|_D$ satisfies the axioms of a norm.

A. Relevance for Level-Crossing Sampling

Recently, we pointed out the special role of Hermann Weyl's discrepancy measure for level-crossing sampling [14], [23]. While uniform sampling guarantees that small deviations of continuous input signals can cause only small deviations of the resulting sequences of sampled data, this, in general, is not the case for level-crossing sampling. In order to illustratively demonstrate the problem that even arbitrarily small deviations in the input can lead to large deviations in the resulting sampled coded data, consider two sinus signals: the first with amplitude $\alpha > 0$, and the other with an amplitude β, arbitrarily close but below α.

In the first case, send-on-delta sampling with threshold $\theta = \alpha$ produces a sequence of infinitely many events (t_k, η_k) with alternating signs, -1 and 1, while in the second case the signal is below the threshold, which means that no event is triggered at all.

Similar effects can also be observed for other types of threshold-based sampling schemes such as level-crossing sampling with hysteresis, send-on-delta and integrate-and-fire. This means that for event-based sampling, small changes of the threshold can lead to drastic changes in the resulting sequence of events. Therefore, in general, it is not possible that deviations between resulting event sequences, w.r.t. a given event metric, can be made arbitrarily small by limiting the deviation between the input signals for threshold-based sampling. Rather, it turns out that stability is a matter of degree and, above all, a matter of the chosen event metric. It turns out that for being a stable event metric it has to be equivalent to Weyl's discrepancy norm.

As pointed out in [23], Weyl's discrepancy satisfies the following asymptotic metric equivalence relation for integrate-and-fire sampling with $\tau = \infty$. See the Appendix for a proof.

Proposition 2.1 (**Asymptotic Isometry,** $\tau = \infty$**):** Let $f_1, f_2 \in L[t_0, t_E]$ be bounded integrable signals. Let $\eta_{f_i}^{(\theta)} = \Delta_{\text{LIF}}^{(\theta)}(f_i)$ be the event sequences that result from LIF sampling with threshold $\theta > 0$ and $\tau = \infty$. Then

$$\|f_1 - f_2\|_D - 4\theta \le \|\eta_{f_1}^{(\theta)} - \eta_{f_2}^{(\theta)}\|_D \le \|f_1 - f_2\|_D + 2\theta,$$

hence,

$$\lim_{\theta \to 0} \|\eta_{f_1}^{(\theta)} - \eta_{f_2}^{(\theta)}\|_D = \|f_1 - f_2\|_D. \qquad (7)$$

Note that Proposition (2.1) provides a solution to the problem stated in (4) for $\tau = \infty$, i.e., if we choose the discrepancy norm $\|.\|_D$ for d_F and d_E in (4) by its continuous and discrete variant, respectively, we obtain the asymptotic isometry property (7).

In this paper we present a sketch for a proof that (7) is also valid for arbitrary time constants $\tau > 0$ for LIF. To be precise, in this paper we prove the asymptotic isometry property (4) in Section III for functions f that vanish only on a finite number of proper intervals I_k. The mathematical problem that arises in the general case is discussed in subsection III-C.

In contrast to the proof of (7) in [23], which relies on the reconstruction of the signal values at the points of time of firing, here, we avoid the reconstruction by following another idea based on so-called MMD intervals. For details regarding MMD intervals see [22] and [24].

B. MMD Intervals for Event Sequences

Let us consider an event sequence $\eta : [t_0, t_E] \to \mathbb{R}$ given by its sequence of events $(t_k, \eta_k)_k$, i.e., $\eta(t_k) = \eta_k$ and $\eta(t) = 0$ whenever t, where the sequence $(t_k)_k$ does not contain any subsequence converging to an accumulation point. Suppose that $\|\eta\|_D = d < \infty$.

We call $J = [a, b) \subseteq [0, \infty)$ a minimal interval with maximal discrepancy (MMD) of η if $|\sum_{t_k \in J} \eta_k| = \|\eta\|_D = d$ and if $L \subsetneq J$ implies $|\sum_{t_k \in L} \eta_k| < d$, where L is a subinterval of J.

Note that MMD intervals are mutually disjoint. Let us enumerate these MMD intervals in an increasing order, $(J_j)_{j=1,\ldots,K}$, such that $j_1 < j_2$ implies $t_r < t_s$ for all $t_r \in J_{j_1}$ and $t_s \in J_{j_2}$. By this we obtain the partition of subintervals

$$[t_0, t_E] = \widetilde{J}_0 \cup \bigcup_{i=1}^{K} (J_j \cup \widetilde{J}_j) \tag{8}$$

with MMD intervals $J_j = [a_j, b_j]$ and in-between intervals $\widetilde{J}_j$ $(j = 1, \ldots, K-1)$ such that $\sum_{t_k \in \widetilde{J}_j} \eta_k = 0$. Observe that, in analogy to the Chebychev alternation theorem, the signs $(\sigma_j)_j \in \{-1, 1\}^K$ of the sums $\sum_{k \in J_j} \eta_k$ are alternating. As a direct consequence of the MMD property we obtain Lemma 2.2. A proof can be found in [22].

