



HAL
open science

Regularization for pipeline impulse response extraction with least square deconvolution

Xun Wang, Mohamed Ghidaoui, Pedro J Lee

► **To cite this version:**

Xun Wang, Mohamed Ghidaoui, Pedro J Lee. Regularization for pipeline impulse response extraction with least square deconvolution. 13th International Conference on Pressure Surges, Nov 2018, Bordeaux, France. <hal-01891797>

HAL Id: hal-01891797

<https://hal.science/hal-01891797v1>

Submitted on 10 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Regularization for pipeline impulse response extraction with least square deconvolution

Xun Wang¹, Mohamed S. Ghidaoui¹, Pedro J. Lee²

1. Department of Civil and Environmental Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China
2. Department of Civil and Natural Resources Engineering, University of Canterbury, Private Bag 4800, Christchurch, New Zealand

ABSTRACT

Impulse response function (IRF) is frequently used for pipeline defect (e.g., leakage and blockage) detection. IRF can be extracted via a least square (LS) deconvolution method if the input signal is known. However, the LS deconvolution is usually an ill-posed inverse problem, thus the regularization approach is needed to solve this problem. The determination of the regularization parameter is essential for regularization problems and is strongly problem-dependent (no universal method that always produces robust and good results). In the present paper, prevailing regularization methods Generalized Cross Validation (GCV) and L-curve are used and tested to decide the optimal regularization parameter. Pseudo random binary sequence (PRBS) signal, which switches between two constant levels with noise-level amplitude, is used as input signal. Numerical and experimental examples show the importance of the determination of regularization parameter and justify that the proposed methods are efficient for extracting the IRF with PRBS signal.

1 INTRODUCTION

Pipeline defect detection is an important issue in urban water supply systems and have attracted a lot of attention from both theoretical and application aspects [1-6]. The time-domain impulse response function (IRF), which is the output of the system when presented with an impulse input signal, clearly shows the reflections of impulse input signal at defects and is thus frequently used for pipeline defect (e.g., leakage and blockage) detection [7-9]. IRF also depicts the relationship between the input signal and the resultant output signal and illustrates how the system modifies the input signal as it propagates through the system.

IRF can be extracted by imposing a pseudo random binary sequence (PRBS) signal [10-12] that switches between two constant levels. The spectrum of a PRBS signal is approximately flat in a wideband of frequency provided that the sequence length is long enough, meaning that a PRBS signal corresponds to a broadband noise and is thus suitable for low frequency system identification [10]. The amplitude of PRBS can be

noise-level such that the system operation disturbance and structural fatigue by other wave generation methods (e.g., step or impulse signals by valve operation) can be avoided. For this reason, recent pipeline leakage detection methods proposed in [13], which is efficient but requires multiple measurements, become easy to implement.

The relationship between designed input signal and measured output signal can be represented by IRF via a convolution, whose discrete form follows a linear model [7]. As a matter of fact, linear relationship has been found and used widely in pipeline detection problem [14-17], particularly the discrete form of the linear models have been proposed [18,19]. For example, the pressure head variation due to leaks equals to contribution from different leaks and each contribution equals to multiplication of leak size and a function of leak location [14,15,20]. However, in the impulse response extraction problem, the least square (LS) solution of the linear model is ill-posed, thus the regularization problem needs to be considered [7].

The determination of the regularization parameter is critical for regularization problems and is strongly problem-dependent. Up to now, there is no universal method for determining the regularization parameter that always produces robust and good results [21]. In a recent paper [7], the Tikhonov regularization for pipeline impulse response extraction with least square deconvolution has been proposed. However, the determination of the regularization parameter in this problem has not been discussed. In the present paper, prevailing regularization methods Generalized Cross Validation (GCV) [22] and L-curve [23,24] are used and tested to decide the optimal regularization parameter.

The organization of this paper is as follows. In Section 2, the linear model for IRF extraction is introduced. Section 3 briefly introduces GCV and L-curve methods. Numerical and experimental results are presented in Section 4. Finally, conclusions are drawn in Section 5.

