Multivaluedness in Networks: Exemplars

Michaël Anton(ie) van Wyk *Senior Member, IEEE* Guanrong Chen *Life Fellow, IEEE*
Eric W. M. Wong *Senior Member, IEEE*

Abstract—This brief presents the first observations of multivaluedness in four systems: a random process, a nonlinear nondynamical system, a nonlinear dynamical system with nonlinearly sensed input and output and an adaptive linear estimator. The preliminary findings reported here, suggest the impact of multivaluedness in different types of networks to range from adverse to benign or even essential.

Index Terms—Counter-cascaded systems, immanence, inverse modeling, multivaluedness, multivalued relation, transcendence.

I. INTRODUCTION

In networks, multivaluedness refers to situations where a single stimulus (cause) is associated with multiple distinct responses (effects), resulting in cause-effect relations that are incompatible with existing network theories. Multivaluedness appears to have been confronted for the first time in [1], where inverse modeling in microwave filters using neural networks was studied. Subsequently, van Wyk [2] proposed a counter-cascaded systems analysis framework, which led to a necessary and sufficient condition for the manifestation of multivaluedness. Furthermore, its application to the structural reduction of complex networks [2], demonstrated the fundamental nature of this framework in the context of signals-and-systems theory.

While the multivaluedness perspective is complementary to existing network theories, it is a unifying perspective which, due to minimal underlying assumptions, unifies linear, nonlinear, discrete- and continuous-time, deterministic and stochastic systems. These attributes equip multivaluedness with potential to produce new insights into complex system behavior and to impact conventional engineering design methodologies.

The contribution reported here is the identification of multivaluedness in four systems ranging from theoretical models to practical systems in machine learning and communication applications. We report that a.) while multivaluedness adversely affects the learning ability of machine learning methods, manifesting as disturbances there, it may be essential for others to function correctly and b.) approximation of a multivalued relation by a mapping may not always be useful.

This work was jointly supported by the Key-Area Research and Development Program of Guangdong Province, China, under Grant 2019B010157002, the Carl and Emily Fuchs Foundation's Chair in Systems and Control Engineering, University of the Witwatersrand, Johannesburg, South Africa, and by the Innovation and Technology Funding of the Hong Kong Special Administrative Region, China, under Grant ITS/359/17.

Michaël A. van Wyk (*Corresponding author*) is with the School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa, and also with the City University of Hong Kong, Hong Kong, China (e-mail: anton.vanwyk@wits.ac.za, mavanwyk@gmail.com).

Guanrong Chen and Eric Wong both are with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China (e-mail: eegchen@cityu.edu.hk and e.wong@cityu.edu.hk, respectively).

Fig. 1. Two counter-cascaded paths $N \circ T$ and M with the common cause, $u \in U$. The implicit relation $S \subset W \times X$ (indicated in gray) takes M 's output as its input and presents N 's output as its output.

II. COUNTER-CASCADED SYSTEMS AND MULTIVALUEDNESS

Figure 1 depicts the *counter-cascaded configuration* of the systems M and $N \circ T$ introduced in [2]. The symbols M , T and N serve the dual purpose of representing the systems in Fig. 1 as well as their respective mathematical descriptions as mappings [3] $M : U \rightarrow W$, $T : U \rightarrow V$ and $N : V \rightarrow X$ that map from and into the (subsets of) vector spaces indicated. The symbol S represents an implicit relation, defined by

$$W \times X \supset S := N \circ T \circ M^{-1}, \quad (1)$$

where M^{-1} denotes the *preimage* [3] of M .

If the image under T of every *equivalence class* [4] induced on U by the preimage of M , is contained in an equivalence class induced on V by the preimage of N , then T is said to be *immanent* with respect to the ordered pair (M, N) or simply (M, N) -*immanent*. If T is not (M, N) -immanent, then T is said to be *transcendent* with respect to the ordered pair (M, N) or simply (M, N) -*transcendent*.

A necessary and sufficient condition for the relation S to be multivalued, now follows [2]:

Theorem II.1. (Multivaluedness) *In Fig. 1, the relation S is multivalued if and only if the mapping T is transcendent with respect to (M, N) .*

An equivalent interpretation of this result is that S is well-defined or single-valued if and only if T is (M, N) -immanent.