Lemma 2.2: Let $\eta : [t_0, t_E] \to \mathbb{R}$ be an event sequence of finite discrepancy norm, $\|\eta\|_D < \infty$. Let η be given by its sequence of events $(t_k, \eta_k)_k$. Further, let $(J_j)_{j=1}^K$, $J_j = [a_j, b_j]$, be the sequence of MMD intervals in increasing order. Then, for any $j \in \{1, \ldots, K\}$ and $c \in [a_j, b_j]$ there holds

$$\sigma_j \sum_{t_k \in [a_j, c]} \eta_k > 0$$

and

$$\sigma_j \sum_{t_k \in [c, b_j]} \eta_k > 0$$

where $\sigma_j \in \{-1, 1\}$ denotes the signum of $\sum_{t_k \in [a_j, b_j]} \eta_k$.

Analogous results can be stated for the Lebesgue measure based version (6).

III. Proof of Asymptotic Isometry for Weyl's Discrepancy

The proof consists of two steps. In the first step, Subsection III-A, we show that for sufficiently small thresholds $\theta > 0$ the generated time events spread densely across the time domain of non-zero signal values, i.e., $\{t \in [t_0, t_E] | \, f(t) \neq 0\}|$. In the second step, Subsection III-B, MMD intervals of the event sequences resulting from applying LIF with thresholds tending to zero, $\theta_k \downarrow 0$, are considered. By selecting a convergent subsequence of these MMD intervals, a subinterval $[a, b] \subseteq [t_0, t_E]$ as its limit is specified. By exploiting the denseness property of Subsection III-A, we finally show that $[a, b]$ is an interval with maximal discrepancy with respect to the signal space.

A. On the Denseness of Spikes for Thresholds Tending to Zero

Since the leakage parameter $\alpha = 1/\tau$ is not vanishing the integrator of LIF also depends on the distribution of the spikes. The closer the neighboring spikes, the better the approximation of the integral part of LIF by an idealized LIF without any leakage, $\alpha = 0$.

Therefore, in this section we wonder how the LIF-based generated spikes are distributed if the thresholds tend to zero. Are there holes or do we get an arbitrarily close-meshed set of time points? Next, we show that there might occur holes, i.e., intervals containing no spikes. However, it turns out that such holes coincide with points in time with vanishing function values which do not contribute to the integration. As a result we get the Dense Co-Spike LIF Representation Lemma 3.1 which tells us that to any spike t_k there are pre-specified close *co-spikes* s_k, v_k such that the LIF integral over $[t_k, t_{k+1}]$ can be split into two integrals with integral ranges $[t_k, s_k]$ and $[v_k, t_{k+1}]$ for sufficiently small thresholds.

*Lemma 3.1 (***Dense Co-Spike LIF Representation***):* Let $\varepsilon > 0$ and $f \in L[t_0, t_E]$. Further, let $(t_k^{(\theta)})_k$ denote the time events (spikes) resulting from applying LIF, (2), with threshold $\theta > 0$ then there are time events (co-spikes) $s_k^{(\theta)}, v_k^{(\theta)} \in [t_k^{(\theta)}, t_{k+1}^{(\theta)})$ in the ε-neighborhoods of $t_k^{(\theta)}$ and $t_{k+1}^{(\theta)}$, respectively, i.e, $\left|t_k^{(\theta)} - s_k^{(\theta)}\right| < \varepsilon$, $\left|t_{k+1}^{(\theta)} - v_k^{(\theta)}\right| < \varepsilon$, satisfying

$$\begin{aligned} \theta &= \left| \int_{t_k^{(\theta)}}^{t_{k+1}^{(\theta)}} f(t) e^{\alpha(t - t_{k+1}^{(\theta)})} dt \right| \\ &= \left| \int_{t_k^{(\theta)}}^{s_k^{(\theta)}} f(t) e^{\alpha(t - t_{k+1}^{(\theta)})} dt + \int_{v_k^{(\theta)}}^{t_{k+1}^{(\theta)}} f(t) e^{\alpha(t - t_{k+1}^{(\theta)})} dt \right| \end{aligned}$$

for all $0 < \theta \le \theta_0$ for some $\theta_0 > 0$.