2 MODEL

A reservoir-pipe-valve system is considered in this paper. A PRBS input signal, denoted as $x(t)$, is used to excite the system. Let $y(t)$ denote the system output signal, the input and output signals have the relationship:

$$y(t) = \int_{-\infty}^{+\infty} g(\tau)x(t-\tau)d\tau, \quad (1)$$

in which $g(t)$ is the IRF of the system. Let $x[n], y[n], g[n], n=1, \dots, N$, be the discretized $x(t), y(t), g(t)$, a linear model describes the convolution Eq. (1):

$$\mathbf{y} = \mathbf{Xg}, \quad (2)$$

in which

$$\mathbf{y} = (y[1], \dots, y[N])^T, \mathbf{g} = (g[1], \dots, g[N])^T \quad (3)$$

and

$$\mathbf{X} = \begin{pmatrix} x[1] & 0 & 0 & \cdots & 0 \\ x[2] & x[1] & 0 & \cdots & 0 \\ x[3] & x[2] & x[1] & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ x[N] & x[N-1] & x[N-2] & \cdots & x[1] \end{pmatrix}. \quad (4)$$

Here, N is the maximum concerned time-step of IRF. For leak detection problem, in general only the first period of IRF is concerned but both the input signal of PRBS and the measured signal can be much longer.

Given $x[n], y[n]$, the discrete IRF can be obtained by solving the least square (LS) problem with Tikhonov regularization [7]:

$$\hat{\mathbf{g}}_{\lambda} = \arg \min_{\mathbf{g}} (\|\mathbf{y} - \mathbf{X}\mathbf{g}\|_2^2 + \lambda \|\mathbf{g}\|_2^2) = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y} \quad (5)$$

in which λ is the regularization parameter and $\|\cdot\|_2$ stands for the L2-norm. Note that according to the Gauss-Markov theorem, the LS solution without regularization (i.e., $\lambda = 0$) of Eq. (2) is optimal in the sense of minimum variance unbiased estimator if the measurement errors of \mathbf{y} are uncorrelated, zero-mean, and have a constant variance. It can be seen from Eq. (4) that \mathbf{X} is a lower triangular matrix which theoretically has full rank. In practice, however, the matrix \mathbf{X} is usually numerically rank deficient and ill-conditioned (in particular for large size of \mathbf{X}), which can be verified via the singular value decomposition (SVD). As a result, the LS solution is ill-posed. Therefore, regularization (i.e., a non-zero λ in Eq. (5)) has to be applied; in this case, a new problem is solved that does not inherit the properties of the LS solution.

The determination of the regularization parameter λ is critical for Tikhonov regularization and is strongly problem-dependent that no universal method always produces robust and good results [21]. This issue is illustrated by Figure 1. Here, experimental results with data that have been used in Ref. [12] are shown, where a steel intact pipe with length being 37.5 m is considered and PRBS signal are used to excite the system. The experimental setup, as well as the input PRBS signal and output measurements, can be found in Ref. [12]. Figure 1 shows the IRF estimation obtained from Eq. (5) with various $\lambda \in [10^{-4}, 10^5]$. It is clear from this figure that the IRF extraction is sensitive to λ and its determination is important for obtaining an IRF with least noise. To the best of our knowledge, this issue has not been reported in the literature, which is the main goal of the present paper.

3 GCV AND L-CURVE FOR DETERMINING THE REGULARIZATION PARAMETER

In this section, the GCV [22] and L-curve [23,24] methods are briefly introduced to decide the regularization parameter λ .

GCV originates from the leave-one-out (LOO) cross-validation method. The basic idea is to minimize the average estimation error of a data point $y[n]$ computed from a (N-1)-

dimensional data/model where the n -th line in Eq. (2) left out [25]. GCV is a rotation-invariant version of the LOO cross-validation which decides the regularization parameter λ by minimizing:

$$GCV(\lambda) = \frac{\|(\mathbf{I} - \mathbf{A}(\lambda))\mathbf{y}\|_2^2}{\text{trace}(\mathbf{I} - \mathbf{A}(\lambda))} \quad (6)$$