Some further consequences of Theorem II.1 follow:

Corollary II.2. *Let M , T , N and S be as depicted in Fig. 1.*

- If M is given and T is of canonical form $T = F \circ M$ for some fixed $F : W \rightarrow V$, then T is (M, N) -immanent for every N .*
- If M is one-to-one then every T is (M, N) -immanent for every N .*
- If M is many-to-one, T is not of canonical form and N is one-to-one, then T is (M, N) -transcendent.*

- d. If the condition for (M, N) -immanence holds everywhere in $B \subset U$, then $T|B$ is $(M|B, N)$ -immanent.¹

Note. Corollary II.2(d) implies that immanence also depends on U ; choosing B to contain exactly one element from each equivalence class of M renders $T|B$ $(M|B, N)$ -immanent.

Corollary II.3. (Existence of a Unique Faithful Model)
For a given mapping T , a unique faithful model or modeling mapping S exists if and only if T is (M, N) -immanent.

Recall that approximation of a multivalued relation S by a function $\hat{S}$, using any optimality criterion, necessarily always yields a nonzero approximation error [2].

III. EXEMPLARS OF MULTIVALUEDNESS

This section starts with the discussion of a random process, followed by the analysis of a nonlinear nondynamical system (i.e. a nonlinear mapping) with nonlinear nondynamical input and output ancillary systems. Next, we consider an instance from a class of problems solved by machine learning methods, namely Input-Measurement-to-Output-Measurement (IMOM) functional modeling of a dynamical system. The final example considers a linear adaptive estimator with imperfections along the signal path, encountered in communications applications.

For each of the last two examples, a first-principles analysis of multivaluedness precedes the analysis based on the proposed framework [2], summarized in Section II. This is done to hone the intuition toward a better understanding of multivaluedness, while also providing a comparative basis and consequent appreciation for the clarity and conciseness of the proposed framework.

A. Random Process

Consider the “system” T to be a real-valued random process $T : \Omega \times \mathbb{N}_0 \rightarrow \mathbb{R}$ defined on the sample space Ω and for all nonnegative integer time instants $\mathbb{N}_0$. The sample-space variable $\omega \in \Omega$ is considered the “input” of T . Furthermore, let $M := I : \Omega \rightarrow \Omega$ and $N := I : \mathbb{R} \rightarrow \mathbb{R}$, the respective identity mappings, be input and output ancillary systems. Then, except for the trivial case of a constant-valued random process, there exists some value $\omega_* \in \Omega$ and time instants $m, n \in \mathbb{N}_0$ such that $T(\omega_*, m) \neq T(\omega_*, n)$. Using (1), we obtain

$$\{T(\omega_*, m), T(\omega_*, n)\} \subseteq S(\omega_*)$$

and hence S is multivalued and T is (I, I) -transcendent.

The motivation for presenting this very simple example here is that it naturally extends to the real-life problem of a dynamical system T with exogenous additive noise contaminating its output, resulting in transcendence and consequently multivaluedness.

B. Nonlinear Nondynamical System

For this example, choose U, V, W and X to be $\mathbb{R}$. For Cases (a) and (b), discussed below, T is the nonlinear mapping defined by

$$v = T(u) := u + \text{sgn}(u),$$

with $\text{sgn}(\cdot)$ is the signum function. Of the several justifications for the term “nondynamical system” used here, the one preferred is the viewpoint that T can be interpreted as a zero-memory Boyd-Chua Volterra system: convolution of the input signal with the Dirac-delta function $\delta(t)$, feeding into an output nonlinearity [5].

Case (a): Let $M(u) := |u|$ and $N(v) := v^3$. We now prove that S in (1) is a multivalued relation, which implies the (M, N) -transcendence of T .