Proof: Consider $\theta_0 > 0$ and suppose that there is a hole. That means there is a point t^* such that

$$\begin{aligned} t_l^* &:= \limsup\{t_k^{(\theta)} | \, t_k^{(\theta)} < t^*, 0 < \theta < \theta_0\} \\ &< \liminf\{t_k^{(\theta)} | \, t_k^{(\theta)} > t^*(\theta), 0 < \theta < \theta_0\} =: t_r^*. \end{aligned} \tag{9}$$

The hole assumption (9) implies that there is an interval $[u_l^*, u_r^*]$ with $t_l^* \le u_l^* < t^* < u_r^* \le t_r^*$ such that

$$\{t_k^{(\theta)} | \, \theta \le \theta_0\} \cap [u_l^*, u_r^*] = \emptyset \tag{10}$$

for some $\theta_0 > 0$. Now fix $\theta < \theta_0$ and consider $\iota(\theta) = \max\{k | \, t_k^{(\theta)} < u_l^*\}$. By construction, we get $\min\{k | \, t_k^{(\theta)} > u_r^*\} = \iota(\theta) + 1$, i.e., for threshold θ the spike at $t_{\iota(\theta)}^{(\theta)}$ is the last one before the hole and the spike at $t_{\iota(\theta)+1}^{(\theta)}$ is the first after the hole. Note that the sequences $(t_{\iota(\theta)}^{(\theta)})_\theta$ and $(t_{\iota(\theta)+1}^{(\theta)})_\theta$

are convergent, i.e., $\lim_{\theta\to 0} t^{(\theta)}_{\iota(\theta)} = \tilde{t}_l$, $\lim_{\theta\to 0} t^{(\theta)}_{\iota(\theta)+1} = \tilde{t}_r$. Thus,

$$\begin{aligned} 0 &= \lim_{\theta\to 0}\theta \\ &= \lim_{\theta\to 0}\left|\int_{t^{(\theta)}_{\iota(\theta)}}^{t^{(\theta)}_{\iota(\theta)+1}} f(t)e^{\alpha(t-t^{(\theta)}_{\iota(\theta)+1})}dt\right| \\ &= \left|\int_{\tilde{t}_l}^{T} f(t)e^{\alpha(t-T)}dt\right| \end{aligned} \quad (11)$$

for any $T \in [\tilde{t}_l, \tilde{t}_r]$. By applying the Radon-Nikodym theorem we conclude from (11) that $f = 0$ a.e. on $[\tilde{t}_l, \tilde{t}_r]$. ■

B. Proof for Functions Vanishing on Finitely Many Intervals

Though the asymptotic isometry property (4) for Weyl's discrepancy $\|.\|_D$ can be proven for the general case of functions with finite discrepancy on arbitrary, not necessarily compact intervals, for sake of adequateness of this paper we will only present a proof under the restrictions that

- the time domain is given by a compact interval $[t_0, t_E]$,
- the functions are Lebesgue integrable,
- the functions only vanish on finitely many intervals.

The last assumption is related to Lemma 3.1. Its relevance will become clear in the proof. The proof for the general case is mathematically more technical and, therefore, is reserved for a mathematical journal. Nethertheless, we give a sketch of the general situation in Subsection III-C.

*Proposition 3.2 (***Asymptotic Isometry, Version I):** Let $f_1, f_2 \in L[t_0, t_E]$ vanish only on a finite number of proper intervals. Then

$$\lim_{\theta\to 0}\|\Delta^{(\theta)}_{\text{LIF}}(f_1) - \Delta^{(\theta)}_{\text{LIF}}(f_2)\|_D = \|f_1 - f_2\|_D. \quad (12)$$

Proof: The concept of the proof relies on considering MMD intervals of the corresponding event sequences induced by LIF with thresholds tending to zero. For convenience, let us set $\Delta f = f_1 - f_2$, $\eta_1 = \Delta^{(\theta)}_{\text{LIF}}(f_1)$, $\eta_2 = \Delta^{(\theta)}_{\text{LIF}}(f_2)$ and $\Delta\eta^{(\theta)} = \eta_1 - \eta_2$. Let $\mathcal{T}$ be the union of time events w.r.t. η_1 and η_2. Note that with f_1, f_2 also Δf satisfies the assumptions of Proposition 3.2. Let $S \subseteq [t_0, t_E]$ be the set of points which are not inner points of proper intervals on which both functions f_1, f_2 are vanishing. Consider a sequence of thresholds θ_k that tends to zero, i.e., $\lim_k \theta_k = 0$.