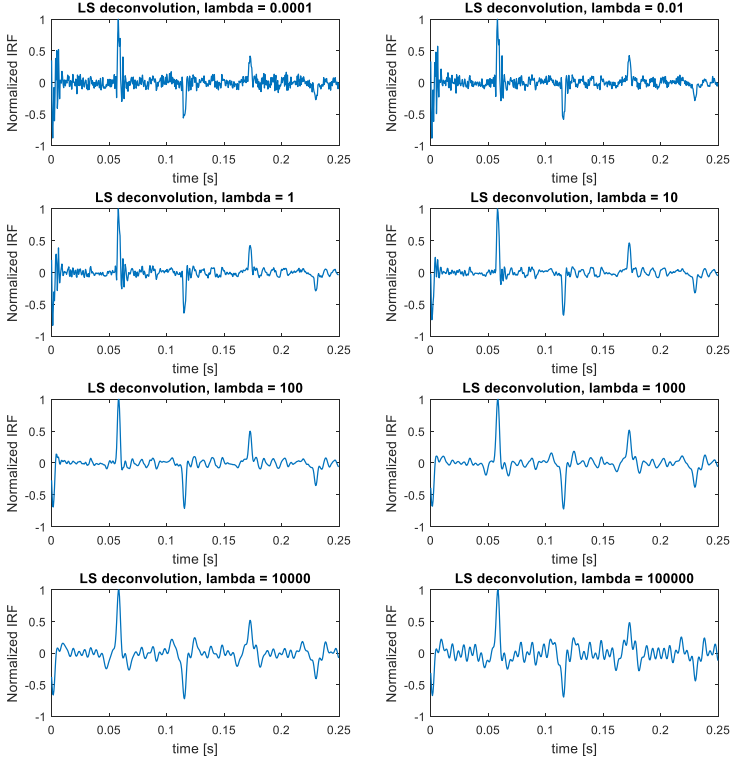


Figure 1. Impulse response function obtained from the Eq. (5) with various Tikhonov regularization parameter λ .

where

$$\mathbf{A}(\lambda) = \mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T. \quad (7)$$

L-curve is a graphical tool with a log-log plot of $\eta(\lambda) = \|\hat{\mathbf{g}}_\lambda\|_2$ versus $\rho(\lambda) = \|\mathbf{y} - \mathbf{X}\hat{\mathbf{g}}_\lambda\|_2$ for all the possible λ . Note that $(\log \rho(\lambda), \log \eta(\lambda))$ usually has a distinct L-shape, and the vertical and horizontal parts of the L-curve correspond to over-regularization and under-regularization, the corner point of the L-curve is seen as a good estimate of the regularization parameter. This corner point is automatically computed by choosing the maximum point of the curvature $\kappa(\lambda)$ of the plot $(\log \rho(\lambda), \log \eta(\lambda))$.

4 NUMERICAL AND EXPERIMENTAL RESULTS

In this section, a numerical example is introduced to show the performance of the proposed methods. The pipe length is $l = 1$ km and the pipe diameter is $D = 0.2$ m. The wave speed is $a = 1000$ m/s. Let the coordinates of the upstream and downstream nodes be $x = 0$ and $x = 1000$ m. A leak is assumed to locate at $x = 300$ m and the leak size is $C^d A^L = 3 \times 10^{-4}$ m².

A transmitter and a pressure sensor are both located at the downstream node $x = 1000$ m. A PRBS signal is emitted by the transmitter, it propagates along the pipe and is measured by the sensor. Both the signal emission and the measurement last $T = 80$ s. The wave propagation is simulated via the Method of Characteristics (MoC) [26], where the grid lengths in space and in time are $\Delta x = 10$ m and $\Delta t = 0.01$ s. Figure 2 (a) shows the input signal of the first two seconds, while Figure 2 (b) displays the output signal for 0-20 s.

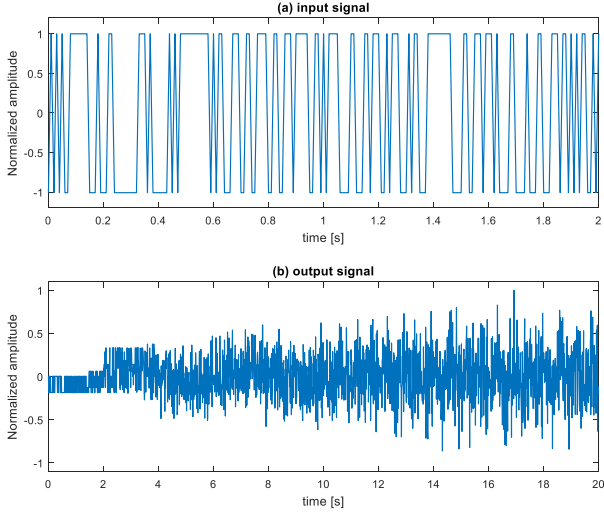


Figure 2. Normalized (a) input signal (PRBS) and (b) output signal at the downstream end of the pipe.