Consider any fixed positive element of $w_* \in W$ and proceed to evaluate $S(w_*)$. We start by determining the preimage $M^{-1}(w_*)$. It is easy to see that $M^{-1}(w_*) = \{-w_*, w_*\}$. The system T , applied to the latter, yields $T(M^{-1}(w_*)) = \{-(w_* + 1), w_* + 1\}$. Now, applying N to the latter two-element set immediately yields

$$S(w_*) = \{-(w_* + 1)^3, (w_* + 1)^3\},$$

which means that the singleton $\{w_*\}$ produces a two-element set under S . Consequently, S is multivalued and T is transcendent with respect to (M, N) . Alternatively, this result follows immediately from Corollary II.2(c).

Finally note that, although here T is (M, N) -transcendent, by Corollary II.2(d), $T|\{0\}$ is $(M|\{0\}, N)$ -immanent.

Case (b): Consider the ancillary mappings $M(u) := |u|$ and $N(v) := v^2$. Now, we prove that T is (M, N) -immanent, from which it then follows that S is well-defined. Consider a fixed positive element $w_* \in M(U)$. Since M is two-to-one, it follows that $M^{-1}(w_*) = \{-w_*, w_*\}$ and $T(M^{-1}(w_*)) = \{-(w_* + 1), w_* + 1\}$. Note that there exists an $x_{w_*} \in N(T(U))$, namely $x_{w_*} := (w_* + 1)^2$, with $N^{-1}(x_{w_*}) = \{-(w_* + 1), w_* + 1\}$, giving the set-inclusion $T(M^{-1}(w_*)) \subseteq N^{-1}(x_{w_*})$ for arbitrary positive w_* and thus for every positive $w_* \in M(U)$. The case $w_* = 0$ follows similarly with these sets now both singletons. This proves that T is (M, N) -immanent and S is well-defined.

Case (c): Suppose M and N are as defined in Case (a). We now derive the expression describing all systems T immanent with respect to (M, N) . Since M is two-to-one and N is bijective, it follows from the contrapositive of Corollary II.2(c) that the only (M, N) -immanent systems T are the piecewise-constant mappings, constant on each of the equivalence classes induced by M . Explicitly, every such T necessarily has the form

$$T_f(u) := f \circ M(u) = f(|u|),$$

for every mapping $f : W \rightarrow V$ (fixed). The corresponding relation S_f ,² in Fig. 1, is the well-defined mapping

$$S_f(w) := N \circ f(w) = (f(w))^3.$$

As an explicit example, for $f(w) := e^{-w}$ the mapping $T(u) = e^{-|u|}$ is (M, N) -immanent and $S(w) = e^{-3w}$ is well-defined. Yet another instance is $f(w) := \sin(\pi w)$ with $T(u) = \sin(\pi|u|)$ and $S(w) = \sin^3(\pi|w|)$.

¹Here, $T|B$ is the restriction of T to the subset B of its domain U .

²Here, the subscript f signifies explicitly the dependency of S on it.

C. IMOM Modeling of a Nonlinear Dynamical System

Consider the nonlinear dynamical system T with its input and output sensed by nonlinear dynamical measurement systems, M and N , respectively. The objective is to determine, for different choices of these measurement systems, whether the resulting IMOM (functional) modeling problem is solvable in a noise-free environment.

An IMOM modeling problem is defined to be *solvable* if the system T is (M, N) -immanent, in which case there exists a faithful model S relating input measurements to output measurements of T . If, however, T is (M, N) -transcendent, this modeling problem is called *unsolvable*. In this circumstance, approximation is the only remaining alternative [2].

Consider the fading-memory Boyd-Chua Volterra system $T : \ell_\infty \rightarrow \ell_\infty$, consisting of a stable first-order linear infinite impulse response (IIR) system feeding into a “soft” deadzone nonlinearity at the output, given by

$$v(n) = h(g * u)(n), \quad n = 0, 1, 2, \dots$$

with $g(\cdot)$ representing the IIR system’s impulse-response (for example, $g(n) := 2^{-n}$) and

$$h(\theta) := \alpha(\theta - \tanh \theta), \quad \alpha := 1/(1 - \tanh 1),$$

the expression for the deadzone nonlinearity. The theoretical results of Section II admit a large class of nonlinear dynamical measurement systems that require advanced mathematical tools and techniques for a complete analysis [3]. For this reason, here we only consider mappings.