Now, let us choose $\varepsilon > 0$. First of all, we point out that there is an index N such that for all $k \geq N$ the mesh size of the union of spikes induced by applying LIF on f_i, $i = 1, 2$, with threshold θ_k restricted to S is bounded by ε. Next, let us turn our focus on MMD intervals. Since the discrepancy measure is based on the supremal partial sum over all subintervals of the given compact time domain $[t_0, t_E]$, the existence of an MMD interval is guaranteed. This means, for any threshold $\theta_k > 0$ there is an MMD interval $[a_k, b_k] \subseteq [t_0, t_E]$ of $\Delta\eta^{(\theta_k)}$. From the compactness of $[t_0, t_E]$ we conclude that there is a convergent subsequence $([a_{k_n}, b_{k_n}])_n$ of $([a_k, b_k])_k$. Let us denote the limits of the borders by $\lim_n a_{k_n} = a$ and $\lim_n b_{k_n} = b$. In the following we will show that $[a, b]$ is an interval of maximal partial integral of Δf with respect to the continuous-time version of $\|.\|_D$. Then, there is a threshold θ_{k_n} such that the MMD borders a_{k_n} and b_{k_n} are within the ε-neighborhood of a and b, respectively. Consequently, also the outermost spikes of η_1 and η_2 of the corresponding MMD interval of $\Delta\eta^{(\theta)}$ are within the ε-neighborhood of a and b, respectively. Let us denote the corresponding spikes by $t^{(1)}_{m_k} \in [a-\varepsilon, b+\varepsilon]$, $k = 1, \ldots, \kappa_1$ and $t^{(2)}_{n_k} \in [a-\varepsilon, b+\varepsilon]$, $k = 1, \ldots, \kappa_2$, respectively. Due to the assumption there are only finitely many proper intervals, $Z^{(1)}_l$, $l = 1, \ldots, z_1$ and $Z^{(2)}_l$, $l = 1, \ldots, z_2$, on which f_1 and f_2 are vanishing, respectively. Note that a spike never can fall inside such an interval. Thus there are only finitely many spikes which bridge such intervals. Let us put the corresponding indices of such bridging spikes together and let us denote this set of indices by $\overline{K}_1$ w.r.t. η_1. Analogously, we define $\overline{K}_2$ w.r.t. η_2. Then, $\{1, \ldots, \kappa_i\} = \overline{K}_i \cup \hat{K}_i$, where $\hat{K}_i = \{1, \ldots, \kappa_i\}\backslash\overline{K}_i$ for $i = 1, 2$. Note that the cardinality of $\overline{K}_i$ is bounded independent from the threshold θ due to the assumption that there are only finitely many such intervals. That is, there is a $K \in \mathbb{N}$ such that $\max\{|\overline{K}_1|, |\overline{K}_2|\} \leq K$. For convenience let us write

$$\begin{aligned} \theta^{(1)}_k &= \int_{t^{(1)}_{m_k}}^{t^{(1)}_{m_k+1}} f_1(t)e^{\alpha(t-t^{(1)}_{m_k+1})}dt, \\ \theta^{(2)}_k &= \int_{t^{(2)}_{n_k}}^{t^{(2)}_{n_k+1}} f_2(t)e^{\alpha(t-t^{(2)}_{n_k+1})}dt. \end{aligned}$$

Due to the boundedness of Δf we obtain

$$\begin{aligned} &\left|\int_a^b \Delta f dt\right| \\ &= \left|\int_a^b f_1 dt - \int_a^b f_2 dt\right| \\ &= \left|4\xi\|\Delta f\|_\infty + \sum_{k=1}^{\kappa_1}\theta^{(1)}_k - \sum_{k=1}^{\kappa_2}\theta^{(2)}_k\right| \\ &= \Bigg|4\xi\|\Delta f\|_\infty + \sum_{k\in\hat{K}_1}\theta^{(1)}_k - \sum_{k\in\hat{K}_2}\theta^{(2)}_k \\ &\qquad + \sum_{k\in\overline{K}_1}\theta^{(1)}_k - \sum_{k\in\overline{K}_2}\theta^{(2)}_k\Bigg| \end{aligned} \quad (13)$$

where $|\xi| \leq \varepsilon$. Note that Lemma 3.1 implies that $\theta^{(i)}_k \leq 2\|f_i\|_\infty\varepsilon$ for $k \in \hat{K}_i$, $i = 1, 2$. Thus

$$\left|\sum_{k\in\overline{K}_1}\theta^{(1)}_k - \sum_{k\in\overline{K}_2}\theta^{(2)}_k\right| \leq 4K\|\Delta f\|_\infty\varepsilon. \quad (14)$$

Note that $\text{MMD}_\eta = \sum_{k=1}^{\kappa_1}\theta^{(1)}_k - \sum_{k=1}^{\kappa_2}\theta^{(2)}_k$ is a maximal partial sum w.r.t. $\eta_1 - \eta_2$. Now, consider some arbitrary interval (c, d). Let $\tilde{c} = \inf\{t \geq c \,|\, |f_1(t) - f_2(t)| > 0\}$ and $\tilde{d} = \sup\{t \leq d \,|\, |f_1(t) - f_2(t)| > 0\}$. If $\tilde{d} \leq \tilde{c}$ then the corresponding partial integral vanishes and is trivially smaller than $\Sigma_f = |\int_a^b \Delta f dt|$. Due to the fact that the mesh size for the set S is bounded by ε