SVD is done for the matrix \mathbf{X} where $N = 8000$; the singular values in descending order normalized by the maximum are shown in Figure 3, which tends to 0 at around 7000-th singular value, meaning that the LS deconvolution is an ill-posed inverse problem. Here, the GCV and L-curve methods are used to regularize this problem. In fact, the IRF can also be obtained from the cross-correlation method [11]:

$$R_{oi}(t) = \frac{1}{N} \sum_{i=1}^{N-t} x[i]y[i+t]. \quad (8)$$

Note that in order for Eq. (8) to be used as a valid estimate of the IRF, the signal is needed to be broadband, i.e., the spectrum of the input function is flat within the considered frequency range. In this section, numerical results of IRF estimation using four methods are compared, i.e., the cross-correlation method (Eq. (8)), LS without

regularization ($\lambda = 0$ in Eq. (5)), GCV, and L-curve. Figure 4 shows the normalized reconstructed IRF. Here, independent and identically distributed Gaussian random noise is added to the measurement and the signal-to-noise ratio (SNR) is 10 dB. Note that the reference signal in the definition of SNR is the average of the output signal, instead of the signal drop due to leak in Refs. [3,14,15]. It can be seen from Figure 4 that the LS without regularization returns a very noisy IRF estimate that the leak cannot be localized. By contrast, from the results of the other three methods, the jump in the IRF due to leak can be clearly observed and GCV returns a less noisy IRF estimate.

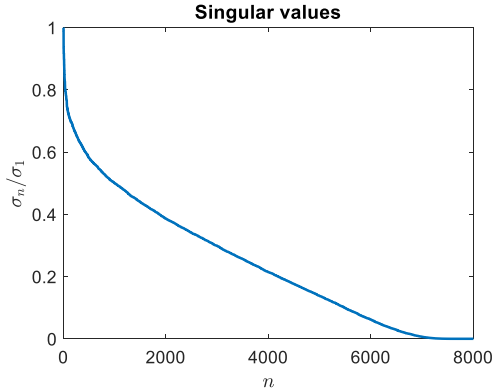


Figure 3. Singular values σ_n in descending order of \mathbf{X} , normalized by their maximum value σ_1 .

In order to detect a leak, the error of IRF estimate must be much lower than the jump in IRF due to wave reflection at the leak. For example, as can be seen in Figure 4 where the leak size is $C^d A^L = 3 \times 10^{-4} \text{ m}^2$, the jump corresponding to the reflection at the leak is 0.2. For a larger leak, the jump in IRF increases, for example if $C^d A^L = 10^{-3} \text{ m}^2$ the height of jump is approximately 0.8 (we numerically test this case but the result is not shown in the present paper), such that the leak detection tolerates a higher IRF estimation error and a higher noise level. However, in order to detect a small leak which is the main concern of the leak detection problem, an IRF extraction method with error as low as possible is desirable.

Finally, let us revisit the results shown in Figure 1 that uses the experimental data in Ref. [12]. Here, the cross-correlation, GCV and L-curve methods are used to extract the IRF; the corresponding results are shown in Figure 5. In this case, the cross-correlation method returns a very noisy result. The Tikhonov regularization parameters λ decided by GCV and L-curve are respectively 20 and 210. The fluctuations in the results of GCV and L-curve are relatively small comparing with the main reflection at the boundary. In particular, L-curve returns almost the optimal results in the sense of least fluctuation in IRF estimation, which can be observed by comparing Figure 5(c) with Figure 1. Note that different from the random noise added in the previous simulation example, model uncertainties appear in real experiments, for example the corner sharpness of the PRBS signal. Tikhonov regularization is able to deal with uncertainties while the cross-correlation method seems to only cancel random noises.

5 CONCLUSION

This paper proposes the generalized cross validation (GCV) and L-curve methods to decide the optimal Tikhonov regularization parameter in the least square (LS) deconvolution method for impulse response function (IRF) extraction of water pipeline. The methods proposed in this paper are also compared with the cross-correlation method. The performance of these methods is evaluated by the estimation of IRF with pseudo random binary sequence (PRBS) input signal. The numerical and experimental results show that the choice of IRF extraction method is important. The availability of the proposed methods is also justified by experimental data.

ACKNOWLEDGEMENT

This work has been supported by research grants from the Research Grant Council of the Hong Kong SAR, China (Project No. T21-602/15R).

