Adaptive Savitzky-Golay Filtering in Non-Gaussian Noise

Arlene John, *Graduate Student Member, IEEE*, Jishnu Sadasivan, and Chandra Sekhar Seelamantula, *Senior Member, IEEE*

***Abstract*—A Savitzky-Golay (SG) filter, widely used in signal processing applications, is a finite-impulse-response low-pass filter obtained by a local polynomial regression on noisy observations in the least-squares sense. The problem addressed in this paper is one of optimal order (or filter length) selection of SG filter in the presence of non-Gaussian noise, such that the mean-squared-error (or risk) between the underlying clean signal and the SG filter estimate is minimized. Since mean-squared-error (MSE) depends on the unknown clean signal, direct minimization is impractical. We circumvent the problem within a risk-estimation framework, wherein, instead of minimizing the original MSE, an unbiased estimate of the MSE (which depends only on the noisy observations and noise statistics) is minimized in order to obtain the optimal order. The proposed method gives an unbiased estimate of the MSE considering SG filtering in the presence of additive noise following any distribution with finite first- and second-order statistics and independent of the signal. The SG filter's order and length are optimized by minimizing the unbiased estimate of MSE. The denoising performance of the optimal SG filter is demonstrated on real-world electrocardiogram (ECG) signals as well as signals from the *WaveLab* Toolbox under Gaussian, Laplacian, and Uniform noise conditions. The proposed denoising algorithm is superior to four benchmark algorithms in low-to-medium input signal-to-noise ratio (SNR) regions (-5 dB to 12.5 dB) in terms of the SNR gain.**

***Index Terms*—Savitzky-Golay filter, local polynomial regression, bias-variance trade-off, mean-squared error, Stein's unbiased risk estimate (SURE), generalized unbiased estimate of MSE (GUE-MSE).**

I. Introduction

REAL-WORLD signals such as speech, biomedical and geophysical signals that are acquired or transmitted are practically never noise-free, which causes errors in downstream processing. Hence, signal denoising is an important preprocessing step. Real-world signals are also time-varying in one or more properties, thereby necessitating denoising based on a localized signal model. The localized signal model employed for denoising is based on local polynomial regression.

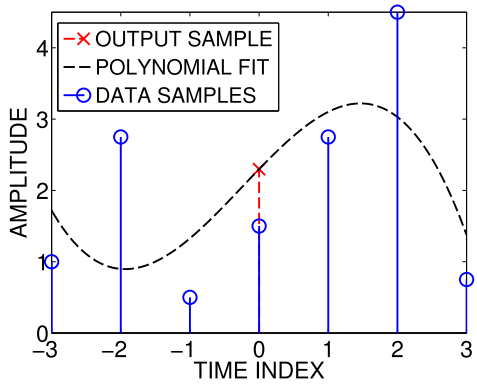

Fig. 1. [Color online] Local least-squares polynomial regression.

The observation model is of the form:

$$x_m = s_m + w_m, \qquad -M \leq m \leq M, \tag{1}$$

where s_m is the clean signal and w_ms are independent and identically distributed (i.i.d.) noise samples with mean zero and variance σ^2. We assume that the underlying continuous domain clean signal is locally smooth. To estimate the signal at an instant, we first fit a polynomial in the least-squares sense to the observations taken from a symmetric window centered at that instant and then evaluate the resulting polynomial value at the center of the window. Local polynomial models have been encountered in various manifestations in the literature – digital smoothing polynomial (DISPO) [1], causal weighted least-squares polynomial estimation [2], locally estimated scatter-plot smoothing (LOESS), and locally weighted scatter-plot smoothing (LOWESS) [3]. An illustration is shown in Fig. 1, wherein a polynomial of order 3 is fitted on 7 observation samples about the origin. The clean signal estimate at the origin (the output sample) is also indicated. The choice of least-squares (LS) regression parameters: window length and polynomial order determine the goodness-of-fit. This process is repeated at each instant to obtain the denoised signal. A detailed review of the technique is available in [4] and [5].

Although such an approach of signal reconstruction may not appear as shift-invariant filtering at first glance, Savitzky and Golay proved that the operation is indeed shift-invariant and equivalent to discrete-time convolution with a finite impulse

Manuscript received December 31, 2019; revised September 30, 2020 and June 28, 2021; accepted August 5, 2021. Date of publication August 24, 2021; date of current version September 15, 2021. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Emilie Chouzenoux. This work was supported by University Grants Commission Center for Advanced Study (UGC-CAS), Phase II. *(Corresponding author: Chandra Sekhar Seelamantula.)*

Arlene John is with the School of Electrical and Electronics Engineering, University College Dublin, Belfield, Dublin 4 D04 V1W8, Ireland (e-mail: arlene.john@ucdconnect.ie).

Jishnu Sadasivan and Chandra Sekhar Seelamantula are with the Department of Electrical Engineering, Indian Institute of Science, Bangalore 560012, Karnataka, India (e-mail: jishnuns@gmail.com; csseelamantula@gmail.com).

Digital Object Identifier 10.1109/TSP.2021.3106450

response with length equal to the window length [6]. The filters are referred to as Savitzky-Golay (SG) filters in literature. Algorithms similar to SG filtering were explored in [7]–[10]. Schafer characterized the frequency-domain behavior of the SG filter and provided an approximate formula for the 3 dB cutoff frequency of the SG filter in terms of the impulse response length and the polynomial order, and showed that they are inversely related [11], [12]. Schafer also showed that SG filtering is linear and shift-invariant, with a Type-1 finite impulse response (FIR).

Ideally, the SG filter should be designed to retain the shape of the original signal, that is, the bias should be small, together with sufficient noise suppression, i.e., low variance. At discontinuities, a longer filter would result in oversmoothing, causing loss of temporal structure of the signal. Hence, the mean-squared-error (MSE) between the estimate and the clean signal would be dominated by the bias. On the other hand, a shorter filter would not cause significant smoothing, thereby reducing the bias, but increased variance. For filters with a variable order but with a fixed filter length, the bias-variance behavior is opposite to that of filtering with a fixed order and varying filter length. Consequently, the optimal choice of filter length for a fixed order or order for a fixed filter length, which minimizes the MSE, must be made at each instant. This is the ubiquitous problem of *bias-variance trade-off* [13], [14]. Estimation of optimal SG filters is equivalent to the optimum *bandwidth selection* problem, in which one selects the optimum 3 dB cutoff frequency of SG filters at each point of reconstruction. The SG filter parameters are optimized at each instant, and hence, the process could be thought of as an adaptive filtering method. Due to the instantaneous optimization, the properties of linearity and shift-invariance are compromised in favor of the MSE-optimal denoising performance.

The optimal filter parameters correspond to the minimum MSE. However, direct minimization of the MSE results in parameters that are functions of the underlying unknown clean signal. We circumvent the problem using a risk estimation framework, wherein instead of minimizing the actual risk, an unbiased estimate of the risk is minimized, which depends only on the observations and noise statistics. In a seminal work, Stein [15] had derived such an unbiased estimate considering Gaussian observations. This gave rise to Stein's unbiased risk estimate (SURE), which has been used successfully for a variety of denoising and deconvolution applications [16]–[21]. Eldar derived a SURE counterpart for distributions from the exponential family [16]. In this paper, we derive a generalized unbiased estimate of MSE (GUE-MSE) without any specific distributional assumption on noise, considering the denoising function to be of the form $\boldsymbol{Gx}$ where $\boldsymbol{G} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{x} \in \mathbb{R}^n$ to be the noisy observation, which is applicable for the SG filtering scenario.

A. Recent Applications of the SG Filter

Candan and Inan [22] proposed a unified SG filter design approach that encompasses applications such as smoothing, differentiation, integration, and fractional delay, by posing the SG filter design problem as one of obtaining the minimum-norm solution to an underdetermined system of equations. Wayt and Khan [23] proposed an integrated SG filter that smooths noise, while estimating values of the signal and its first $k-1$ non-zero polynomial derivatives from the signal denoised using an SG filter of order k by using an inverse Taylor-series approach. Shmaliy *et al.* [24] proposed an unbiased FIR smoothing filter for discrete-time polynomial signals considering a state-space representation, where the filter gains are represented using a polynomial.

Gan *et al.* [25] utilized the SG filter as a prefilter in the design of a drive system dynamics compensator for a mechanical emulator. Hargittai [26] showed that the SG filter is effective at reducing noise in electrocardiogram (ECG) signals while preserving clinically relevant information. Schettino *et al.* [27] employed a second-order differentiator based on SG filter for detecting saturation regions in current transformers. Seo *et al.* [28] introduced SG filtering to obtain the envelope of the vibro-acoustic signal and to extract the arcing signal from the noisy data obtained through a measurement system installed on a transformer's on-load tap changer. King *et al.* [29] used the SG filter to detect absorption band positions from reflectance spectra of vegetation classes. Chinrungrueng and Toonkum [30] proposed two-dimensional (2-D) weighted SG filters for real-time speckle noise reduction and coherence enhancement of ultrasound images. Sa-Ing *et al.* [31], [32] also employed the SG filter to remove speckle noise in ultrasound imaging. Jose *et al.* [33] applied an SG filter for determining the orientation of ridges in images by computing partial derivatives. Koluguri *et al.* [34] proposed a 2-D SG-filter based denoising scheme to enhance spectrograms for robust bird sound detection.

Ramirez *et al.* [35] considered certain properties of the mean-squared value between the observation and the estimate to select the optimum window-length. However, the computation of the mean-squared value requires multiple noisy realizations. Krishnan and Seelamantula [36] proposed a SURE-optimized SG filter for 1-D signal denoising. The SURE formulation is specific to Gaussian noise. To demonstrate the denoising capability of the SURE-optimized SG-filter, ECG signal denoising in white Gaussian noise was considered and it was shown that the SG-filter outperforms wavelet-based denoising techniques. They also proposed a SURE-optimal 2-D SG filter for image denoising [37]. Menon and Seelamantula [38] proposed a robust SG filter design framework in which the ℓ_1 cost is minimized using iteratively reweighted least-squares (IRLS) technique to determine the polynomial coefficients.

B. Contributions of This Paper

In [36], [37], we successfully used SURE to solve the optimum order (or filter length) selection problem. However, as previously discussed, SURE can be used only when the observations are Gaussian distributed. In this paper, we develop an unbiased estimate of MSE for linear denoising functions (Section II). Since the proposed unbiased estimate of the MSE does not assume a distribution of the additive noise except that the first- and second-order moments are finite, we refer to it as the generalized unbiased estimate of MSE (GUE-MSE). Since the denoising operation performed by the SG filter is linear in

the observations, it becomes easy to derive the corresponding GUE-MSE (Section III). Both optimum order selection and filter length selection problem can be addressed by minimizing GUE-MSE and a regularized GUE-MSE (Sections III-C and III-E). The proposed GUE-MSE framework is capable of estimating the optimum SG filter even in non-Gaussian noise conditions.

We compare the denoising performance of the GUE-MSE based SG filter against four benchmarking denoising techniques proposed in [19], [20], [39], and [38] (Section III-G). The signals considered for evaluation are ECG signals (taken from PhysioBank database [40]) and *Piece-Regular* signal (taken from *WaveLab* toolbox [41]). The noise distributions considered are Gaussian, Laplacian, and Uniform. We also analyze the performance of the GUE-MSE based SG filter with simultaneous optimization of window-length and order (Section III-H). The denoising performance and subsequent improvements due to combination of the estimates in the sample neighborhood are also studied (Section III-I). We extend the GUE-MSE to the correlated noise scenario and analyze the denoising performance of the SG filters on signals corrupted with correlated noise (Section III-J). The results show that the optimal SG-filter exhibits a competitive denoising performance.

C. Notations

Boldfaced symbols denote vectors and matrices. For instance, $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{G} \in \mathbb{R}^{n\times n}$ denote a vector and a matrix, respectively. The i^{th} component of the vector $\boldsymbol{x}$ is denoted by x_i. Similarly, the $(i,j)^{\text{th}}$ element of the matrix $\boldsymbol{G}$ is denoted by $g_{i,j}$. The ℓ_2-norm of the vector $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$. The symbol $\mathcal{E}$ denotes the expectation. The symbol $\boldsymbol{h}(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}^n$ is a vector-valued function of a vector argument. The notation $h_i(\boldsymbol{x})$ denotes the i^{th} component of $\boldsymbol{h}(\boldsymbol{x})$.