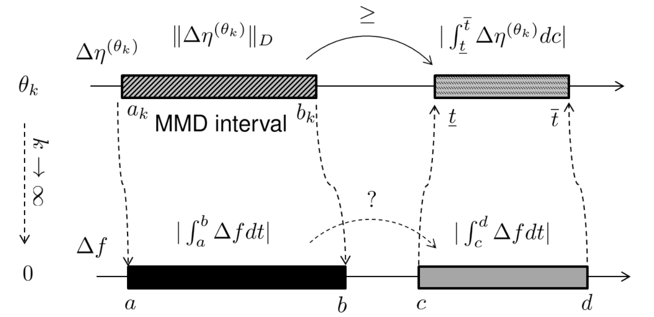

Fig. 2. Illustration of the idea of step (16) in the proof of Proposition 3.2: 1. MMD intervals $[a_k, b_k]$ of LIF induced event sequences are considered; 2. take a convergent subsequence of these intervals; 3. check that the integral over the limit interval $[a, b]$ is greater or equal than another arbitrary interval $[c, d]$; 4. prove this inequality by exploiting the maximality property of the partial sum w.r.t. $[a_k, b_k]$.

there are spikes in the ε-neighborhood of $\tilde{c}$ and $\tilde{d}$, respectively. Let us denote their points in time by

$$\begin{aligned} \underline{t} &\in (\tilde{c}-\varepsilon, \tilde{c}+\varepsilon), \\ \bar{t} &\in (\tilde{d}-\varepsilon, \tilde{d}+\varepsilon). \end{aligned} \tag{15}$$

By taking (13), (15), the characterizing properties of MMD intervals and $|\int_c^d \Delta f dt| = |\int_{\tilde{c}}^{\tilde{d}} \Delta f dt|$ into account, we obtain

$$\begin{aligned} & \left|\int_a^b \Delta f dt\right| + 4K\|\Delta f\|_\infty \varepsilon + 2\cdot 4\|\Delta f\|_\infty \varepsilon \\ \geq\ & \mathrm{MMD}_\eta + 4K\|\Delta f\|_\infty \varepsilon + 4\|\Delta f\|_\infty \varepsilon \\ \geq\ & \left|\sum_{t_i \in \mathcal{T}\cap[\underline{t},\bar{t}]} \Delta\eta^{(\theta_k)}(t_i)\right| + 4K\|\Delta f\|_\infty \varepsilon + 4\|\Delta f\|_\infty \varepsilon \\ \geq\ & \left|\sum_{t_i \in \mathcal{T}\cap[\underline{t},\bar{t}], i\in\hat{K}_1} \theta_i^{(1)} - \sum_{t_i \in \mathcal{T}\cap[\underline{t},\bar{t}], i\in\hat{K}_2} \theta_i^{(2)}\right| + 4\|\Delta f\|_\infty \varepsilon \\ \geq\ & |\int_{\tilde{c}}^{\tilde{d}} \Delta f dt| \\ =\ & |\int_c^d \Delta f dt|. \end{aligned} \tag{16}$$

See Figure 2 which illustrates the idea of this step. (16) proves that $[a, b]$, indeed, is an MMD interval of Δf. Together with (13), which shows that $\lim_k \|\Delta\eta^{(\theta_k)}\|_D = \int_a^b \Delta f dt$, we finally come to the conclusion that $\lim_k \|\Delta\eta^{(\theta_k)}\|_D = \|\Delta f\|_D$, which proves (21). ■

C. Discussion of Proof for General Case

From the Lemma 3.1 and its proof we immediately obtain Corollary 3.3, stating that for a function f that does not vanish on any interval $[a, b]$, $a < b$, eventually, for sufficiently small thresholds the LIF-induced set of spikes are pre-specified close meshed, i.e., a neighboring spike is at most a pre-specified distance $\varepsilon > 0$ away.

*Corollary 3.3 (***Denseness of Spikes***):* Let $\varepsilon > 0$ and $f \in L[t_0, t_E]$ such that there is no proper interval $[a, b] \subseteq [t_0, t_E]$, $a < b$, with $f = 0$ a.e. on $[a, b]$. Then, there is a $\theta_0 > 0$ such that for all $\theta < \theta_0$ the mesh size of the corresponding spikes $(t_k^{(\theta)})_k$ is bounded by ε, i.e.,

$$\sup_k \left(t_{k+1}^{(\theta)} - t_k^{(\theta)}\right) < \varepsilon. \tag{17}$$

It is interesting to observe that the set of functions of Corollary 3.3 is dense in $L[t_0, t_E]$ with respect to the discrepancy norm $\|.\|_D$. For convenience let us denote the set of such functions which do not vanish on any proper interval by $\mathcal{D} \subseteq L[t_0, t_E]$. Moreover, we can show that to any given $\varepsilon > 0$ and function $f \in L[t_0, t_E]$ there is a function $\tilde{f} \in \mathcal{D}$ such that they have the same number of MMD intervals and that the Hausdorff distance between such corresponding MMD intervals is bounded by $\varepsilon > 0$.