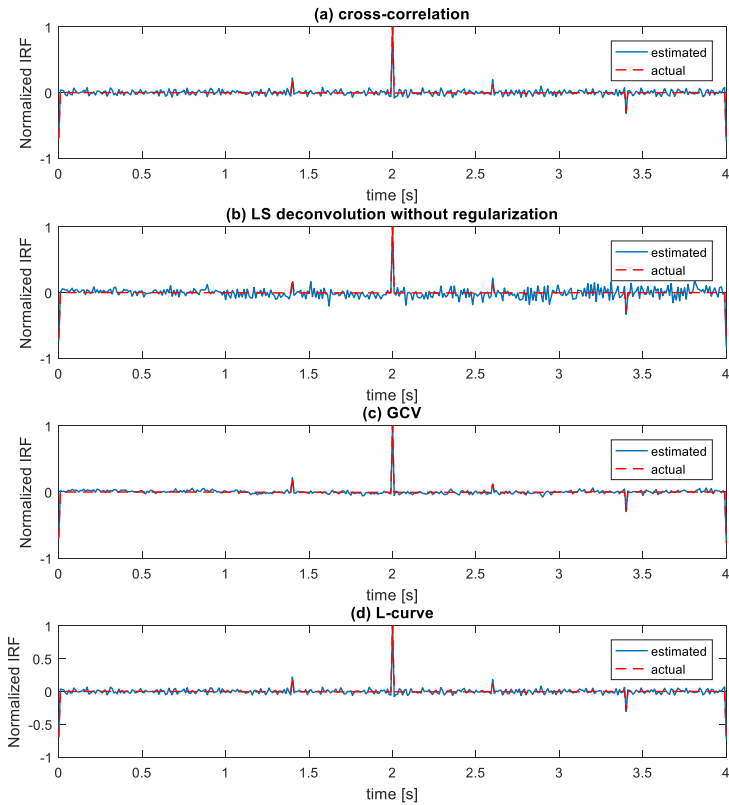


Figure 4. Estimation of IRF normalized by its maximum value. SNR=10 dB.

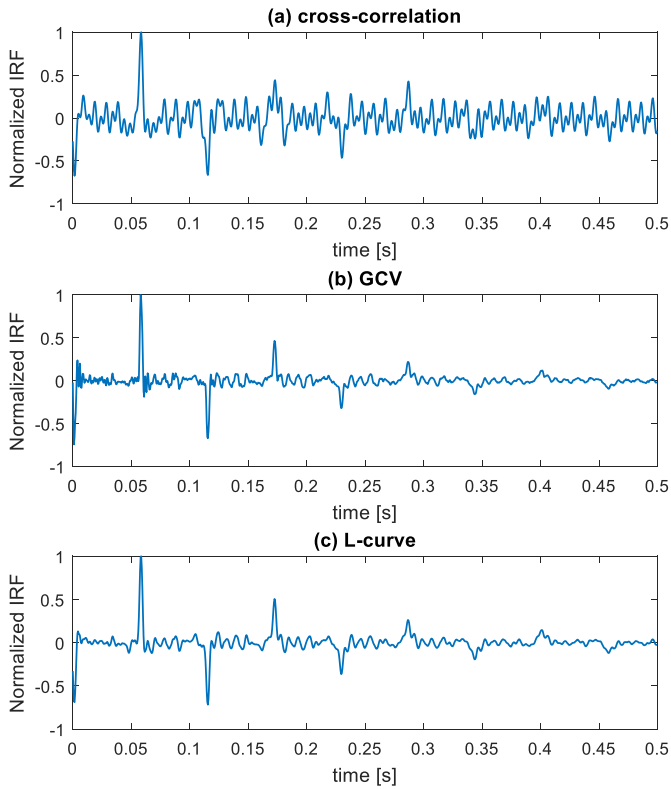


Figure 5. Impulse response function obtained from the Eq. (5) using the cross-correlation, GCV and L-curve methods.

REFERENCE

- [1] Colombo, A. F., Lee, P. J., and Karney, B. W. (2009). A selective literature review of transient-based leak detection methods. *Journal of Hydro-Environment Research*, 2(4), 212-227
- [2] Brunone, B. (1999). Transient test-based technique for leak detection in outfall pipes. *Journal of Water Resources Planning and Management*, 125(5), 302-306
- [3] Ferrante, M., Brunone, B., and Meniconi, S. (2007). Wavelets for the analysis of transient pressure signals for leak detection. *Journal of Hydraulic Engineering*, 133(11), 1274-1282