Solvable Case. Motivated by the facts that all electronic systems saturate when driven beyond their linear regimes and that dynamical systems with sufficiently fast time-constants can be approximated by nondynamical mappings, we choose the input and output measurement systems to be mappings $M, N : \mathbb{R} \rightarrow \mathbb{R}$, defined by

$$w = M(u) := \beta \tanh u \quad \text{and} \quad x = N(v) := \beta \tanh v,$$

with $\beta := 1/\tanh 1$ and hence $W = X = \ell_\infty$. Since the mapping M is invertible, with inverse

$$M^{-1}(w) = \frac{1}{2} (\ln(1 + w/\beta) - \ln(1 - w/\beta)),$$

an explicit expression can be derived for the model S as solution to this modeling problem.

However, usually an explicit expression for S is difficult to derive, if not impossible. For such situations, Theorem II.1 and its corollaries are indispensable. To demonstrate this approach, first note that the invertibility of M follows from the fact that it is monotonically increasing and hence the existence of the model S immediately follows from Corollary II.2(b).

Unsolvable Case. Consider the same setup, but suppose that the input measurement system now consists of an absolute-value amplifier feeding into the saturation nonlinearity. For this case, the input measurement system is described by

$$w = M(u) := \beta \tanh |u|.$$

Proceeding from first principles, assume initial conditions to be zero and the measured input value at time $n = 0$ to

TABLE I
SUBSYSTEMS’ OUTPUTS AT TIME $n = 0$.

$u(0)$	$v(0)$	$w(0) \mapsto x(0)$
−1	−1	+1 $\mapsto$ −1
+1	+1	+1 $\mapsto$ +1

be $w(0) = +1$.³ Then, the input value $u(0)$ must either be +1 or −1, with the corresponding output value $v(0)$ of T being +1 or −1, respectively, and the measurement system’s output value $x(0)$ then +1 or −1, respectively. This implies that S is required to map $w(0) = +1$ to either $x(0) = +1$ or $x(0) = -1$, based exclusively on knowing that $w(0) = +1$, as shown in Table I. No mapping can achieve this and hence this is impossible to realize. Repeating this reasoning for $n = 1$, assuming as input measurement $w(1) = +1$, yields four distinct output measurement values for $x(1)$, either of which is to be produced by S based only on the knowledge that $w(0) = w(1) = 1$, yet again contradicting the requirement for S to be a mapping; and so on for subsequent values of n . We conclude that this modeling problem is unsolvable.

Alternatively, note that with M two-to-one, N one-to-one and T not of canonical form since $T(-u) = -T(u)$, Corollary II.2(c) yields that this problem is not solvable.

Here, unsolvability implies that every (single-valued) model $\hat{S}$ considered for approximating S yields a nonzero error. Thus, being unaware of the multivaluedness of S , this error could incorrectly be attributed to either a disturbance or uncertainty entering or inaccurate (single-valued) model selection. Even optimal (single-valued) model selection may be futile. To see this, note that due to the sign ambiguity here the minimax optimal approximation to S calculated over any *balanced* [10] subset of the vector space W yields the trivial (single-valued) model $\hat{S}_{\text{opt}} \equiv 0$ which produces zero as output irrespective of the actual input; this is not at all a useful approximation.

Observe that the modeling interpretation of Theorem II.1 suggests that either M or N (or both) may be altered in search of a solvable IMOM modeling problem. Conversely, for given systems M and N , Corollary II.2(a) implies that approximating T by some canonical form, would yield a solvable modeling problem that approximates the original (unsolvable) modeling problem, if solvability is strictly required.

D. Adaptive Linear Estimator

Next, consider an adaptive (linear) single-tone frequency estimator [6][7]. By employing an iterative prediction-correction strategy [6], an implicit estimate of the sinusoidal input signal’s frequency is obtained once convergence occurs. The frequency is then extracted from the adapted coefficient of the estimator as discussed below and hence the term “implicit estimate.”

Specifically, during the prediction step, for a fixed $\Delta > 0$, two past signal values, $u(t - 2\Delta)$ and $u(t - \Delta)$, are linearly combined to form a prediction $\hat{u}(t)$ of the current signal value $u(t)$. During the correction step, the prediction error value $e(t) := u(t) - \hat{u}(t)$ is used by a search algorithm

³For convenience, signal-values will be treated as dimensionless quantities.