*Theorem 3.4 (***$\|.\|_D$-denseness of $\mathcal{D}$ in $L[t_0, t_E]$***):* Let $\varepsilon > 0$. Let $f \in L[t_0, t_E]$ and suppose that $\|f\|_D < \infty$. Further, let I_k denote the MMD intervals of f on $[t_0, t_E]$. Then there is a $\tilde{f} \in \mathcal{D}$ such that f and $\tilde{f}$ have the same number of MMD intervals satisfying

$$\|f - \tilde{f}\|_D \leq \varepsilon, \tag{18}$$

$$d_H(I_k, \tilde{I}_k) \leq \varepsilon, \tag{19}$$

where $\tilde{I}_k$ denote the corresponding MMD intervals of $\tilde{f}$, and d_H denotes the Hausdorff metric.

Proof: Given a function $f \in L[t_0, t_E]$, which vanishes on some interval I. The idea is to bridge this hole by some auxiliary function that fluctuates around the zero line with a maximal pre-specified discrepancy. As a candidate for such a function we consider a sinus with amplitude $\varepsilon > 0$ and frequency π/ε, i.e.,

$$\begin{aligned} \chi_I^{(\varepsilon)}(t) &= \varepsilon \sin\left(t\frac{\pi}{\varepsilon}\right) 1_I(t), \\ \chi_f^{(\varepsilon)}(t) &= \sum_k \chi_{I_k}^{(\varepsilon)}(t), \end{aligned} \tag{20}$$

where I_k denote the proper intervals on which f vanishes. Note that

$$\|\chi_I\|_D \leq \frac{2}{\pi}\varepsilon^2$$

for arbitrary intervals I and that the statements (18) are a direct consequence of construction (20) by setting $\tilde{f} = f + \chi_f^{(\varepsilon)}$. ■

The $\|.\|_D$-denseness of $\mathcal{D}$ in $L[t_0, t_E]$ allows to circumvent the problem with the zero intervals (proper intervals on which f vanishes) and is probably the first step towards proving the asymptotic isometry property (4) in the general case:

Conjecture (**Asymptotic Isometry, Version II**): Let $f_1, f_2 \in L_D[t_0, \infty)$. Then

$$\lim_{\theta\to 0} \|\Delta_{\text{LIF}}^{(\theta)}(f_1) - \Delta_{\text{LIF}}^{(\theta)}(f_2)\|_D = \|f_1 - f_2\|_D. \tag{21}$$

Its analysis is reserved for future research.

IV. Conclusion

In this paper we presented an introduction in the analysis of metric preserving properties of LIF as mapping between metric spaces. We concentrated on Weyl's discrepancy measure as metric as it is the only one (up to equivalence) which guarantees stability in the metric sense. We outlined a proof that LIF is an asymptotic isometry with respect to this metric. We presented a detailed proof for the restricted scenario that is characterized by compact time domains and functions that only vanish on a finite number of proper intervals, and gave an outlook for the general case.

APPENDIX

The appendix deals with the special case of $\tau = \infty$. Thus the LIF model (2) becomes

$$\pm\theta = \int_{t_k}^{t_{k+1}} f(s)ds$$

and, by setting $g(t) = \int_{t_0}^{t} f(s)ds$, its firing condition reduces to an on-delta-send sampling scheme (SOD). That is, an event is fired if g crosses a level $k\theta$, $k \in \mathbb{Z}$, for the first time (subsequent repeated crosses at the same level are not counted). Note that for (essentially) bounded f we obtain a Lipschitz continuous g. Therefore, Proposition 2.1 can be reformulated in terms of SOD and Lipschitz continuous functions. For details see [23].

Proposition A.1: Let $g_1, g_2 : [t_0, t_E] \to \mathbb{R}$, be bounded Lipschitz continuous-time signals, $\theta > 0$, and let $\eta_1^{(\theta)}$, $\eta_2^{(\theta)}$ be the event sequences induced by g_1, g_2, respectively, w.r.t. the on-delta-send sampling principle, . Then, with $\eta_{\text{diff}}^{(\theta)} = \eta_2^{(\theta)} - \eta_1^{(\theta)}$, $f_{\text{diff}} = g_2 - g_1$ we obtain

$$\max\{2\|f_{\text{diff}}\|_\Phi - 4\theta, 0\} \le \|\eta_{\text{diff}}^{(\theta)}\|_D \le 2\|f_{\text{diff}}\|_\Phi + 2\theta, \quad (22)$$

where $\|g\|_\Phi = \sup_{t\in[t_0,t_E]} g(t) - \inf_{t\in[t_0,t_E]} g(t)$.