- [4] Lee, P. J., Vítkovský, J. P., Lambert, M. F., Simpson, A. R., and Liggett, J. A. (2005). Leak location using the pattern of the frequency response diagram in pipelines: a numerical study. *Journal of Sound and Vibration*, 284(3-5), 1051-1073
- [5] Covas, D., Ramos, H., Graham, N., and Maksimovic, C. (2005). Application of hydraulic transients for leak detection in water supply systems. *Water Science and Technology: Water Supply*, 4(5-6), 365-374
- [6] Meniconi, S., Brunone, B., Ferrante, M., Capponi, C., Carrettini, C. A., Chiesa, C., ... and Lanfranchi, E. A. (2015). Anomaly pre-localization in distribution–transmission mains by pump trip: preliminary field tests in the Milan pipe system. *Journal of Hydroinformatics*, 17(3), 377-389
- [7] Nguyen, S. T. N., Gong, J., Lambert, M. F., Zecchin, A. C., and Simpson, A. R. (2018). Least squares deconvolution for leak detection with a pseudo random binary sequence excitation. *Mechanical Systems and Signal Processing*, 99, 846-858
- [8] Lee, P. J., Vítkovský, J. P., Lambert, M. F., Simpson, A. R., and Liggett, J. (2007). Leak location in pipelines using the impulse response function. *Journal of Hydraulic Research*, 45(5), 643-652
- [9] Vitkovsky, J. P., Lee, P. J., Stephens, M. L., Lambert, M. F., Simpson, A. R., and Liggett, J. A. (2003). Leak and blockage detection in pipelines via an impulse response method. *Pumps, electromechanical devices and systems applied to urban water management*, 1, 423-430
- [10] Newland, D. E. (2012). *An introduction to random vibrations, spectral & wavelet analysis*. Courier Corporation
- [11] Liou, C. P. (1998). Pipeline leak detection by impulse response extraction. *Journal of Fluids Engineering*, 120(4), 833-838
- [12] Lee, P. J., Vítkovský, J. P., Lambert, M. F., and Simpson, A. R. (2008). Valve design for extracting response functions from hydraulic systems using pseudorandom binary signals. *Journal of Hydraulic Engineering*, 134(6), 858-864
- [13] Wang, X., Palomar, D. P., Zhao, L., Ghidaoui, M. S., and Murch, R. D. (2019). Spectral-based methods for pipeline leakage detection. *Journal of Hydraulic Engineering*, accepted for publication
DOI: 10.1061/(ASCE)HY.1943-7900.0001572
- [14] Wang, X. and Ghidaoui, M. S. (2018). Pipeline leak detection using the matched-field processing method. *Journal of Hydraulic Engineering*, 144(6), 04018030
- [15] Wang, X. and Ghidaoui, M. S. (2018). Identification of multiple leaks in pipeline: Linearized model, maximum likelihood, and super-resolution localization. *Mechanical Systems and Signal Processing*, 107, 529-548
- [16] Wang, X. and Ghidaoui, M. S. (2019). Identification of multiple leaks in pipeline II: Iterative beamforming and leak number estimation. *Mechanical Systems and Signal Processing*, 119, 346-362
- [17] Zouari, F., Wang, X., Louati, M., and Ghidaoui, M. (2017). Single extended blockage identification using a model-based matched-field processing approach. In 37th IAHR World Congress
- [18] Keramat, A., Ghidaoui, M., and Wang, X. (2017). Inverse transient analysis for pipeline leak detection in a noisy environment. In 37th IAHR World Congress
- [19] Zhou, B., Liu, A., Wang, X., She, Y., and Lau, V. (2018). Compressive Sensing-Based Multiple-Leak Identification for Smart Water Supply Systems. *IEEE Internet of Things Journal*, 5(2), 1228-1241
- [20] Lee, P. J., Vítkovský, J. P., Lambert, M. F., Simpson, A. R., and Liggett, J. A. (2005). Frequency domain analysis for detecting pipeline leaks. *Journal of Hydraulic Engineering*, 131(7), 596-604

- [21] Bauer, F., and Lukas, M. A. (2011). Comparing parameter choice methods for regularization of ill-posed problems. *Mathematics and Computers in Simulation*, 81(9), 1795-1841
- [22] Golub, G. H., Heath, M., and Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2), 215-223
- [23] Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the L-curve. *SIAM review*, 34(4), 561-580
- [24] Hansen, P. C., and O'Leary, D. P. (1993). The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing*, 14(6), 1487-1503
- [25] Allen, D. M. (1974). The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, 16(1), 125-127
- [26] Chaudhry, M. H. (2014). *Applied Hydraulic Transients*, Third Edition. Springer