[6] to update the estimated frequency of the signal $u(\cdot)$.⁴ This is done by adjusting the frequency estimate in the direction (in parameter space) that produces more accurate future signal-sample predictions during subsequent prediction steps. This iterative frequency estimation process converges as the prediction error converges.

Mathematically, the signal-sample prediction step is based on an elementary identity that relates three uniformly spaced signal-values on a sinusoid $u(\cdot)$ of fixed frequency ω_0 , namely

$$u(t) - 2 \cos(\omega_0 D) u(t - D) + u(t - 2D) = 0, \quad (2)$$

for arbitrary values of $t, D \in \mathbb{R}$. Equation 2 is easily derived by writing the difference equation for the zero-input response of a second-order IIR filter with poles located on the z -plane unit circle, at angles $\pm \omega_0 D$. Obviously, there exist nonsinusoidal signals that also satisfy this expression, but then only for proper subsets of time. Considering the signal value $u(t)$ in (2), to be the predicted signal value $\hat{u}(t)$, yields the adaptive filter's prediction-step equation,

$$\hat{u}_n(t) = 2 \cos(\hat{\omega}_n \Delta) u(t - \Delta) - u(t - 2\Delta). \quad (3)$$

Here, t represents time that varies, while Δ represents a fixed nonzero time delay and $\hat{u}_n(t)$ is the n th-step prediction based on $\hat{\omega}_n$, namely the n th-step updated frequency estimate. We can express (3) in operator form, as

$$\hat{u}_n(\cdot) := M_n(u(\cdot)), \quad (4)$$

where M_n is the input ancillary mapping operating on $u(\cdot)$.

The convergence of the prediction step is adversely affected by imperfections in the system. In the presence of small nonidealities such as distortion or noise, the above identity holds only approximately. For example, when faced with a slight compromise in the accuracy of the time delays between consecutive delayed signal values, (2) becomes an approximate equality, namely

$$u(t - \tau) - 2 \cos(\omega_0 \Delta) u(t - \Delta) + u(t - 2\Delta) \approx 0,$$

for some $0 < |\tau| \ll \Delta$, with τ, Δ and ω_0 appropriately defined. Here we will assume that $\tau > 0$. However, instead of focusing on just the effect of timing imperfections, we consider a larger class of nonidealities which we represent by an unknown perhaps nonlinear operator T that operates on the input signal $u(\cdot)$ and satisfies the following approximate form of (3):

$$T(u(\cdot)) - \hat{u}_n(\cdot) \approx 0(\cdot), \quad 0(t) := 0 \quad \forall t.$$

However, if we cannot directly observe the effects of the nonidealities and need some specific measurement techniques to extract these, then the above expression becomes

$$N(T(u(\cdot))) - \hat{u}_n(\cdot) \approx 0(\cdot),$$

where N embodies the required measurement technique. In order to transform the latter expression into a strict equality, we introduce S , some implicit and unknown relation,⁵ to obtain

$$\varepsilon_n(\cdot) := N(T(u(\cdot))) - S(\hat{u}_n(\cdot)) = 0(\cdot). \quad (5)$$

⁴The notation $u(\cdot)$ is used to emphasize time-dependency of the signal u .

⁵Alternatively [2], define S such that $S(N(T(u(\cdot)))) - \hat{u}(\cdot) \equiv 0(\cdot)$.

Fig. 2. Adaptive estimator with Theorem II.1's partitioning indicated; the implicit relation S and its implication $\varepsilon_n(\cdot) = 0$ are indicated in gray.

To be consistent with Section II, we set $w(\cdot) := \hat{u}_n(\cdot)$ and $x(\cdot) := N(v(\cdot))$. Favoring clarity over generality, we assume that no measurement extraction is required, i.e. $N = I$. Equations (4) and (5) are graphically represented by Fig. 2. In addition, the systems partitioning according to Theorem II.1, is overlaid onto the resulting systems configuration.

In the context of this work, we are interested in the multivaluedness of S and therefore omit the correction step by fixing the estimated frequency to be ω_0 and writing M and $\hat{u}(\cdot)$ for M_n and $\hat{u}_n(\cdot)$, respectively.