Proof: Note that because of $g(t) = \int_{t_0}^{t} f(s)ds$ we obtain $\|g\|_\Phi = \|f\|_D$ Proposition A.1 is equivalent to Proposition 2.1. Further, note that the diameter semi-norm $\|.\|_\Phi$ can also be represented by

$$\|g\|_\Phi = 2\inf_{c\in\mathbb{R}} \sup_{t\in[t_0,t_E]} |g(t) - c|. \quad (23)$$

Let $\mathcal{T}$ be the ordered set of sampling points of the event sequences $\eta_1^{(\theta)}$ and $\eta_2^{(\theta)}$, respectively. Then for some $t_i \in \mathcal{T}\backslash\{t_0\}$ we have either $\eta_1(t_i)^{(\theta)} \neq 0$ or $\eta_2(t_i)^{(\theta)} \neq 0$. Set

$$\kappa_1(j) := \max(\{k \le j \,|\, \eta_1(t_k) \neq 0\} \cup \{0\}), \quad (24)$$

and, analogously, for η_2. If $\eta_1^{(\theta)}(t_j) \neq 0$ we get $\kappa_1(j) = j$. If $\eta_1^{(\theta)}$ is vanishing on $[t_0, t_j]$ (24) yields $\kappa_1(j) = 0$. In all other cases $\kappa_1(j)$ denotes the index of the closest sampling point $t_{\kappa_1(j)} \le t_j$ at which $\eta_1^{(\theta)}$ deviates from zero. Let us assume that $|\mathcal{T}| \ge 2$ and consider a $t_j \in \mathcal{T}\backslash\{t_0\}$. Without loss of generality we may assume that $\eta_2^{(\theta)}(t_j) \neq 0$. Note that

$$\begin{aligned} & |g_1(t_{\kappa_1(j)}) - g_2(t_{\kappa_2(j)})| \\ = \; & |g_1(t_{\kappa_1(j)}) - g_2(t_j)| \\ \le \; & |g_1(t_j) - g_2(t_j)| + \theta. \end{aligned} \quad (25)$$

Inequality (25) results from $f_1(t_j) \in (f_1(t_{\kappa_1(j)}) - \theta, f_1(t_{\kappa_1(j)}) + \theta)$ due to the definition of SOD. According to (23) and (25) we obtain

$$\begin{aligned} & \|\eta_2^{(\theta)} - \eta_1^{(\theta)}\|_D \\ = \; & 2\inf_c \max_{i\ge 0, t_i\in\mathcal{T}} |\sum_{j=0}^{i} \eta_2^{(\theta)}(t_j) - \eta_1^{(\theta)}(t_j) + c| \\ = \; & 2\inf_c \max_{i\ge 0, t_i\in\mathcal{T}} |g_2(t_{\kappa_2(i)}) - g_2(t_0) \\ & -g_1(t_{\kappa_1(i)}) + g_1(t_0) + c| \\ = \; & 2\inf_c \max_{i\ge 0, t_i\in\mathcal{T}} |g_2(t_{\kappa_2(i)}) - g_1(t_{\kappa_1(i)}) + c| \\ \le \; & 2\inf_c \max_{i\ge 0, t_i\in\mathcal{T}} \{|g_2(t_i) - g_1(t_i) + c| + \theta\} \\ \le \; & 2\|g_2 - g_1\|_\Phi + 2\theta. \end{aligned}$$

The case $|\mathcal{T}| = 1$ ($\eta_1^{(\theta)}$ and $\eta_2^{(\theta)}$ are vanishing for all $t \ge t_0$) trivially satisfies $\|\eta_2^{(\theta)} - \eta_1^{(\theta)}\|_D = 0 \le 2\|g_2 - g_1\|_\Phi + 2\theta$, which proves the right part of inequality (22).

Now, let us turn to the left part of inequality (22). Due to the intermediate value theorem for continuous functions for all $\rho \ge t_0$ there is a $t_\rho > t_0$ such that

$$\begin{aligned} & 2\|(g_2 - g_1)|_{[t_0,\rho]}\|_\Phi \\ = \; & 2\inf_c \sup_{t\in[t_0,\rho]} |g_2(t) - g_1(t) + c| \\ = \; & 2\inf_c \sup_{t\in[t_0,\rho]} |(g_2(t) - g_2(t_0)) - (g_1(t) - g_1(t_0)) + c| \\ = \; & 2\inf_c |(g_2(t_\rho) - g_2(t_0)) - (g_1(t_\rho) - g_1(t_0)) + c| \\ \le \; & 2[\inf_c | \sum_{\substack{t_j \in \mathcal{T}, \\ t_j \le t_\rho}} \eta_2^{(\theta)}(t_j) - \sum_{\substack{t_j \in \mathcal{T}, \\ t_j \le t_\rho}} \eta_1^{(\theta)}(t_j) + c| + 2\theta] \\ \le \; & 2\inf_c \max_i |\sum_{j=0}^{i} \eta_2^{(\theta)}(t_j) - \eta_1^{(\theta)}(t_j) + c| + 4\theta \\ = \; & \|\eta_2^{(\theta)} - \eta_1^{(\theta)}\|_D + 4\theta. \end{aligned} \quad (26)$$

The last line of (26) follows from (23). ■

REFERENCES

[1] A. Burkitt, "A review of the integrate-and-fire neuron model: I. homogeneous synaptic input.," *Biol Cybern.*, vol. 95, no. 1, pp. 1–19, 2006.