II. Generalized Unbiased Estimate of MSE (GUE-MSE)

Consider the vector observation in $\mathbb{R}^n$:

$$\boldsymbol{x} = \boldsymbol{s} + \boldsymbol{w}, \tag{2}$$

where $\boldsymbol{w}$ is random with mean zero and diagonal covariance matrix $\text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$, and the signal $\boldsymbol{s}$ is independent of $\boldsymbol{w}$. Let $\boldsymbol{h}(\boldsymbol{x})$ be an estimate of $\boldsymbol{s}$. The objective of the estimation stage is to determine the optimum $\boldsymbol{h}(\boldsymbol{x})$ that minimizes the MSE:

$$\begin{aligned}\mathcal{J} &= \mathcal{E}\left\{\|\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{s}\|^2\right\} \\ &= \mathcal{E}\left\{\boldsymbol{h}(\boldsymbol{x})^{\mathsf{T}}\boldsymbol{h}(\boldsymbol{x})\right\} - 2\mathcal{E}\left\{\boldsymbol{h}(\boldsymbol{x})^{\mathsf{T}}\boldsymbol{s}\right\} + \mathcal{E}\left\{\boldsymbol{s}^{\mathsf{T}}\boldsymbol{s}\right\},\end{aligned} \tag{3}$$

where, $\mathcal{E}$ denotes the expectation and is taken with respect to the noise. The second term in (3) is a function of $\boldsymbol{s}$, which is the unknown clean signal vector, thereby making direct minimization of the cost function infeasible. Next, we state a result that helps to determine $\hat{\mathcal{J}}$, such that $\mathcal{E}\{\hat{\mathcal{J}}\} = \mathcal{J}$. It does not affect the calculations below because the expectations are taken with respect to the noise.

Lemma 1: Consider the additive observation model in (2). Let $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{G}\boldsymbol{x}$ be an estimate of $\boldsymbol{s}$ and $\mathcal{Q} = \mathcal{E}\{\boldsymbol{h}(\boldsymbol{x})^{\mathsf{T}}\boldsymbol{s}\}$. Then, $\hat{\mathcal{Q}} = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x} - \sum_{\ell=1}^{n} g_{\ell,\ell}\sigma_\ell^2$ is an unbiased estimate of $\mathcal{Q}$, where $g_{\ell,\ell}$ is the $(\ell,\ell)^{\text{th}}$ element of $\boldsymbol{G}$.

Proof: Consider the estimator $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{G}\boldsymbol{x}$, where $\boldsymbol{G} \in \boldsymbol{R}^{n\times n}$, resulting in $\mathcal{Q} = \mathcal{E}\{\boldsymbol{h}(\boldsymbol{x})^{\mathsf{T}}\boldsymbol{s}\} = \mathcal{E}\{(\boldsymbol{G}\boldsymbol{x})^{\mathsf{T}}\boldsymbol{s}\}$. We have

$$\begin{aligned}\mathcal{E}\left\{\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{s}\right\} &= \mathcal{E}\left\{\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x}\right\} - \mathcal{E}\left\{\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{w}\right\}, \\ &= \mathcal{E}\left\{\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x}\right\} - \mathcal{E}\left\{\boldsymbol{s}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{w}\right\} - \mathcal{E}\left\{\boldsymbol{w}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{w}\right\}, \\ &= \mathcal{E}\left\{\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x}\right\} - \mathcal{E}\left\{\boldsymbol{w}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{w}\right\}, \\ &= \mathcal{E}\left\{\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x}\right\} - \sum_{\ell=1}^{n}\sum_{k=1}^{n} g_{\ell,k}\underbrace{\mathcal{E}\{w_\ell w_k\}}_{\sigma_k^2\,\delta[\ell-k]}, \\ &= \mathcal{E}\left\{\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x}\right\} - \sum_{\ell=1}^{n} g_{\ell,\ell}\sigma_\ell^2.\end{aligned} \tag{4}$$

Therefore, we have the unbiased estimate $\hat{\mathcal{Q}} = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x} - \sum_{\ell=1}^{n} g_{\ell,\ell}\sigma_\ell^2$, that is, $\mathcal{E}\{\hat{\mathcal{Q}}\} = \mathcal{Q}$. ■

Using $\mathcal{E}\{\|\boldsymbol{h}(\boldsymbol{x})\|^2\} = \mathcal{E}\{\|\boldsymbol{G}\boldsymbol{x}\|^2\}$, $\mathcal{E}\{\|\boldsymbol{s}\|^2\} = \mathcal{E}\{\|\boldsymbol{x}\|^2\} - \sum_{\ell=1}^{n}\sigma_\ell^2$, and Lemma 1, we obtain

$$\begin{aligned}\mathcal{J} = \mathcal{E}\left\{\|\boldsymbol{G}\boldsymbol{x}\|^2\right\} &- 2\mathcal{E}\left\{\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x}\right\} + 2\sum_{\ell=1}^{n} g_{\ell,\ell}\sigma_\ell^2 \\ &+ \mathcal{E}\left\{\|\boldsymbol{x}\|^2\right\} - \sum_{\ell=1}^{n}\sigma_\ell^2,\end{aligned} \tag{5}$$

and consequently, an unbiased estimate of $\mathcal{J}$ is given by

$$\begin{aligned}\hat{\mathcal{J}} = \|\boldsymbol{G}\boldsymbol{x}\|^2 &- 2\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x} + 2\sum_{\ell=1}^{n} g_{\ell,\ell}\sigma_\ell^2 \\ &+ \|\boldsymbol{x}\|^2 - \sum_{\ell=1}^{n}\sigma_\ell^2.\end{aligned} \tag{6}$$

In deriving $\hat{\mathcal{J}}$, we assumed knowledge of the first and second-order statistics only. We did not make any assumption on the distribution of the noise. Hence, we refer to $\hat{\mathcal{J}}$ as the generalized unbiased estimate of the MSE (GUE-MSE).

Corollary 1: Consider the model in Lemma 1. The GUE-MSE is equivalent to Stein's unbiased risk estimate (SURE) [16].

Proof: The estimator $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{G}\boldsymbol{x}$ satisfies the hypotheses of Stein's lemma. More precisely, it is weakly differentiable $\mathcal{E}\{|\frac{\partial h_\ell(\boldsymbol{x})}{\partial x_\ell}|\} < \infty$, $\ell = 1, 2, \ldots, n$ where h_ℓ denotes the ℓ^{th} entry in $\boldsymbol{h}$. The corresponding SURE is given as [16]:

$$\begin{aligned}\hat{\mathcal{J}}_{\text{SURE}} &= \|\boldsymbol{h}(\boldsymbol{x})\|^2 - 2\boldsymbol{h}(\boldsymbol{x})^{\mathsf{T}}\boldsymbol{x} + 2\sigma^2\sum_{\ell=1}^{n}\frac{\partial h_\ell(\boldsymbol{x})}{\partial x_\ell} + \|\boldsymbol{s}\|^2, \\ &= \|\boldsymbol{G}\boldsymbol{x}\|^2 - 2\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x} + 2\sigma^2\sum_{\ell=1}^{n}\frac{\partial h_\ell(\boldsymbol{x})}{\partial x_\ell} + \|\boldsymbol{s}\|^2.\end{aligned} \tag{7}$$

Substituting $\sum_{\ell=1}^{n} \frac{\partial h_\ell(\boldsymbol{x})}{\partial x_\ell} = \sum_{\ell=1}^{n} g_{\ell,\ell}$ in (7) yields

$$\hat{\mathcal{J}}_{\text{SURE}} = \|\boldsymbol{G}\boldsymbol{x}\|^2 - 2\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x} + 2\sigma^2 \sum_{\ell=1}^{n} g_{\ell,\ell} + \|\boldsymbol{x}\|^2 - n\sigma^2,$$

which matches with GUE-MSE (6), considering $\sigma_\ell^2 = \sigma^2, \forall \ell$.■

If the noise is zero-mean and correlated, i.e., the noise has a full covariance matrix, then the GUE-MSE could be modified to account for the noise covariance. The counterpart of Lemma 1 for the correlated noise scenario follows.

Lemma 2: Consider the additive observation model $\boldsymbol{x} = \boldsymbol{s} + \boldsymbol{w}$, where $\boldsymbol{w} \in \mathbb{R}^n$ is random with mean zero and covariance matrix Σ with $(i,j)^{th}$ element denoted by $r_{i,j}$. Let $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{G}\boldsymbol{x}$ be an estimate of $\boldsymbol{s}$ and $\mathcal{Q}_{\text{c}} = \mathcal{E}\{\boldsymbol{h}(\boldsymbol{x})^{\mathsf{T}}\boldsymbol{s}\}$. Then, $\hat{\mathcal{Q}}_{\text{c}} = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x} - \sum_{\ell=1}^{n}\sum_{k=1}^{n} g_{k,\ell} r_{k,\ell}$ is an unbiased estimate of $\mathcal{Q}_{\text{c}}$, where $g_{k,\ell}$ is the $(k,\ell)^{\text{th}}$ element of $\boldsymbol{G}$.

The proof is straightforward and proceeds along the same lines as in case of Lemma 1 with the minor change that $\mathcal{E}\{w_k w_\ell\} = r_{k,\ell}$. Using $\mathcal{E}\{\|\boldsymbol{h}(\boldsymbol{x})\|^2\} = \mathcal{E}\{\|\boldsymbol{G}\boldsymbol{x}\|^2\}, \mathcal{E}\{\|\boldsymbol{s}\|^2\} = \mathcal{E}\{\|\boldsymbol{x}\|^2\} - \sum_{\ell=1}^{n} r_{\ell,\ell}$, and Lemma 2, we obtain the following unbiased estimate of $\mathcal{J}$:

$$\hat{\mathcal{J}} = \|\boldsymbol{G}\boldsymbol{x}\|^2 - 2\boldsymbol{x}^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{x} + 2\sum_{k,\ell=1}^{n} g_{k,\ell} r_{k,\ell} + \|\boldsymbol{x}\|^2 - \sum_{\ell=1}^{n} r_{\ell,\ell}. \tag{8}$$

The objective in this paper is to apply the GUE-MSE obtained to solve the problem of SG filter parameter optimization.

III. Optimal Savitzky-Golay Filter Selection

The SG filter aims to fit a polynomial

$$\hat{s}_m = \sum_{k=0}^{p} a_k m^k, \tag{9}$$

to the observations $\{x_m, -M \le m \le M\}$ by minimizing the least-squares (LS) cost function:

$$C_p = \sum_{m=-M}^{M} (\hat{s}_m - x_m)^2 = \sum_{m=-M}^{M} \left(\sum_{k=0}^{p} a_k m^k - x_m \right)^2. \tag{10}$$

The clean signal estimate at the point of reconstruction is the polynomial evaluated at zero, that is, $\hat{s}_0 = a_0$. To estimate the clean signal at the next instant, one shifts the analysis interval to the right by one sample and repeats the procedure.

To obtain the polynomial coefficients that minimize the cost function C_p, we solve $\frac{\partial C_p}{\partial a_j} = 0$, which results in a set of $p+1$ normal equations with $p+1$ unknowns:

$$\sum_{k=0}^{p} \left(\sum_{m=-M}^{M} m^{j+k} a_k \right) = \sum_{m=-M}^{M} m^j x_m, \;\; j = 0, 1, 2, \ldots, p. \tag{11}$$

The above set of equations can be written in matrix form. Let $\boldsymbol{A}$ be a matrix of dimensions $(2M+1) \times (p+1)$ with $(m,j)^{\text{th}}$ element:

$$A_{m,j} = m^j, \;\; \text{where } m \in [-M, M] \text{ and } j \in [0, p]. \tag{12}$$

Denoting $\boldsymbol{a} = [a_0, a_1, \ldots, a_p]^{\mathsf{T}}$ and $\boldsymbol{x} = [x_{-M}, \cdots, x_{-1}, x_0, x_1, \cdots, x_M]^{\mathsf{T}}$, one can write $\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{a} = \boldsymbol{A}^{\mathsf{T}}\boldsymbol{x}$. The polynomial coefficient vector $\boldsymbol{a}$ is obtained as

$$\boldsymbol{a} = (\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{x} = \boldsymbol{H}\boldsymbol{x}, \tag{13}$$

where $\boldsymbol{H} = (\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathsf{T}}$, which is independent of the input samples. The coefficient a_p is the p^{th} row of $\boldsymbol{H}$ multiplied with $\boldsymbol{x}$. Therefore, in order to obtain the output at $m = 0$, that is a_0, only the first row of $\boldsymbol{H}$ needs to be computed, which leads to the convolution equation $\hat{s}_m = \sum_{v=-M}^{M} h_{m-v} x_v$, the output sample at the index 0 is obtained as $a_0 = \hat{s}_0 = \sum_{v=-M}^{M} h_{-v} x_v$ where h_v is the first row element of $\boldsymbol{H}$ [11]. Therefore, one can observe that the elements of the zeroth row of $\boldsymbol{H}$ matrix are essentially the flipped impulse response of a filter called the SG filter, which can be used to implement LS regression as a convolution [11]. A detailed discussion about the design and the properties of SG filter is presented in [11].