We now turn to the aspect of multivaluedness, starting with the first-principles approach for the original problem (3) without nonidealities i.e. $T = I$. First, the prediction-equation's output response to an arbitrary input signal $u(\cdot)$ has to be determined. To this end, fix the signal values $u(t - \Delta)$ and $u(t - 2\Delta)$. By (3), the linear predictor interprets these arbitrary signal values as belonging to some sinusoidal signal of frequency ω_0 , with implied amplitude and phase, $A = A(u(t - \Delta), u(t - 2\Delta))$ and $\phi = \phi(u(t - \Delta), u(t - 2\Delta))$, respectively. A worst-case upper bound on A and hence on $\hat{u}(t)$ follows immediately by assuming these two signal values to be the same and $\cos(\omega_0 \Delta) = -1$, giving $|\hat{u}(t)| \leq A \leq 3|u(t - 2\Delta)|$. This implies that $\hat{u}(\cdot)$ is bounded if $u(\cdot)$ is bounded in magnitude.

Let U be the class of all power signals with power less than 1W, say. Then for an arbitrary $u(\cdot) \in U$, any value $u(t)$ at time t is allowable which means $v(t)$ (equivalently, $S(w(t))$ in Fig. 2) can be made any possible value while $w(t)$ remains fixed for fixed $u(t - \Delta)$ and $u(t - 2\Delta)$. Consequently, S is a multivalued relation theoretically with infinite (i.e. unbounded) spread for signals $u(\cdot) \in U$.

Now, consider the subclass $B_C \subset U$ of all continuous power signals. It is possible to construct a continuous function that passes through the two previously fixed points as well the point $(t, x(t))$, for an arbitrary choice of $x(t)$. This can be achieved from first principles or by multiscale methods [8] [9], and hence S is multivalued with unbounded spread for signals $u(\cdot) \in B_C$.

The still smaller subclass $B_W \subset B_C$ of bandlimited power signals with bandwidth at most W , is considered next. By Bernstein's inequality [4], the time-derivative of every bandlimited signal is bounded and hence Lipschitz-continuous. It follows that S is a multivalued relation with bounded spread

for signals $u(\cdot) \in B_W$.

Further reducing the scope to the subclass $B_S \subset B_C$ of sinusoidal input signals $u(\cdot)$ of all possible frequencies, yields a multivalued relation S with a bounded spread. This result also applies to $B \subset B_S$ the class of all sinusoidal signals with frequencies that are integer multiples of ω_0 .

Finally, consider the subclass of power signals of particular interest here, $B_0 \subset B$ consisting of all sinusoids having frequency ω_0 . Since $T = I$, it follows from (2) that the outputs of M and T are identical which means that $S = I$ and thus single-valued (i.e. has zero spread). This completes the analysis from the first-principles perspective for the ideal case (3) only. For an arbitrary nonideality a first-principles analysis would require additional explicit analytical expressions and assumptions, whereas the alternative approach below, applies even when nonidealities are of a more conceptual nature.

The analysis based on Theorem II.1 now follows. The phenomenon of aliasing implies that M , which linearly combines time distributed samples of the signal that it operates on, is a linear projection and thus necessarily many-to-one. Linear projection operators are bounded [10]. Considering U as specified above, Corollary II.2(c) then implies that every T is (M, I) -transcendent. Equivalently, S is multivalued for *every* T , i.e. not just for the case $T = I$ considered in the above first-principles analysis. So, whether T is a linear or nonlinear, bounded or unbounded operator, S is multivalued. The results for all the other subclasses of input signals considered in the above, follow similarly. However, we single out the case $u(\cdot) \in B_0$. Here, aliasing implies that $M|_{B_0} = I$ (see note below Corollary II.2). By Corollary II.2(b), the S is well-defined for every T ; direct calculation using (1) gives $S = T$.

Finally, note that the use of uniformly spaced samples of continuous-time signals here, was a matter of choice and did not affect the conclusions reached.

E. Summary of Observations

A brief summary of the notable observations, that arose during the analysis of the above examples, follows.