[2] W. Gerstner and W. Kistler, *Spiking Neuron Models: An Introduction*. New York, NY, USA: Cambridge University Press, 2002.

[3] S. Gerwinn, J. H. Macke, and M. Bethge, "Reconstructing stimuli from the spike-times of leaky integrate and fire neurons," *Frontiers in Neuroscience*, vol. 5, no. 1, 2011.

[4] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: exploring the neural code*. Cambridge, MA, USA: MIT Press, 1999.

[5] C. Bartolozzi and G. Indiveri, "Selective attention implemented with dynamic synapses and integrate-and-fire neurons," *NeuroComputing*, vol. 69, no. 16–18, pp. 1971–1976, 2006.

[6] G. Indiveri, "VLSI reconfigurable networks of integrate-and-fire neurons with spike-timing dependent plasticity." The Neuromorphic Engineer, 2005.

[7] G. Indiveri and S. Fusi, "Spike-based learning in VLSI networks of integrate-and-fire neurons," in *International Symposium on Circuits and Systems, (ISCAS), 2007*, pp. 3371–3374, IEEE, 2007.

[8] G. Indiveri, F. Stefanini, and E. Chicca, "Spike-based learning with a generalized integrate and fire silicon neuron," in *International Symposium on Circuits and Systems, (ISCAS), 2010*, (Paris, France), pp. 1951–1954, IEEE, 2010.

[9] S.-C. Liu, T. Delbruck, G. Indiveri, A. Whatley, and R. Douglas, *Event-based neuromorphic systems*. Wiley, 2014.

[10] A. A. Lazar and L. T. Toth, "Time encoding and perfect recovery of bandlimited signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 6, pp. 709–712, 2003.

[11] A. A. Lazar, E. K. Simonyi, and L. T. Toth, "Fast recovery algorithms of time encoded bandlimited signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 4, pp. 237–240, 2005.

[12] H. G. Feichtinger, J. C. Príncipe, J. L. Romero, A. Singh-Alvarado, and G. A. Velasco, "Approximate reconstruction of bandlimited functions for the integrate and fire sampler.," *Adv. Comput. Math.*, vol. 36, no. 1, pp. 67–78, 2012.

[13] R. B. Stein, "Some models of neuronal variability," *Biophysical Journal*, vol. 7, no. 1, pp. 37–68, 1967.

[14] B. A. Moser, "Stability of threshold-based sampling as metric problem," in *IEEE Int. Conf. on Event-based Control, Communications & Signal Processing.*, pp. 1–8, 2015.

[15] H. Weyl, "Über die Gleichverteilung von Zahlen mod. Eins," *Mathematische Annalen*, vol. 77, pp. 313–352, 1916.

[16] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1992.

[17] B. Chazelle, *The Discrepancy Method: Randomness and Complexity*. New York, NY, USA: Cambridge University Press, 2000.

[18] B. Moser, "Geometric characterization of Weyl's discrepancy norm in terms of its n-dimensional unit balls," *Discrete and Computational Geometry*, vol. 48, no. 4, pp. 793–806, 2012.

[19] H. Neunzert and B. Wetton, "Pattern recognition using measure space metrics," Tech. Rep. 28, University of Kaiserslautern, Department of Mathematics, November 1987.

[20] P. Bauer, U. Bodenhofer, and E. P. Klement, "A fuzzy system for image pixel classification and its genetic optimization," in *Cybernetics and Systems '96* (R. Trappl, ed.), vol. 1, (Vienna), pp. 285–290, Austrian Society for Cybernetic Studies, April 1996.

[21] B. Moser, "A similarity measure for image and volumetric data based on Hermann Weyl's discrepancy," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2321–2329, 2011.

[22] B. A. Moser, "The range of a simple random walk on $\mathbb{Z}$: An elementary combinatorial approach," *The Electronic Journal of Combinatorics*, vol. 21, no. 4, p. #P4.10, 2014.

[23] B. A. Moser and T. Natschläger, "On stability of distance measures for event sequences induced by level-crossing sampling.," *IEEE Transactions on Signal Processing*, vol. 62, no. 8, pp. 1987–1999, 2014.

[24] B. A. Moser, "Matching event sequences approach based on Weyl's discrepancy norm," in *IEEE Int. Conf. on Event-based Control, Communications & Signal Processing.*, pp. 1–5, 2015.

ACKNOWLEDGMENT

The author would like to thank the Austrian COMET Program.