A. Bias-Variance Trade-off

The measurements $\{s_m\}$ denote the samples of the continuous-time signal $s(t)$ [42]. The basis for SG filtering is that the underlying continuous-time signal $s(t)$ is smooth and permits a local polynomial approximation. Asymptotically, when the number of observation samples in the data window is sufficiently large, the expressions for bias and variance of the estimator are given by [4], [36]:

$$\text{Bias}\{\hat{s}(t)\} \approx \begin{cases} \dfrac{s^{(p+1)}(t)}{(p+2)!}\left(\dfrac{L}{2}\right)^{p+1}, & \text{if } p \text{ is odd,} \\ \dfrac{s^{(p+2)}(t)}{(p+3)!}\left(\dfrac{L}{2}\right)^{p+2}, & \text{if } p \text{ is even, and} \end{cases}$$

$$\text{Var}\{\hat{s}(t)\} = \frac{\sigma^2}{L},$$

respectively, where $s^{(q)}(t)$ denotes the q^{th} derivative of s at t, p is the polynomial order, L denotes the duration of the window specified symmetrically about the point of reconstruction, and σ^2 is the noise variance. The trade-off between bias and variance is evident from these expressions. For a fixed order, as L increases, the bias increases and variance decreases, and vice versa. This is intuitive because, for a fixed order, as the filter length increases, the filter cutoff frequency decreases [11], the noise gets smoothed, implying less variance, whereas fine variations present in the clean signal also undergo smoothing, implying greater bias. If the filter length decreases, the noise also gets modeled (increased variance), although fine signal variations can be captured (implying reduced bias). A similar argument holds for a fixed filter length and variation with respect to the order. The bias-variance trade-off as a function of the filter length is illustrated in Fig. 2, where the clean signal is

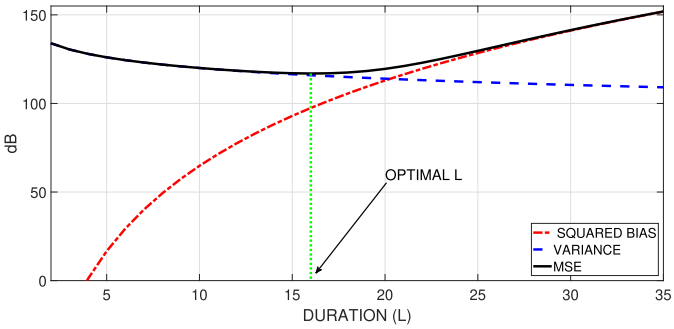

Fig. 2. [Color online] Illustration of the bias-variance trade-off as a function of the filter length.

a fourth-order polynomial, and the estimation is done using third-order regression.

For denoising a time-varying signal corrupted with noise, it becomes important to adapt the filter length and the order of the SG filter such that the MSE is minimized at every instant. Effectively, the goal is to obtain p and/or M such that $\mathcal{R} = \mathcal{E}\left\{\frac{1}{2M+1}\sum_{m=-M}^{M}(\hat{s}_m - s_m)^2\right\}$ is minimized. Since $\mathcal{R}$ depends on the unknown clean signal, direct minimization is not feasible. Hence, in order to obtain the optimum SG filter parameters, we use a risk estimation framework, where instead of minimizing $\mathcal{R}$, we minimize an unbiased estimate of $\mathcal{R}$ following the results presented in Section II.

B. GUE-MSE for the SG Filter

From (9), (13), and from the definition of the matrix $\boldsymbol{A}$ given in (12), we obtain

$$\hat{\boldsymbol{s}}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{H}\boldsymbol{x}. \tag{14}$$

Using (6) and (14), the GUE-MSE for the SG filter is given as follows:

$$\hat{\mathcal{R}} = \frac{1}{2M+1}\left(\|\boldsymbol{A}\boldsymbol{H}\boldsymbol{x}\|^2 - 2\boldsymbol{x}^{\mathsf{T}}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{x} + 2\sigma^2\sum_{\ell=-M}^{M}(\boldsymbol{H}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}})_{\ell,\ell} + \|\boldsymbol{x}\|^2\right) - \sigma^2, \tag{15}$$

where the additive noise is assumed to be i.i.d. with standard deviation σ and mean zero. In other words, $\sigma_\ell^2 = \sigma^2, \forall \ell$.

C. Optimum Order Selection

Algorithm 1 determines the optimum order LS fit at the point of reconstruction n_0 such that the GUE-MSE (15) between the clean signal and its estimate is minimized, where $\boldsymbol{x} = [x_{(n_0-M)} \cdots x_{n_0} \cdots x_{(n_0+M)}]^{\mathsf{T}}$. The optimal order p_{opt} is the one corresponding to the minimum $\hat{\mathcal{R}}$ value. Since the maximum polynomial order is $p_{\max}$, we fix $2M > p_{\max}$. If the number of samples is less than $p_{\max}+1$, the matrix $\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}$ in (13) will be singular.

Fig. 3 shows GUE-MSE and MSE computed for different orders at instant 800 of a *Piece-Regular* (PR) signal, corrupted

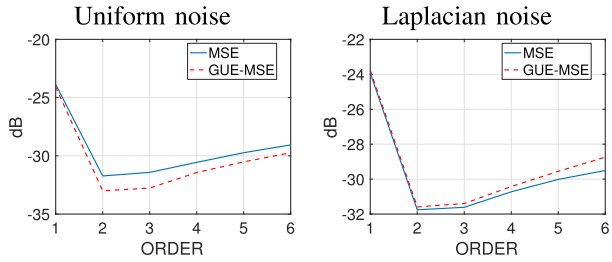

Fig. 3. [Color online] Plots of GUE-MSE and MSE versus order at the time instant 800 for a *Piece-Regular* (PR) signal corrupted with Uniform and Laplacian noise at input SNR 15 dB ($M = 15$).

Algorithm 1: Order Selection Algorithm: Calculate Optimum Order p_{opt} at an Instant n_0.

Require: $2M > p_{\max}$
$p \leftarrow p_{\min}$
while $p \leq p_{\max}$ **do**
  Employ least-squares fit over $[(n_0 - M), (n_0 + M)]$
  Evaluate $\hat{\mathcal{R}}$ using (15)
  $p \leftarrow p + 1$
end while
$p_{\text{opt}} = \arg\min_p \hat{\mathcal{R}}(p)$

with Uniform and Laplacian noises (input SNR of 15 dB). The results are shown after averaging over 100 independent noise realizations. The MSE and GUE-MSE match closely, and the order at which the minimum occurs is the same in both cases.

The results for a *Piece-Regular* (PR) signal corrupted with noise are illustrated in Fig. 4, corresponding to Uniform and Laplacian noise. The parameters are $p_{\min} = 1$, $p_{\max} = 5$, and $M = 15$. The optimum order of the SG filter at each instant and the corresponding denoised signals are also shown. We observe that, for both noise types, the optimum order SG filters obtained by minimizing GUE-MSE results in noise suppression and preservation of sharp transitions present in the signal. At the transitions, the optimal order is high, and over intervals where the signal is relatively smooth, the optimal order is small. This behavior reflects the capability of the algorithm to resolve the bias-variance trade-off at each instant.

Consider time instants from 1500 to 1750 in Fig. 4 (region highlighted by green boxes). Although the signal is smooth in this interval, one can observe the presence of spurious high values in the optimum order estimate (cf. Figs. 4(c) and (d)), which is because of the sample size. In order to overcome this problem, we consider regularization.

D. The Regularized GUE-MSE Approach

The denoising performance of an algorithm can be improved by incorporating a suitable regularizer [16]. To penalize spurious transitions, we propose the regularized risk:

$$\mathcal{R}_\lambda = \mathcal{R} + \frac{\lambda}{2M+1}\mathcal{E}\left\{\sum_{m=-M}^{M}\left(\mathcal{L}\left\{\hat{s}_m(\boldsymbol{x})\right\}\right)^2\right\}, \tag{16}$$

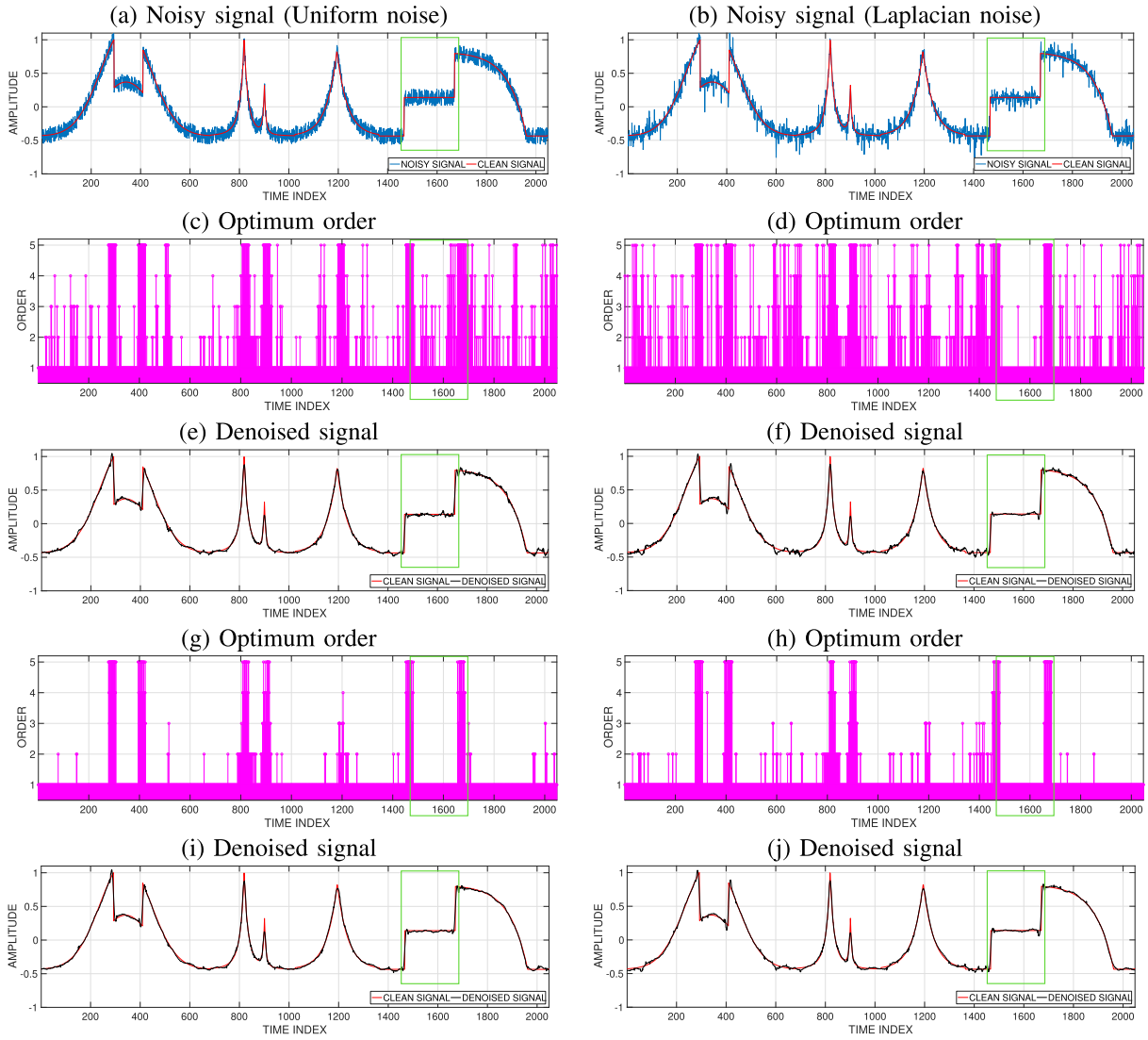

Fig. 4. [Color online] Denoising using order-optimized SG filters obtained by minimizing GUE-MSE and regularized GUE-MSE: (a) and (b) show a Piece-Regular signal corrupted with Uniform noise and Laplacian noise (input SNR 15 dB), respectively; (c) and (d) show the optimum order at each instant that minimizes GUE-MSE; the corresponding denoised signals are shown in (e) and (f), respectively; and (g) and (h) show the optimum order at each instant that minimizes regularized GUE-MSE; the corresponding denoised signals are shown in (i) and (j), respectively. $p_{\min} = 1$, $p_{\max} = 5$, $M = 15$.

where $\mathcal{L}$ denotes the regularization operator. We choose $\mathcal{L}$ as the first derivative operator, corresponding to which

$$\hat{\mathcal{R}}_\lambda = \hat{\mathcal{R}} + \frac{\lambda}{2M+1} \sum_{m=-M}^{M} \underbrace{\left(\frac{\partial \hat{s}_m(\boldsymbol{x})}{\partial x_m} \right)^2}_{(\boldsymbol{AH})^2_{m,m}}. \tag{17}$$

The regularizer based on the ℓ_2-norm of the first derivative measures the sensitivity of the estimate $\hat{s}_m$ to perturbations in the observation caused by noise. Penalizing the sensitivity would enforce a certain amount of continuity and reduce the variance. The optimum SG filter parameters are obtained by minimizing $\hat{\mathcal{R}}_\lambda$ instead of $\hat{\mathcal{R}}$. We set $\lambda = 12\sigma^2$, which was determined empirically. Figs. 4(g) and (h) show that, similar to GUE-MSE, regularized GUE-MSE also results in high orders at instants where the signal exhibits sharp transitions. However, regularization prevents spurious high orders in smooth signal regions (cf. Figs. 4(g) and (h)) compared to GUE-MSE without regularization (cf. Figs. 4(c) and (d)). Consequently, the quality of denoising is considerably improved (cf. Figs. 4(i) and (j)).