1. Randomness is *not* the only cause of multivaluedness since it was demonstrated that multivaluedness can manifest in the complete absence of randomness. It is important to note that neglecting multivaluedness could lead to erroneous deductions such as the presence of modeling errors, disturbances or uncertainties.
2. The impact of multivaluedness in different types of systems ranges from adverse (Section III-C) to benign or even essential (Section III-D).
3. Corollary II.2(d) implies that the detection of multivaluedness of S , using a subset of input signals, is susceptible to failure.
4. The compromise of approximation of a multivalued relation by some function, may not be useful as was demonstrated in Section III-C.
5. It would be technically more correct to refer to T as immanent or transcendent with respect to (U, M, N) as suggested by Corollary II.2(d) and Section III-D.

IV. CONCLUSION

In this brief, we presented examples of the manifestation of multivaluedness in four different setups ranging from simple to realistic. These examples confirm that the framework, based on the notion of counter-cascaded systems, is not simply a curiosity; it gives insight into why some problems cannot always be solved and suggests that multivaluedness is required for some systems to function correctly. It therefore broadens our ways of thinking about networks and systems. Moreover, it provides a generalization, containing cascaded systems as a special case, when M is the identity mapping, thereby making these results directly applicable to ordinary cascaded systems e.g. communications systems [11].

Currently, work is in progress to extend the theoretical results in [2] to situations where both measurement noise and injected noise will be accounted for. It is foreseen that such results will provide a more general analysis and modeling framework leading toward a more complete understanding of systems in general. In addition, work is under way that applies this framework to distributed measurement systems and mixed modeling as well as to machine learning.

Regarding future work, from Section III-E, it is evident that much further work is required in order to develop a deeper understanding of the significance as well as potential advantages and disadvantages of multivaluedness in systems. Furthermore, work is needed to determine how this framework can be used to design for multivaluedness in accordance with design specifications and objectives for particular applications.

ACKNOWLEDGMENT

The authors thank Proff. B.J. van Wyk and B. Wigdorowitz, Dr. S. Dinger and Mr. A.M. McDonald for technical discussions related to and proofreading of several versions of the manuscript.

REFERENCES

- [1] C. Zhang, J. Jin, W. Na, Q. Zhang and M. Yu, "Multivalued neural network inverse modeling and applications to microwave filters," *IEEE Trans Microw. Theory Techn.*, vol. 66, no. 8, pp. 3781–3797, Aug. 2018.
- [2] M.A. van Wyk, "Multivaluedness in networks: Theory," *IEEE Trans. Circuits Syst., II, Exp. Briefs*, vol. 67, no. 12, pp. 3462–3466, Dec. 2020.
- [3] M. Deng, *Operator-based nonlinear control systems: design and applications*, IEEE Press, 2014.
- [4] A. Lapidoth, *A Foundation in Digital Communication*, 2nd ed., Cambridge University Press, 2017.
- [5] M.A. van Wyk and G. Chen, " L_p -Stability of a class of Volterra systems," *IEEE Trans. Circuits Syst., II, Exp. Briefs*, vol. 67, no. 8, pp. 1460–1471, Aug. 2020.
- [6] F. Zhao, H. Zhao, and W. Wang, "A robust M-shaped error weighted algorithms for censored regression," *Circuits Syst. Signal Process.*, vol. 39, pp. 324–343, Oct. 2019.
- [7] S. Haykin, *Adaptive filter theory*, 5th Edition, Prentice Hall, 2013.
- [8] M.A. van Wyk and T.S. Durrani, "A framework for multiscale and hybrid RKHS-based approximators," *IEEE Trans. Signal Proc.*, vol. 48, no. 12, pp. 3559–3568, Dec. 2000.
- [9] A. Cohen, *Numerical Analysis of Wavelet Methods* (Studies in Mathematics and Its Applications vol. 32), Elsevier, 2003.
- [10] W. Rudin, *Functional Analysis*, 2nd Edition, McGraw-Hill, 1991.
- [11] M.A. van Wyk, L. Ping, G. Chen, "Multivaluedness in networks: Shannon's Noisy Channel Coding Theorem," submitted and under review: *IEEE Trans. Circuits Syst., II, Exp. Briefs*, Nov. 2020.