II. INVERSE ESTIMATION ALGORITHMS

A. Direct inversion approach

One of the simplest solution to the inverse recovery problem is based on the following basic relation between a function and its inverse

$$m^{-1}(u) = \int_0^1 \mathbf{1}\{m(v) \leq u\}\, dv, \tag{2}$$

which, given a set of the (sorted increasingly w.r.t. values of u_k's) input-output measurements $\{(u_k, y_k = m(u_k) + z_k)\}, n = 1, \ldots, N$, leads to the natural (*empirical distribution function*-like) estimate of the nonlinearity inverse

$$\tilde{m}^{-1}(u) = N^{-1} \sum_{k=1}^{N} \mathbf{1}\{y_k \leq u\}. \tag{3}$$

By definition, the estimate is *isotone*. One can however expect that the above estimate works well only when the input signal is uniform (which is nevertheless the case in some applications; see *e.g.* [6], [7]).

B. Dette's et al. approach

In [8], Dette *et al.* presented a novel algorithm which estimates a monotonous regression function in a nonparametric fashion, that is, without implying any other assumption on its shape. The resulting estimate preserves the isotonicity property and works for non-uniform input *probability density functions* as well. The algorithm has three steps:

1) **Initial estimation of a nonlinearity**. This (preliminary) step can be performed with the help of virtually any nonparametric regression estimate, $\hat{m}(u)$ say, like a Nadaraya-Watson or a Gasser-Müller one, see *e.g.* [9], [10], [11]. Observe that the estimate is in general not isotone and will serve in the subsequent algorithm steps.
2) **Estimation of the inverse of the nonlinearity**. This step has two phases. In the first, the density of the output signal, *i.e.* of the function $\left(m^{-1}\right)'(u)$, is estimated using any density estimate with a non-negative kernel (see *e.g.* [12], [13]), and the initial nonlinearity estimate $\hat{m}(u)$ obtained in the the previous step. The inverse, $\hat{m}^{-1}(u)$, is then estimated by virtue of the basic observation that, given an estimate $\left(\hat{m}^{-1}\right)'(u)$, the nonlinearity $\hat{m}^{-1}(u)$ can be recovered from the following formula
$$\hat{m}^{-1}(u) = \int_0^u \left(\hat{m}^{-1}\right)'(v)\, dv. \tag{4}$$
Note here, that the resulting inverse estimate, $\hat{m}^{-1}(u)$, is isotone (by virtue of the assumption that the density estimate $\left(\hat{m}^{-1}\right)'(u)$ is based on non-negative kernel functions).
3) **Recovering the isotone nonlinearity** by inverting the estimate $\check{m}(u) = \left(\hat{m}^{-1}\right)^{-1}(u)$.

In [8, Theorems 3.1-3.2], consistency of the algorithm has been proven for the case of *i.i.d.* data.

C. Haar series implementations

In the original algorithm, the kernel-based estimates of the output density probability function and of the nonlinearity itself were employed (with the only restriction that a kernel function is non-negative in order to preserve monotonicity of the estimate). In the note we present the algorithm instances based on orthogonal Haar wavelet bases; see *e.g.* [11]. Furthermore, since we apply their algorithm to the linearization problem, we only use the first two steps of the algorithm. The details describing the Haar orthogonal series estimates, which can be used in the implementations, are presented in the Appendix. The two-step counterpart of the Dette's algorithm above has the following form:

1) Estimate the nonlinearity using any Haar wavelet-based regression function estimate (*cf.* (10)-(13) in Appendix). For instance, the *order-statistic (OS)* one, where

$$\hat{m}(u) = \sum_{n=0}^{2^M-1} \hat{\alpha}_{Kn}\varphi_{Kn}(u), \text{ where } \hat{\alpha}_{Kn} = \sum_{k=1}^{N} y_k \int_{u_{k-1}}^{u_k} \varphi_{Kn}(u)\, du. \tag{5}$$

2) Given the initial estimate $\hat{m}(u)$, the Haar estimate (of a histogram-like shape) of the output density has a form

$$\left(\hat{m}^{-1}\right)'(u) = \sum_{n=0}^{2^M-1} \hat{a}_{Mn}\varphi_{Mn}(u), \text{ where } \hat{a}_{Mn} = \frac{1}{N}\sum_{k=1}^{N} \varphi_{Mn}\left(\hat{m}\left(\tfrac{k}{N}\right)\right). \tag{6}$$

Subsequently, the Haar estimate of the inverse of the nonlinearity is then computed from $\left(\hat{m}^{-1}\right)'(u)$ using the empirical distribution-like formula (*cf.* (2) and (4))

$$\hat{m}^{-1}(u) = \int_0^u \left(\hat{m}^{-1}\right)'(v)\, dv. \tag{7}$$

In (5) and (6), K and M are the respective scale factors of the Haar series estimates.

Remark 1: Essentially, we compute the density estimate from the estimate of the nonlinearity and then integrate it as in the Dette's algorithm. Nevertheless, exploiting a simple and explicit form of the Haar function, we can readily improve the computation burden using the already integrated scaling functions (*viz.* their indefinite integrals; *cf.* (7))

$$\hat{m}^{-1}(u) = \sum_{n=0}^{2^M-1} \hat{a}_{Mn} \int_0^u \varphi_{Mn}(v)\, dv = \sum_{n=0}^{2^M-1} \hat{a}_{Mn}\Phi_{Mn}(u), \tag{8}$$

where $\Phi_{Mn}(u)$ is the indefinite integral of the scaling function $\varphi_{Mn}(u)$, *i.e.* the function of the form

$$\Phi_{Mn}(u) = 2^{-\frac{M}{2}} \begin{cases} 0 & if \qquad\quad 2^M u < n \\ \frac{2^M u - n}{2^M} & if \quad n \le 2^M u < n+1 \\ 1 & if \quad n+1 \le 2^M u \end{cases}.$$

Note that the normalization factors $2^{-\frac{M}{2}}$ in $\Phi_{Mn}(u)$ and $2^{\frac{M}{2}}$ in $\varphi_{Mn}(u)$ compensate each other. In a similar manner, the integration in (5) can be replaced by subtraction of the indefinite integrals $\Phi_{Kn}(u)$ – see Appendix B