E. Optimum Filter Length Selection

The next goal is to determine the optimum filter length for a given order p. We need a minimum of $p+1$ samples to ensure that $\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}$ in (13) is non-singular. We use Algorithm 2 at instant n_0 to determine the optimal filter length, where M_{opt} is the optimum half-window length corresponding to the minimum observed $\hat{\mathcal{R}}$. To obtain M_{opt} for a fixed order p, we evaluate $\hat{\mathcal{R}}$ for different values of M ranging from $M_{\min}$ to $M_{\max}$, and choose the value that yields the minimum $\hat{\mathcal{R}}$.

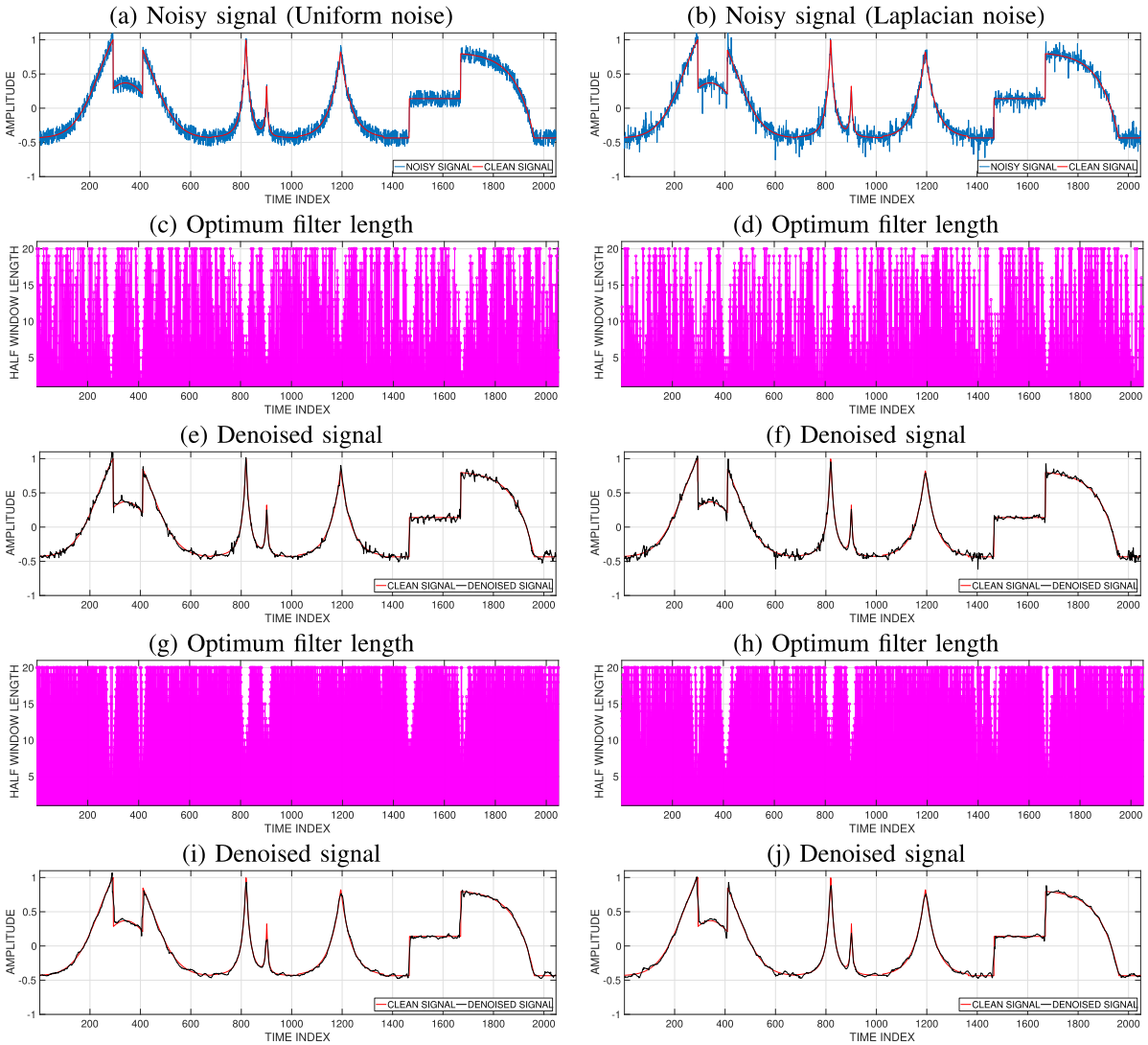

Fig. 5. [Color online] Denoising using filter length-optimized SG filters obtained by minimizing GUE-MSE and regularized GUE-MSE: (a) and (b) show the *Piece-Regular* signal corrupted with Uniform noise and Laplacian noise (input SNR 15 dB), respectively; (c) and (d) show the filter length at each instant that minimizes GUE-MSE, corresponding denoised signals are shown in (e) and (f), respectively; and (g) and (h) show the filter length at each instant that minimizes regularized GUE-MSE, corresponding denoised signals are shown in (i) and (j), respectively. The algorithm parameters are $p = 3$, $M_{\max} = 20$.

We used SG filters obtained using Algorithm 2 to denoise *Piece-Regular* signal corrupted with Uniform noise and Laplacian noise. Figs. 5(a) and (b) show noisy signals (input SNR 15 dB) corresponding to Uniform noise and Laplacian noise, respectively. First, we consider the case without a regularizer. We fix the algorithm parameters $p = 3$ and $M_{\max} = 20$. The optimal half window-length of the SG filter at each instant is shown in Figs. 5(c) and (d), and the corresponding denoised signals are shown in Figs. 5(e) and (f), respectively. From Figs. 5(c) and (d), we observe that, at transitions, the optimal window size is small, whereas in smooth signal regions, the optimal window size assumes the maximum value ($M_{\max}$).

Next, we consider optimum window-length selection by minimizing GUE-MSE with a regularizer. The optimal half window-length of the SG filter obtained by minimizing regularized GUE-MSE at each instant is shown in Figs. 5(g) and (h). Figs. 5(i) and (j) show the denoised signal obtained using optimum SG filter with regularizer. From Figs. 5(i) and (j), we observe that the denoising performance of the optimum SG filter obtained by minimizing regularized GUE-MSE is superior to without.

F. Noise Standard Deviation Estimation

Both GUE-MSE (15) and regularized GUE-MSE (17) are functions of the standard deviation σ of the additive noise. In order to estimate σ, we use a median-based approach. Recall the observation model $x_m = s_m + w_m, -M \leq m \leq M$, where w_m are independent and identically distributed (i.i.d.) with mean zero and variance σ^2. Considering s_m to be largely smooth, except at discontinuities, one can consider the approximation $|x_m - x_{m-1}| \approx |w_m - w_{m-1}|$. Consider the following robust estimator:

$$\hat{\kappa} = \text{median}(|x_m - x_{m-1}|\,; m = 2, 3, \cdots, N),$$

Algorithm 2: Optimum Filter Length Selection: Calculates the Optimum M_{opt} at an Instant n_0

Require: $2M_{\min} > p$
$M \leftarrow M_{\min}$
while $M \leq M_{\max}$ **do**
  Employ least-squares fit over $[(n_0 - M), (n_0 + M)]$
  Evaluate $\hat{\mathcal{R}}$ using (15)
  $M \leftarrow M + 1$
end while
$M_{\text{opt}} = \arg\min_{M} \hat{\mathcal{R}}(M)$

$$\approx \text{median}(|w_m - w_{m-1}|\,; m = 2, 3, \cdots, N). \tag{18}$$

Using the p.d.f. of w_m, the median κ is expressed as a function of σ, that is $\kappa = u(\sigma)$, and σ is estimated as $u^{-1}(\hat{\kappa})$. The median-based estimates for different noise distributions are given below (cf. Appendix for derivations):

- Uniform distribution: $\hat{\sigma} = 0.9853\,\hat{\kappa}$;
- Gaussian distribution: $\hat{\sigma} = 1.0485\,\hat{\kappa}$; and
- Laplacian distribution: $\hat{\sigma} = 1.2337\,\hat{\kappa}$.

G. Comparison With Benchmark Denoising Algorithms

For performance evaluation, we consider a *Piece-Regular* signal taken from *WaveLab* toolbox and two real-world ECG signals taken from the *PhysioBank* database[1], in Gaussian, Laplacian, and Uniform noise conditions. We benchmark the performance of the proposed denoising algorithms against four performant algorithms: (i) A wavelet-based denoising algorithm that uses SURE and a linear expansion of thresholds to model the denoising function (SURE-LET)[2] [19]; (ii) Smooth sigmoidal shrinkage in wavelet domain (SS) [20][3] in which the parameters of the sigmoid that control the degree of attenuation are obtained by optimizing SURE; (iii) Wavelet-thresholding based denoising, where, for an N-length observation sequence, $\sigma\sqrt{2\ln(N)}$ is chosen as the threshold (ST)[4] [39]; and (iv) the robust SG filter that minimizes ℓ_1-norm of the data-fitting error through iteratively reweighted least-squares (IRLS) technique, which is suitable for handling heavy-tailed noise distributions [38]. We consider Daubechies-14 wavelet in SS, SURE-LET, and ST for performance evaluation.

The SG filter based denoising schemes used are:

- GUE-MSE-based filter-length selection for a fixed order (G-FL);
- GUE-MSE-based filter-length selection for a fixed order with regularization (G-FL-R);
- GUE-MSE-based order selection for a fixed filter length (G-O); and
- GUE-MSE-based order selection for a fixed filter length with regularization (G-O-R).

[1]PhysioNet: http://physionet.org/physiobank/database/aami-ec13/.
[2]SURE-LET software: http://bigwww.epfl.ch/demo/suredenoising/.
[3]MATLAB implementation courtesy Pastor *et al.* [20].
[4]WAVELAB toolbox from: http://www-stat.stanford.edu/~wavelab/Wavelab_850/index_wavelab850.html.

In the experiments, for G-FL and G-FL-R, we fix the order $p = 3$ and the maximum half-filter length $M_{\max} = 20$. For G-O and G-O-R, we fixed the half filter-length $M = 15$ and varied the order from $p_{\min} = 1$ to $p_{\max} = 5$. For regularized approaches, we chose $\lambda = 12\sigma^2$.

The results are presented in Fig. 6 after averaging over 100 independent noise realizations. Figs. 6(a), (b), and (c), show the performance of the proposed algorithms in the case of ECG-1 signal corrupted by Gaussian, Laplacian, and Uniform noise, respectively. We observe that for all the noise types considered, G-O-R outperforms the other algorithms for input SNRs ranging from -5 dB to 12.5 dB. Among the algorithms compared, SS shows a higher denoising performance. In the case of Gaussian and Uniform noise, for input SNR range -5 dB to 10 dB, G-O-R exhibits an SNR improvement of 0.5 to 4 dB over SS (cf. Figs. 6(a), and (c)), and for Laplacian noise case G-O-R exhibits an SNR improvement of $8 - 2$ dB over SS (cf. Fig. 6(b)). As the input SNR increases, the SNR gains offered by G-O-R and SS do not differ much.

For low-to-medium input SNR (-5 to 12.5 dB), G-O, G-FL-R, and G-O-R outperform the benchmark denoising algorithms. The denoising trends are similar for the other ECG signal as well (cf. Figs. 6(d)-(f)). The performance comparison in the case of *Piece-Regular* signal is shown in Figs. 6(g)-(i). The trends are similar to those of the ECG signal scenario.

The results in Fig. 6 indicate that the GUE-MSE-based order selection algorithm consistently exhibits the highest performance in the low-to-medium input SNR range ($0 - 12.5$ dB). The SG filter obtained by minimizing regularized GUE-MSE yields a better denoising performance than the one obtained without regularization. We observe that the proposed regularization improves the denoising performance in low SNR regions where GUE-MSE has a high variance. A fixed-order variable-length filter (G-FL) has a performance that is lowest amongst the proposed algorithms and is consistent with the observations in [36], [43]. In most of the cases, G-O and G-FL-R yield nearly equal output SNR. For the G-O-R denoising algorithm, the relative gain compared with the SURE-LET and SS (which are state-of-the-art SURE-based wavelet denoising algorithms) is higher for signals corrupted with Laplacian in comparison with Gaussian and Uniform noise cases. The G-FL-R, G-O, and G-O-R methods outperform the robust SG filtering method for all the noise scenarios considered. The robust SG filter has better performance for Laplacian noise than other noise distributions, which is due to the ℓ_1 data-fitting term. The proposed algorithms perform well in all three noise types, and interestingly, for a given signal and SNR, their performance is similar across different noise distributions.

For ECG signals, preserving both the transitions and the fine structure is crucial from a clinical perspective. Fig. 7 shows the denoised signals obtained by G-O-R and SS (among the benchmarking techniques, SS shows the best performance (cf. Fig. 6)) where the noisy signals are obtained by adding a Laplacian noise and Uniform noise to ECG-1 signal at input SNR 5 dB. We observe that for both noise types, G-O-R exhibits better denoising in terms of preserving the signal transitions and suppressing noise (cf. Figs. 7(b) and (f)) compared with SS (cf. Figs. 7(d) and

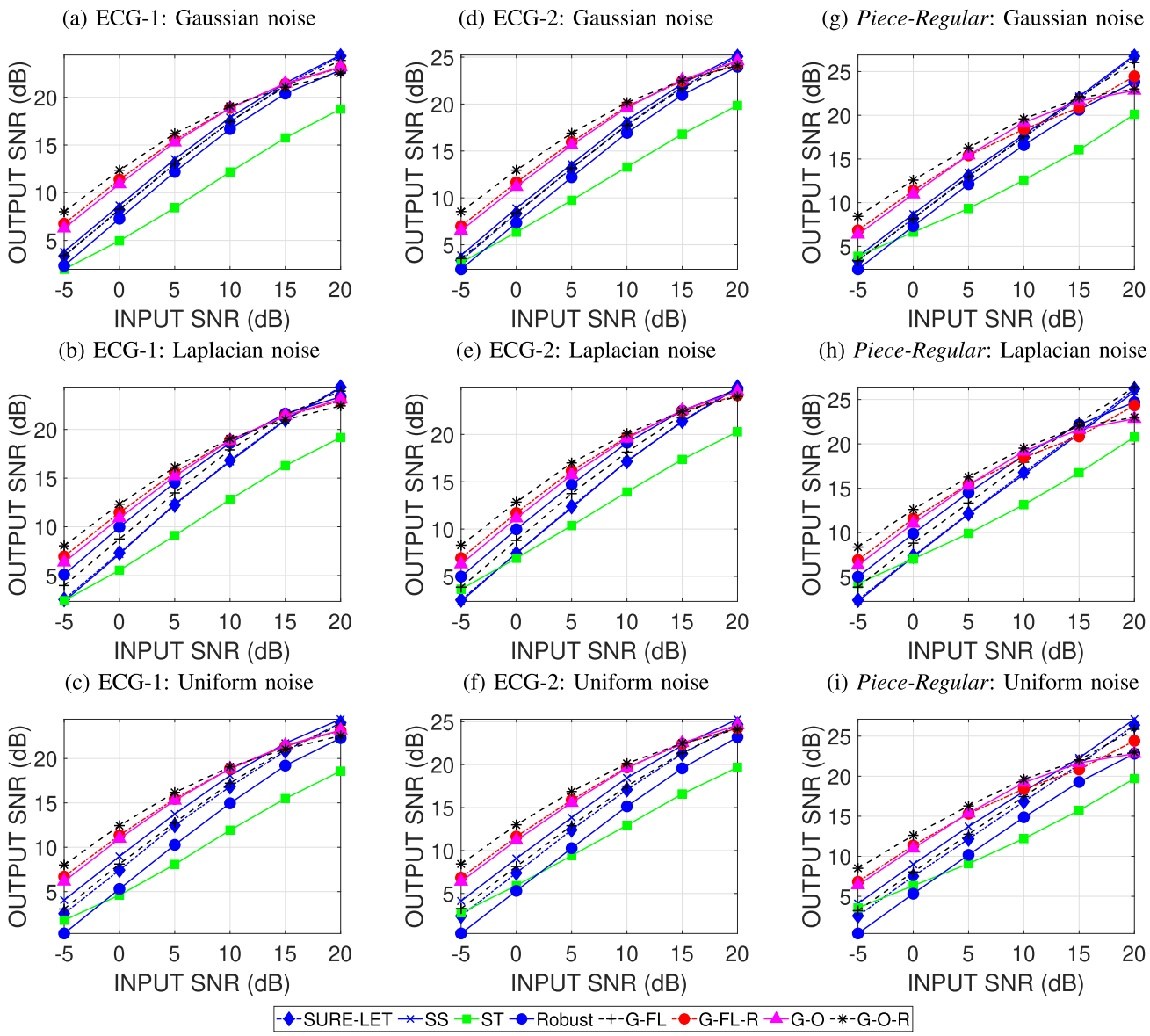

Fig. 6. [Color online] Denoising performance comparison for various signals in different noise conditions.

(h)). Fig. 8 shows the denoising performance for another ECG signal. The variance of the output SNR over the 100 realizations was also calculated. The proposed technique is superior to the state-of-the-art at low SNRs despite the variance. The details are provided in the Supplementary Document. The experimental results validate the use of the GUE-MSE as a surrogate for the MSE and demonstrate the ability of the GUE-MSE based optimal SG filter to act as a competitive denoising technique.

H. Simultaneous Optimization of Order and Window-Length

Next, we perform simultaneous optimization of M and p by minimizing GUE-MSE considering both regularized and unregularized risk scenarios: (i) GUE-MSE-based filter length and order selection method with regularization (G-FL-O-R); and (ii) GUE-MSE-based filter length and order selection method (G-FL-O). We consider half filter-length from $M_{\min} = 10$ to $M_{\max} = 20$ and the order is varied from $p_{\min} = 1$ to $p_{\max} = 7$. The algorithm for the G-FL-O is discussed in Algorithm 3. The algorithm corresponding to G-FL-O-R is obtained by replacing $\hat{\mathcal{R}}$ in Algorithm 3 with $\hat{\mathcal{R}}_\lambda$. Fig. 9 compares the performance of G-FL-O and G-FL-O-R with the G-FL, G-FL-R, G-O, and G-O-R in terms of SNR, where we consider the denoising of ECG 1, ECG 2, and PR signal corrupted with Gaussian, Laplacian, and Uniform noise types. The results presented are averaged over 100 independent noise realizations. Fig. 9 shows that G-FL-O-R outperforms the other algorithms for input SNRs ranging from -5 dB to 12.5 dB for all signals, when the signal is corrupted by Laplacian noise. However, the performance improvement is not significant. The G-FL-O method exhibits a marginal performance improvement over G-O and G-FL methods when denoising a *Piece-Regular* signal in Laplacian noise.

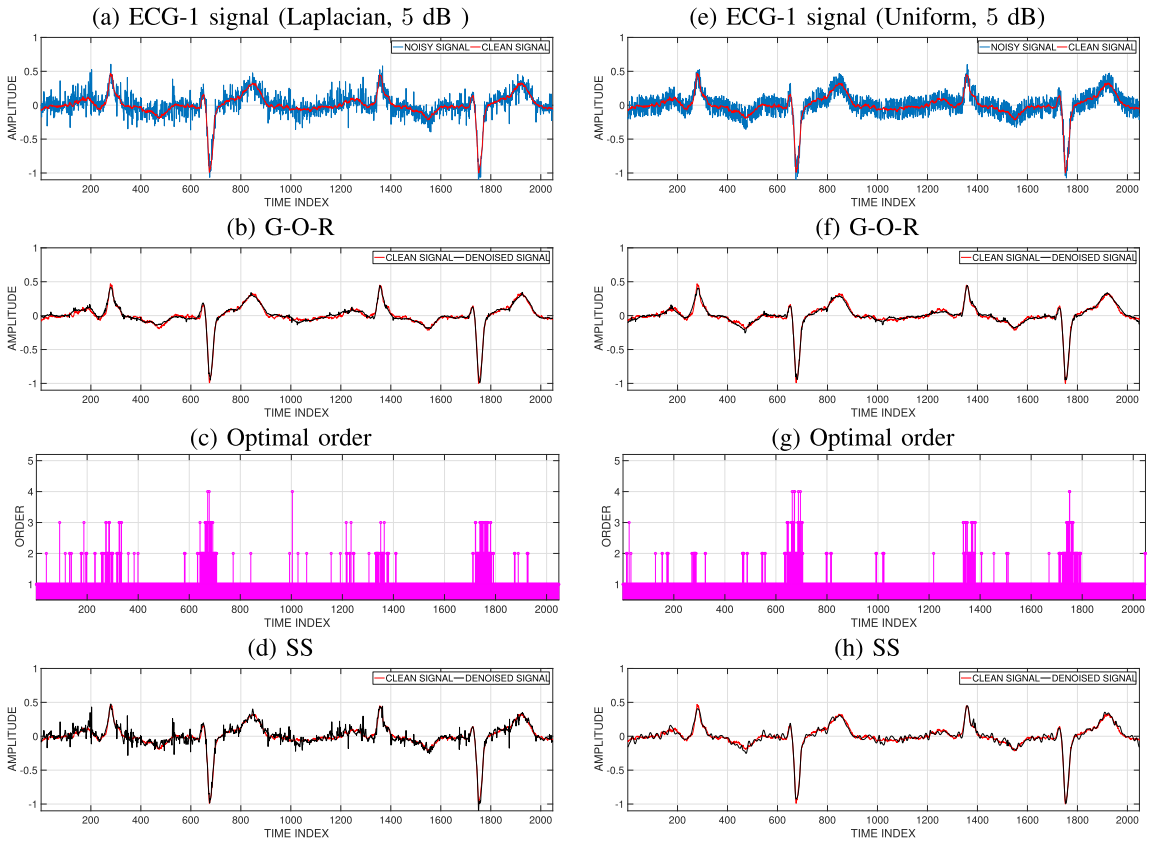

Fig. 7. [Color online] Denoised signals using G-O-R and SS (the best performing benchmarking technique). The first and second columns correspond to the denoising of ECG-1 signal corrupted with a Laplacian and Uniform noise, respectively.

Algorithm 3: Order and window-length selection: Determine optimum order p_{opt} and window length M_{opt} at instant n_0.

Require: $2M_{\min} > p_{\max}$
$p \leftarrow p_{\min}$
while $p \leq p_{\max}$ **do**
  $M \leftarrow M_{\min}$
  while $M \leq M_{\max}$ **do**
    Least-squares polynomial fit over $[n_0 - M, n_0 + M]$
    Evaluate $\hat{\mathcal{R}}$ using (15)
    $M \leftarrow M + 1$
  end while
  $p \leftarrow p + 1$
end while
$p_{\text{opt}}, M_{\text{opt}} = \arg\min_{p,M} \hat{\mathcal{R}}(p, M)$

In all scenarios, G-FL-O-R shows comparable performance to the G-O-R algorithm. Similarly, G-FL-O shows comparable performance to G-O. Simultaneous optimization of both window length and order does not result in a significant improvement over individual optimization.

I. Combining Neighbourhood Estimates

To estimate the clean signal at a given point, we fit an optimum local polynomial to a window of samples considered symmetrically about that point. To further improve the performance, we combine the estimates at a given point obtained from the neighboring windows using a convex combination of the estimates obtained from windows corresponding to M past and M future samples about the point of reconstruction. The optimal estimate from the window centered at the point of interest is given the highest weighting, and on either side, the weight reduces. We used a Gaussian weight kernel with standard deviation of six, normalized to sum of unity. The G-FL, G-FL-R, G-O, and G-O-R schemes with weighting incorporated are denoted as G-FL-W, G-FL-R-W, G-O-W, and G-O-R-W, respectively.

Fig. 10 shows the performance comparison of G-O-W and G-O-R-W, with G-O and G-O-R, where the results presented are averaged over 100 independent noise realizations. From Fig. 10, we observe that in the low SNR regime (-5 dB to 10 dB), G-O-W and G-O-R-W show a better denoising performance compared with the corresponding unweighted versions (G-O and G-O-R, respectively). However, for SNRs greater than 10 dB, weighting the neighborhood estimates did not turn out to be beneficial. Fig. 11 compares G-FL-W and G-FL-R-W with G-FL and G-FL-R, and the trends are similar to that in

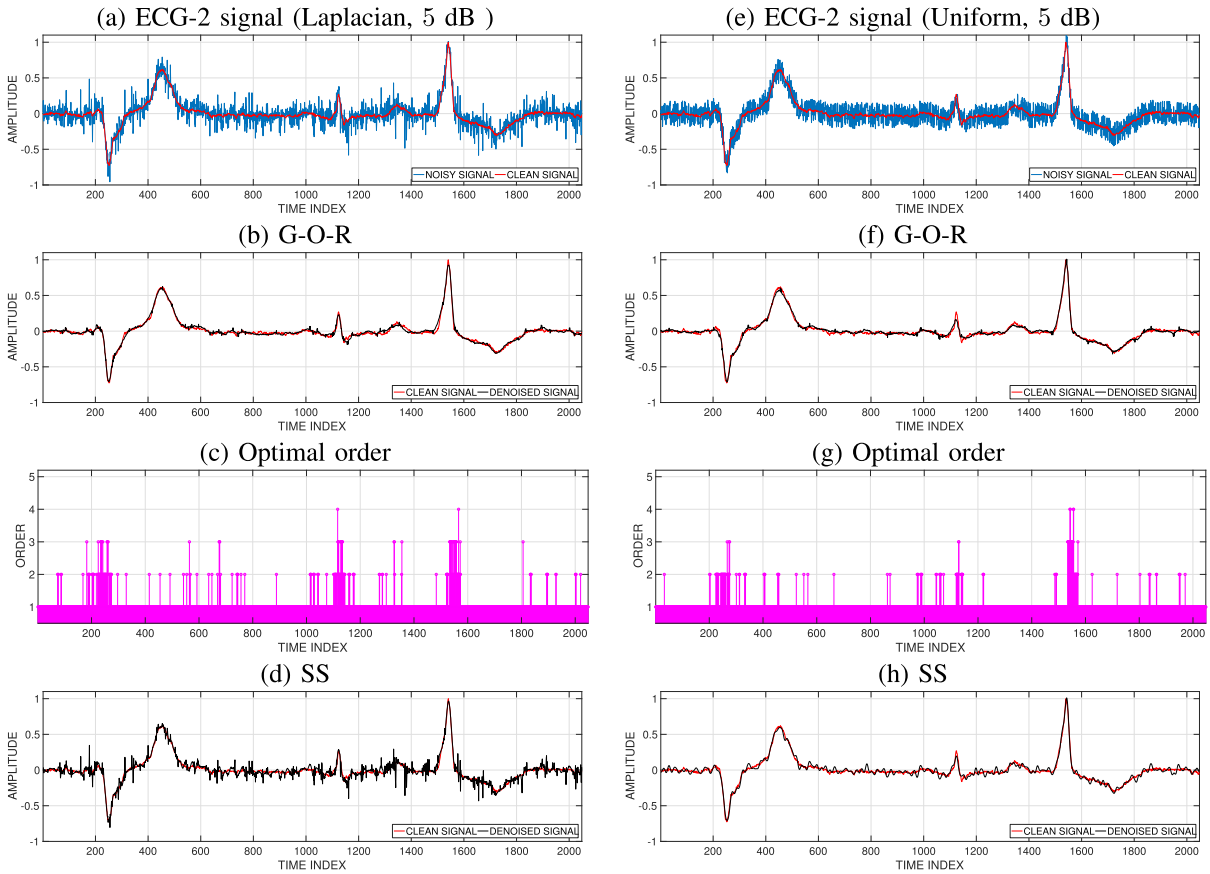

Fig. 8. [Color online] Denoised signals using G-O-R and SS (the best performing benchmarking technique). The first and second columns correspond to the denoising of ECG-2 signal corrupted with a Laplacian and Uniform noise, respectively.

Fig. 10. Weighted averaging offers an advantage in the low SNR regime because it reduces the variance. One could also obtain additional performance improvements by optimizing the choice of the weight function.

J. SG Filtering for Correlated Noise

Consider the GUE-MSE of the SG filter in correlated noise. Substituting $\hat{\boldsymbol{s}}(\boldsymbol{x}) = \boldsymbol{AHx}$ in (8) gives

$$\hat{\mathcal{R}} = \frac{1}{2M+1}\left(\|\boldsymbol{AHx}\|^2 - 2\boldsymbol{x}^{\mathsf{T}}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{x} + 2\sum_{\ell=1}^{n}\sum_{k=1}^{n}(\boldsymbol{H}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}})_{k,\ell}r_{k,\ell} + \|\boldsymbol{x}\|^2 - \sum_{\ell=1}^{n} r_{\ell,\ell}\right). \quad (19)$$

The corresponding regularized GUE-MSE can be written as:

$$\hat{\mathcal{R}}_{\lambda} = \hat{\mathcal{R}} + \frac{1}{2M+1}\sum_{m=-M}^{M}\lambda_m \underbrace{\left(\frac{\partial \hat{s}_m(\boldsymbol{x})}{\partial x_m}\right)^2}_{(\boldsymbol{AH})^2_{m,m}}. \quad (20)$$

In the experiments, we choose $\lambda_m = 12\, r_{m,m}$. We consider the *Piece-Regular* signal and real-world ECG signals, in Gaussian, Laplacian, and Uniform noise, in the following scenarios:

- GUE-MSE-based filter-length selection for a fixed order for correlated noise (G-FL-C);
- GUE-MSE-based filter-length selection for a fixed order with regularization for correlated noise (G-FL-R-C);
- GUE-MSE-based order selection for a fixed filter length for correlated noise (G-O-C); and
- GUE-MSE-based order selection for a fixed filter length with regularization for correlated noise (G-O-R-C).

We add correlated noise $w_m = 0.3w_{m-1} + e_m$ with SNRs varying from -5 dB to 20 dB, and e_m drawn from Gaussian, Laplacian, or Uniform distributions according to the noise scenario. We assume that the covariance matrix of the additive noise is known a priori in order to remove the effect of errors in the estimation of the noise covariance matrix. From Fig. 12 , we observe that in correlated noise also, GUE-MSE-optimized SG filter is able to suppress noise to an appreciable extent.

The algorithms explored in this work and the parameters used are summarised in Table I.

K. Computational Complexity

The matrices $\boldsymbol{A}$ and $\boldsymbol{H}$, which correspond to the flipped filter impulse response coefficients for obtaining the polynomial coefficients, can be precomputed. The computational complexity is dominated by the convolution for a given window length, which can be implemented using the fast Fourier transform (FFT) algorithm. For the G-FL algorithm with fixed order

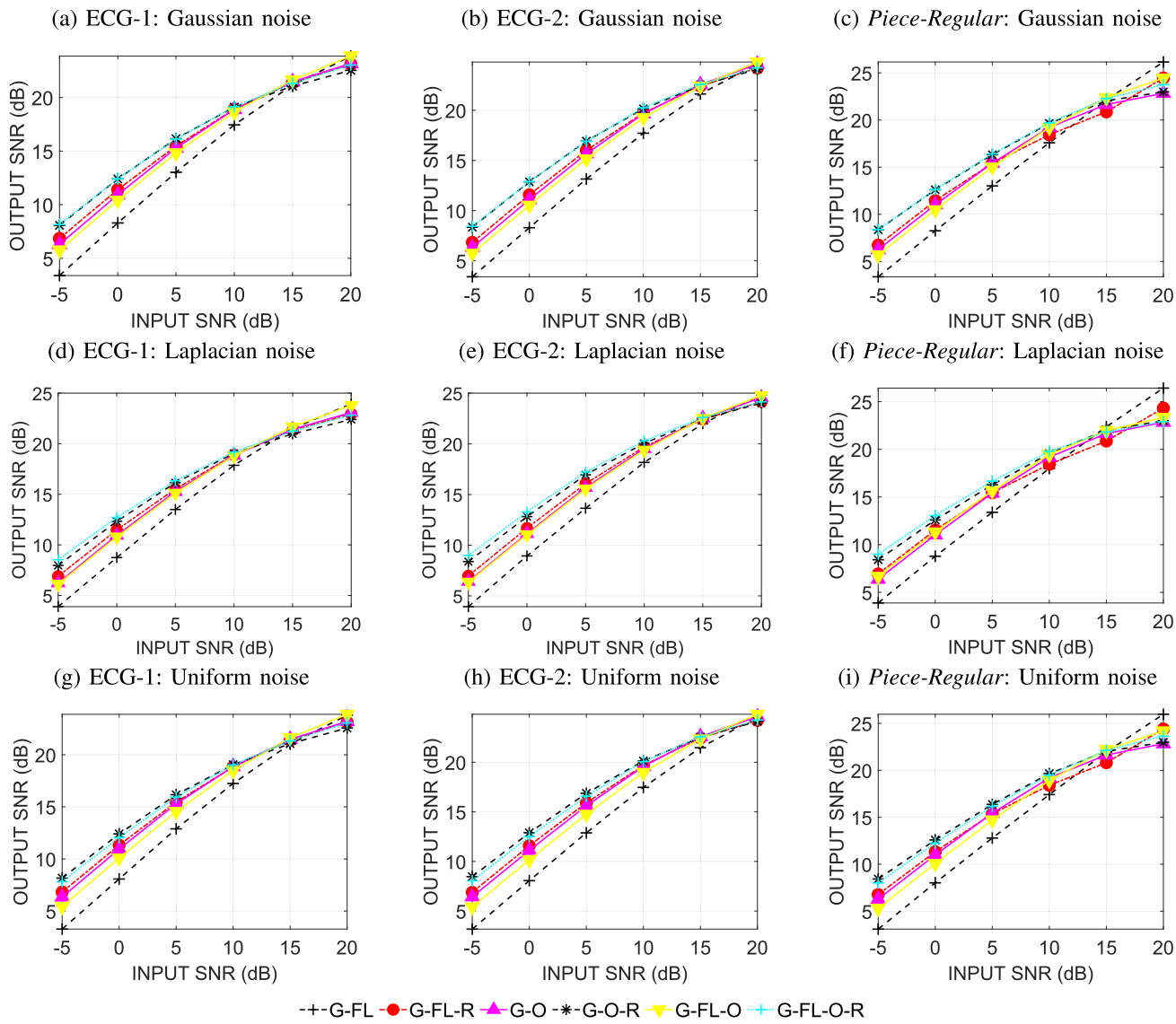

Fig. 9. [Color online] Denoising performance comparison of various schemes under consideration: GUE-MSE based simultaneous optimization of filter length and order (G-FL-O) and G-FL-O with regularization (G-FL-O-R) against GUE-MSE based filter-length-optimization (G-FL), G-FL with regularization (G-FL-R), GUE-MSE based order optimization (G-O), and G-O with regularization (G-O-R) methods.

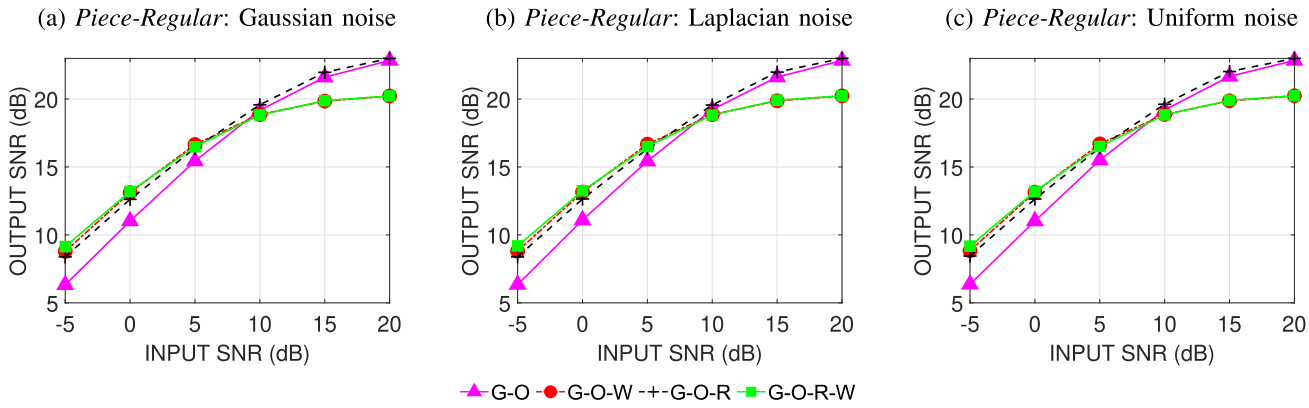

Fig. 10. [Color online] Denoising performance comparison of various schemes: GUE-MSE based optimized-order and neighborhood averaging (G-O-W) and G-O-W with regularizer (G-O-R-W) with GUE-MSE based optimized-order G-O and G-O with regularizer (G-O-R).

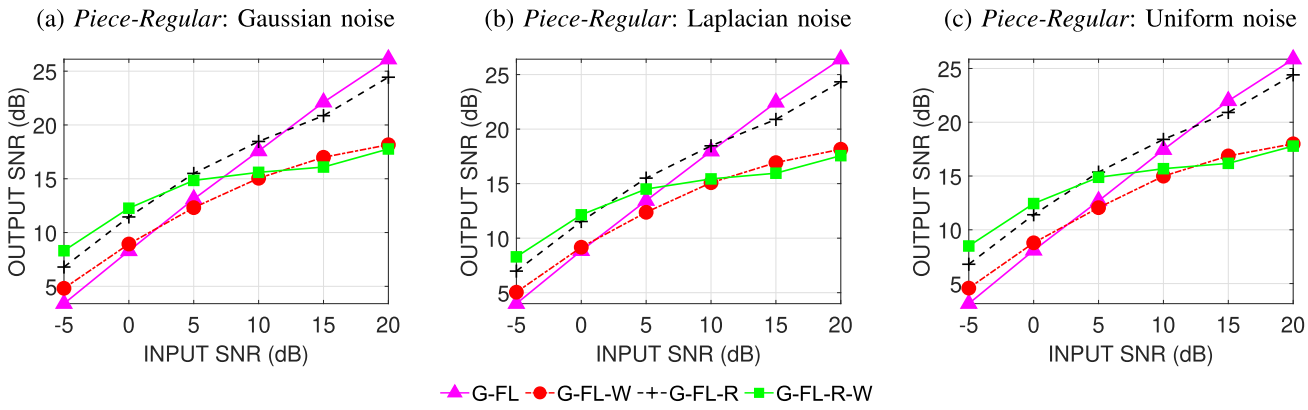

Fig. 11. [Color online] Denoising performance comparison of various schemes: GUE-MSE based filter-length-optimization and neighborhood averaging (G-FL-W) and G-FL-W with regularizer (G-FL-R-W) with GUE-MSE based filter-length-optimization (G-FL) and G-FL with regularizer (G-FL-R).

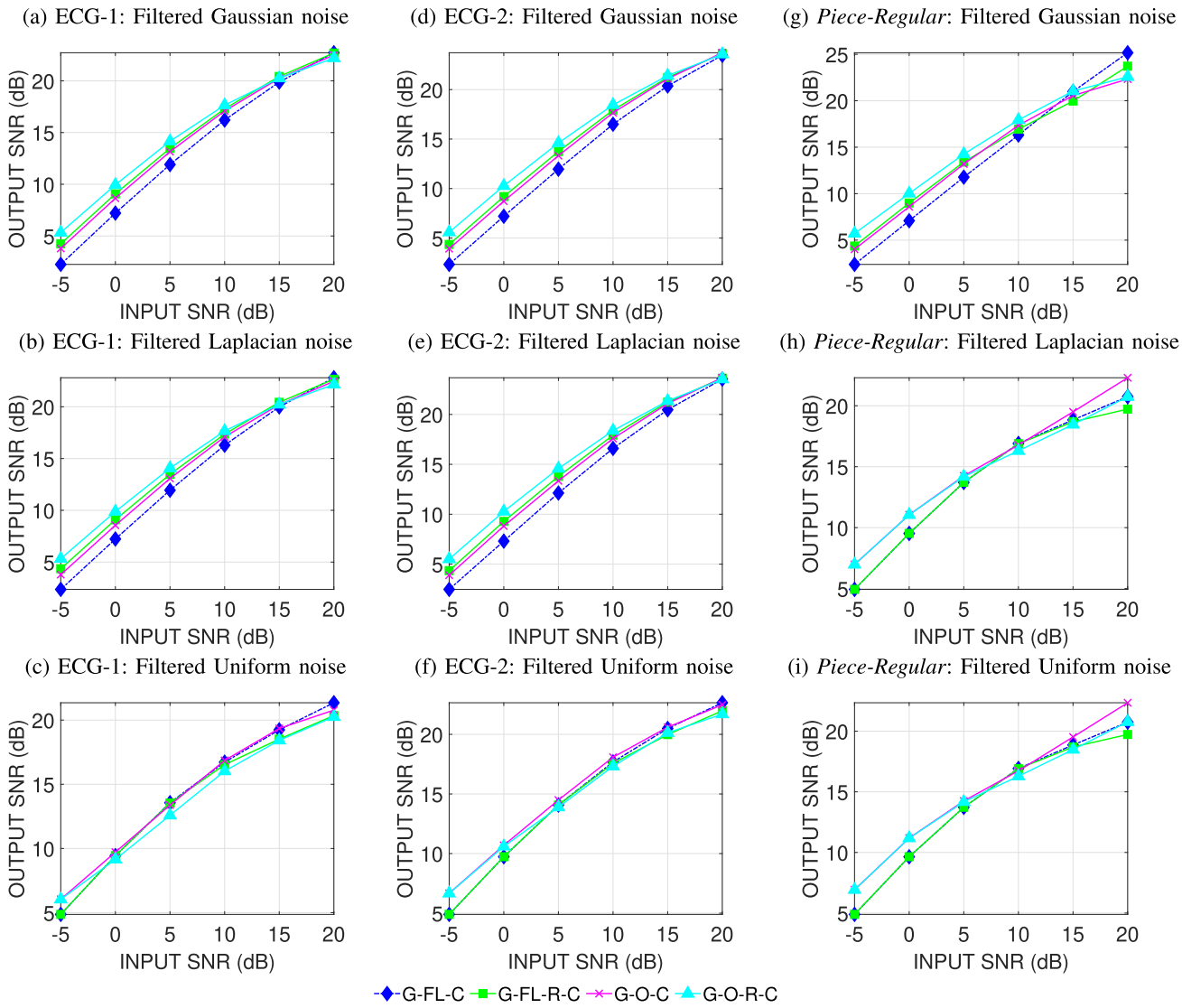

Fig. 12. [Color online] Denoising performance of the GUE-MSE based SG filter in correlated noise. G-FL-C: GUE-MSE based filter-length optimization in correlated noise; G-FL-R-C: G-FL-C with regularization; G-O-C: GUE-MSE based optimized-order filter in correlated noise; G-O-R-C: G-O-C with regularization.

TABLE I
LIST OF ALGORITHMS USED FOR COMPARISON AND PARAMETER SETTINGS

Method	Settings
SURE-LET: SURE-based linear expansion of thresholds	Daubechies-14 wavelet
SS: Smooth sigmoidal shrinkage in wavelet domain	Daubechies-14 wavelet
ST: Wavelet soft-thresholding	Daubechies-14 wavelet
Robust SG filter: ℓ_1 data-fitting error	$M = 9, p = 2$
G-FL: GUE-MSE-based filter-length selection for a fixed order	$p = 3, M_{\text{min}} = 2, M_{\text{max}} = 20$
G-FL-R: GUE-MSE-based filter-length selection for a fixed order with regularization	$p = 3, M_{\text{min}} = 2, M_{\text{max}} = 20,\ \lambda = 12\,\sigma^2$
G-O: GUE-MSE-based order selection for a fixed filter length	$M = 15, p_{\text{min}} = 1, p_{\text{max}} = 5$
G-O-R: GUE-MSE-based order selection for a fixed filter length with regularization	$M = 15, p_{\text{min}} = 1, p_{\text{max}} = 5, \lambda = 12\sigma^2$
G-FL-O: GUE-MSE-based filter length and order selection	$M_{\text{min}} = 2, M_{\text{max}} = 20, p_{\text{min}} = 1, p_{\text{max}} = 5$
G-FL-O-R: GUE-MSE-based filter length and order selection with regularization	$M_{\text{min}} = 2, M_{\text{max}} = 20, p = 1:5, \lambda = 12\sigma^2$
G-FL-W: GUE-MSE-based filter-length selection for a fixed order and neighborhood weighting	$p = 3, M_{\text{min}} = 2, M_{\text{max}} = 20$, Gaussian weights (mean = 0, variance = 6)
G-FL-R-W: GUE-MSE-based filter-length selection for a fixed order with regularization and neighborhood weighting	$p = 3, M_{\text{min}} = 2, M_{\text{max}} = 20, \lambda = 12\sigma^2$, Gaussian weights (mean = 0, variance = 6)
G-O-W: GUE-MSE-based order selection for a fixed filter length and neighborhood weighting	$M = 15, p_{\text{min}} = 1, p_{\text{max}} = 5$, Gaussian weights (mean = 0, variance = 6)
G-O-R-W: GUE-MSE-based order selection for a fixed filter length with regularization and neighborhood weighting	$M = 15, p_{\text{min}} = 1, p_{\text{max}} = 5, \lambda = 12\sigma^2$, Gaussian weights (mean = 0, variance = 6)
G-FL-C: GUE-MSE-based filter-length selection for a fixed order for colored noise	$p = 3, M_{\text{min}} = 2, M_{\text{max}} = 20, r_{k,l} = \sigma^2 \frac{0.3^{\lvert k-l\rvert}}{1-0.3^2}$
G-FL-R-C: GUE-MSE-based filter-length selection for a fixed order with regularization for colored noise	$p = 3, M_{\text{min}} = 2, M_{\text{max}} = 20, \lambda_m = 12\, r_{m,m} = \frac{12\sigma^2}{1-0.3^2}$
G-O-C: GUE-MSE-based order selection for a fixed filter length for colored noise	$M = 15, p_{\text{min}} = 1, p_{\text{max}} = 5, r_{k,l} = \sigma^2 \frac{0.3^{\lvert k-l\rvert}}{1-0.3^2}$
G-O-R-C: GUE-MSE-based order selection for a fixed filter length with regularization for colored noise	$M = 15, p_{\text{min}} = 1, p_{\text{max}} = 5, \lambda_m = 12\, r_{m,m} = \frac{12\sigma^2}{1-0.3^2}$

$p = 3$ and half-filter length M_r varied from M_{min} to M_{max} at the r^{th} iteration, the algorithmic complexity is $4\mathcal{O}((2M_r + 1)\log_2(2M_r + 1)) + 9(2M_r + 1) + 4$ per sample. The total computational load can be obtained by summing up these values over the filter lengths at various iterations, and then multiplying by the number of observations. For the G-O algorithm with fixed window length of $2M + 1 = 31$, and the order p_r varied from p_{min} to p_{max} at the r^{th} iteration, the algorithmic complexity is $(p_r + 1)\mathcal{O}((2M + 1)\log_2(2M + 1)) + (2M + 1)(0.5p_r^2 + 0.5p_r + 3) + 4$. The number of iterations in the G-O algorithm was found to be smaller than that of the window length selection algorithm G-FL, resulting in computational savings.

IV. CONCLUSIONS

We solved the optimum Savitzky-Golay filter selection problem in non-Gaussian noise using a generalized risk estimation framework. The *bias-variance trade-off* was resolved by optimizing the order and filter length. We proposed an unbiased estimate of MSE called the generalized unbiased estimate of MSE (GUE-MSE) in order to address the optimal order (or filter length) selection problem. Unlike SURE, which is compatible with a Gaussian noise distribution only, the proposed risk estimation framework can accommodate any additive noise distribution whose first- and second-order statistics are available provided that the denoiser is linear in the observations. Validations were reported on *Piece-Regular* signals in Gaussian, Laplacian, and Uniform distributions. The performance improved when a regularizer was incorporated. The results with optimum order selection with regularization exhibited better performance than all other denoising algorithms compared in low-to-medium input SNRs (-5 to 12.5 dB). The ECG denoising problem was chosen as an application of GUE-MSE-based optimum SG filter selection. The work reported herein could be extended in several ways. For instance, one could develop the 2-D counterpart of the methodology proposed herein to perform image denoising in non-Gaussian noise conditions with applications in microscopy, ultrasound imaging, etc. While we considered uniform sampling, one could also explore the nonuniform sampling counterpart and deploy the GUE-MSE-optimized SG filter for performing interpolation. Such an approach is more likely to preserve the local signal properties.

APPENDIX
MEDIAN-BASED NOISE VARIANCE ESTIMATOR

Recall the observation model $x_m = s_m + w_m, \quad 1 \le m \le N$, where $\{w_m\}$ are i.i.d. random variables with variance σ^2 and mean zero. Since s_m are assumed to come from a smooth signal, $|x_m - x_{m-1}| \approx |w_m - w_{m-1}|$. Denote $z = |w_m - w_{m-1}|$. Let the estimate of the median be $\hat{\kappa}$. To obtain an estimator for σ,

we rely on the relationship between the median κ of the p.d.f. of z and σ.

A. Uniform Noise

For $\{w_m\}$ that follow the Uniform distribution

$$f(w_m) = \begin{cases} \dfrac{1}{2a}, & -a \le w_m \le a, \\ 0, & \text{otherwise}, \end{cases}$$

z has the p.d.f.

$$f(z) = \begin{cases} -\dfrac{z}{2a^2} + \dfrac{1}{a}, & 0 \le z \le 2a, \\ 0, & \text{otherwise}. \end{cases} \tag{21}$$

The median of (21) satisfies

$$\int_0^{\kappa} \left(-\frac{z}{2a^2} + \frac{1}{a} \right) \mathrm{d}z = \frac{1}{2} \qquad \Rightarrow \kappa = (2 \pm \sqrt{2})a.$$

However, $\kappa = (2+\sqrt{2})a$ is not a valid solution as it exceeds the support $(2a)$ and contradicts the definition of the median. Hence, $\kappa = (2-\sqrt{2})a$ is the only valid solution. Since $\sigma = \dfrac{a}{\sqrt{3}}$, we obtain the estimator $\hat{\sigma} = \dfrac{\hat{\kappa}}{\sqrt{3}(2-\sqrt{2})} = 0.9853\,\hat{\kappa}$.

B. Gaussian Noise

When $\{w_m\}$ are Gaussian distributed with mean zero and variance σ^2, the distribution of z becomes a half-normal distribution [44, pp. 19]. From the definition of the median (denoted by κ) for a half-normal distribution, we know that

$$\frac{1}{\sqrt{2\pi}\sigma_z} \int_0^{\kappa} 2 \exp\left(-\frac{z^2}{2\sigma_z^2} \right) \mathrm{d}z = \frac{1}{2},$$

where $\sigma_z^2 = 2\sigma^2$. The above equation results in $\kappa = \sigma_z \sqrt{2} \mathrm{erf}^{-1}(\dfrac{1}{2})$, which leads to the estimator $\hat{\sigma} = 1.0485\,\hat{\kappa}$.

C. Laplacian Noise

In this case, $f(w_m; b) = \dfrac{1}{2b} \exp(-\dfrac{|w_m|}{b})$, where $\sigma^2 = 2b^2$. From the distribution of the corresponding z [45, pp. 43–49], [46] we obtain

$$\int_0^{\kappa} \left[\frac{z}{2b^2} + \frac{1}{2b} \right] \exp\left(-\frac{z}{b} \right) \mathrm{d}z = \frac{1}{2} \Rightarrow \frac{\kappa}{b} \approx 1.1461,$$

which results in the estimator $\hat{\sigma} = \frac{\hat{\kappa}\sqrt{2}}{1.1461} = 1.2337\,\hat{\kappa}$.

REFERENCES

[1] H. Ziegler, "Properties of digital smoothing polynomial (DISPO) filters," *Appl. Spectrosc.*, vol. 35, pp. 3588–3592, 1981.

[2] A. Ben Khaled-El Feki, L. Duval, C. Faure, D. Simon, and M. Ben Gaid, "CHOPtrey: Contextual online polynomial extrapolation for enhanced multi-core co-simulation of complex systems," *Simul. Trans. Soc. Model. Simul.*, vol. 93, no. 3, pp 185–200, Mar. 2017.

[3] M. Avery, "Literature review for local polynomial regression," Accessed: Aug. 13, 2021. [Online]. Available: https://casual-inference.com/post/local-polynomial-smoothing.pdf

[4] J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications*, 1st ed. London, U.K: Chapman & Hall, 1996.

[5] W. Hardle, *Applied Nonparametric Regression*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[6] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964.

[7] T. A. Brubaker, F. N. Cornett, and C. L. Pomernacki, "Linear digital filtering for laboratory automation," *Proc. IEEE*, vol. 63, no. 10, pp 1475–1486, Oct. 1975.

[8] M. R. Smith, "Interpolation, differentiation, data smoothing, and least squares fit to data with decreased computational overhead," *IEEE Trans. Ind. Electron.*, vol. IE-32, no. 2, pp 135–141, May 1985.

[9] P. Steffen, "On digital smoothing filters: A brief review of closed form solutions and two new filter approaches," *Circuits Syst. Signal Process.*, vol. 5, no. 2, pp 187–210, 1986.

[10] S. J. Orfanidis, *Introduction to Signal Processing*. Englewood Cliffs, NJ, USA: Prentice Hall, 2010.

[11] R. W. Schafer, "What is a Savitzky-Golay filter? [Lecture Notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111–117, Jul. 2011.

[12] R. W. Schafer, "On the frequency-domain properties of Savitzky-Golay filters," in *Proc. Digit. Signal Process. Signal Process. Educ. Meeting Workshop*, 2011, pp. 54–59.

[13] S. Kay and Y. C. Eldar, "Rethinking biased estimation," *IEEE Signal Process. Mag.*, vol. 25, no. 3, pp. 133–136, May 2008.

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, 1st ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[15] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Statist.*, vol. 9, no. 6, pp. 1135–1151, Nov. 1981.

[16] Y. C. Eldar, "Generalized SURE for exponential families: Applications to regularization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 471–481, Feb. 2009.

[17] T. Qiu, A. Wang, N. Yu, and A. Song, "LLSURE: Local linear SURE-based edge-preserving image filtering," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 80–90, Jan. 2013.

[18] T. Blu and F. Luisier, "The SURE-LET approach to image denoising," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2778–2786, Nov. 2007.

[19] F. Luisier, T. Blu, and M. Unser, "A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 593–606, Mar. 2007.

[20] A. M. Atto, D. Pastor, and G. Mercier, "Smooth adaptation by Sigmoid Shrinkage," *EURASIP J. Image Video Process.*, Oct. 2009, Art. no. 532312.

[21] H. Kishan and C. S. Seelamantula, "SURE-fast bilateral filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2012, pp. 1129–1132.

[22] C. Candan and H. Inan, "A unified framework for derivation and implementation of Savitzky-Golay filters," *Signal Process.*, vol. 104, pp. 203–211, Nov. 2014.

[23] H. J. Wayt and T. R. Khan, "Integrated Savitzky-Golay filter from inverse taylor series approach," in *Proc. 15th Int. Conf. Digit. Signal Process.*, 2007, pp. 375–378.

[24] Y. S. Shmaliy and L. J. Morales-Mendoza, "FIR smoothing of discrete-time polynomial signals in state space," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2544–2555, May 2010.

[25] C. Gan, R. Todd, and J. M. Apsley, "Drive system dynamics compensator for a mechanical system emulator," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 70–78, Jan. 2015.

[26] S. Hargittai, "Savitzky-Golay least-squares polynomial filters in ECG signal processing," in *Proc. Comput. Cardiol.*, 2005, pp. 763–766.

[27] B. M. Schettino, C. A. Duque, and P. M. Silveira, "Current-transformer saturation detection using Savitzky-Golay filter," *IEEE Trans. Power Del.*, vol. 31, no. 3, pp. 1400–1401, Jun. 2016.

[28] J. Seo, H. Ma, and T. K. Saha, "On Savitzky-Golay filtering for online condition monitoring of transformer on-load tap changer," *IEEE Trans. Power Del.*, vol. 33, no. 4, pp. 1689–1698, Aug. 2018.

[29] R. L. King, C. Ruffin, F. E. LaMastus, and D. R. Shaw, "The analysis of hyperspectral data using Savitzky-Golay filtering practical issues. 2," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 1999, vol 1, pp. 398–400.

[30] C. Chinrungrueng and P. Toonkum, "Real-time speckle reduction and coherence enhancement of ultrasound images based on mixture of anisotropic Savitzky-Golay filters," in *Proc. IEEE Conf. Nucl. Sci.*, 2004, vol. 7, pp. 4212–4216.

[31] V. Sa-ing, P. Vorasayan, N. C. Suwanwela, S. Auethavekiat, and C. Chinrungrueng, "Multiscale adaptive regularisation Savitzky-Golay method for speckle noise reduction in ultrasound images," *IET Image Process.*, vol. 12, no. 1, pp. 105–112, Jan. 2018.

[32] V. Sa-ing, P. Vorasayan, N. C. Suwanwela, S. Auethavekiat, and C. Chinrungrueng, "Real-time 3D ultrasound denoising based on adaptive regularisation Savitzky-Golay filter," *Electron. Lett.*, vol. 53, no. 15, pp. 1029–1031, Jul. 2017.
[33] A. Jose, S. R. Krishnan, and C. S. Seelamantula, "Ridge detection using Savitzky-Golay filtering and steerable second-order Gaussian derivatives," in *Proc. IEEE Int. Conf. Image Process*., 2013, pp. 3059–3063.
[34] N. R. Koluguri, G. N. Meenakshi, and P. K. Ghosh, "Spectrogram enhancement using multiple window Savitzky-Golay (MWSG) filter for robust bird sound detection," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 6, pp. 1183–1192, Jun. 2017.
[35] F. Ramirez-Echeverria, A. Sarr and Y. S. Shmaliy, "Optimal memory for discrete-time FIR filters in state-space," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 557–561, Feb. 2014
[36] S. R. Krishnan and C. S. Seelamantula, "On the selection of optimum Savitzky-Golay filters," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 380–391, Jan. 2013.
[37] S. V. Menon and C. S. Seelamantula, "SURE-optimal two-dimensional Savitzky-Golay filters for image denoising," in *Proc. IEEE Int. Conf. Image Process*., 2013, pp. 459–463.
[38] S. V. Menon and C. S. Seelamantula, "Robust Savitzky-Golay filters," in *Proc. 19th Int. Conf. Digit. Signal Process*., 2014, pp. 688–693.
[39] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
[40] *PhysioBank database*, Last Accessed on: Aug. 13, 2021. [Online]. Available: http://physionet.org/physiobank/database/aami-ec13/.
[41] "WAVELAB toolbox," Accessed: Aug. 13, 2021. [Online]. Available: http://www-stat.stanford.edu/wavelab/Wavelab_850/index_wavelab850.html
[42] M. H. Stone, "The generalized weierstrass approximation theorem," *Math. Mag.*, vol. 21, no. 4, pp. 167–184, 1948.
[43] J. Fan and I. Gijbels, "Adaptive order polynomial fitting: Bandwidth robustification and bias reduction," *J. Comput. Graphical Statist.*, vol. 4, no. 3, pp. 213–227, 1995.
[44] M. Ahsanullah, B. M. G. Kibria, and M. Shakil, *Normal and Student's-t Distributions and Their Applications*. Boston, MA, USA: Atlantis Press, 2014.
[45] S. Kotz, T. J. Kozubowski, and K. Podgörski, *The Laplace Distribution and Generalizations: A Revisit With New Applications*. Boston, MA, USA: Birkhäuser Boston, 2001.
[46] Q. Zhao, H. W. Li, and Y. T. Shen, "On the sum of generalized Gaussian random signals," in *Proc. 7th Int. Conf. Signal Process*., 2004, vol. 1, pp. 50–53.

Arlene John (Student Member, IEEE) received the Bachelor of Technology degree in electrical and electronics engineering from the National Institute of Technology, Calicut, India, in 2017. She is currently working toward the Ph.D. degree with the School of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland, and a Machine Learning Intern with Qualcomm, Cork, Ireland. In the summer of 2016, she was a Research Intern with the Indian Institute of Science, Bangalore, India. During June 2017–June 2018, she was with Bosch India Ltd. as a Project Manager and in engineering and strategy development for hybrid electric vehicles. From March to June, 2019, she was a Senior Visiting Researcher with the Beijing University of Technology, Beijing, China. Her research interests include multisensor data fusion, biomedical signal processing, and machine learning. Her research focuses on the development of data fusion frameworks for wearable health monitoring devices. In 2020, she was a finalist in the young female STEM pioneer category of the Diversity in Tech Awards. She was a recipient of the prestigious University College Dublin President's Award 2021.

Jishnu Sadasivan received the Bachelor of Technology degree in electronics and communications engineering from the Cochin University of Science and Technology, Kerala, India, in 2007, the Master of Technology degree in signal processing from the College of Engineering Trivandrum, Thiruvananthapuram, India, in 2009, and the Ph.D. degree from the Indian Institute of Science, Bangalore, India, in 2018. He is currently a Research Scientist with Cogknit Semantics Pvt. Ltd., Bangalore, India. His research interests include statistical signal processing, speech processing, and machine learning.

Chandra Sekhar Seelamantula (Senior Member, IEEE) received the Bachelor of Engineering degree in electronics and communication engineering from Osmania University College of Engineering, Hyderabad, India, in 1999 and the Ph.D. degree from the Department of Electrical Communication Engineering, Indian Institute of Science (IISc), Bangalore, India, in 2005. During April 2005–March 2006, he was a Technology Consultant for ESQUBE Communication Solutions Private Limited, Bangalore, India, where he developed proprietary audio coding solutions. During April 2006–July 2009, he was a Postdoctoral Fellow with Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, where he specialized in optical-coherence tomography, holography, splines, sparse signal processing, and sampling theories. In July 2009, he joined the Department of Electrical Engineering, IISc, where he is currently a Professor and directs research at Spectrum Lab. He is also an Associate Faculty with the Centre for Neuroscience, IISc. He was the Vice-Chair (2013–2017) and Student's-t Chair (2017–2021) of IEEE Signal Processing Society Bangalore Chapter, an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS (2013–2017), an Associate Editor for the *SPIE Journal of Electronic Imaging* (2014–2018), the Senior Area Editor of the IEEE SIGNAL PROCESSING LETTERS (2017–2021), and is an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING (2018-), an Area Chair for the IEEE International Conference on Acoustics, Speech, and Signal Processing 2020, and Member of IEEE Technical Committee on Computational Imaging (2020-). He was on the Organizing Committee of Interspeech 2018, IEEE International Symposium on Biomedical Imaging 2020, and served as the General Chair of International Conference on Signal Processing and Communications (SPCOM) 2020. He was the recipient of Prof. Priti Shankar Teaching Award from IISc, Digital Health Prize at National Bio-Entrepreneurship Competition (NBEC) 2018, the Grand Challenges Exploration – India (Round 5) Research Award funded by Bill and Melinda Gates Foundation and Biotechnology Industrial Research Assistance Council in 2020, and Qualcomm Innovation Fellowship (joint recipient with Ph.D. students) in 2019, 2020, and 2021, respectively. He was also the recipient of Gold Medal and the Best Thesis Award for Bachelor of Engineering thesis in 1999.