Spectral-Based Methods for Pipeline Leakage Localization

Xun Wang[1]; Daniel P. Palomar[2]; Licheng Zhao[3];
Mohamed S. Ghidaoui, M.ASCE[4]; and Ross D. Murch[5]

Abstract: In this paper, the pipeline leak localization problem using transient data is investigated. Signal processing techniques that proved successful in wireless communications and acoustics are adapted and tested for leak identification. More specifically, Bartlett's beamforming (BF) (also known as conventional BF, matched field, or phased array), Capon's BF (also known as the minimum variance distortionless response filter), Lagunas' BF, and multiple signal classification (MUSIC) methods are used. The localization is realized by a one-dimensional search for the leak location along the pipe, where one-dimensional search means that the wave model used includes one leak only. The one-dimensional search is advantageous in that it involves low computational cost. The performance of the different techniques in the cases of a single leak and multiple leaks is discussed. In the single-leak case, the proposed spectral methods accurately localize the leak even for a high level of noise. For the multiple-leak case, the proposed spectral methods are able to localize all leaks provided that the leak-to-leak distance is of the same order or larger than half the shortest probing wavelength. However, the localization deteriorates when the leaks are too close together because a model with a single leak is being used to identify multiple leaks. Although not accurate, the application of the one-dimensional search to multiple leaks is still valuable because it provides a fast initial estimate of the leak locations, which serves as prior information for more precise but computationally expensive multidimensional search techniques. **DOI: 10.1061/(ASCE)HY.1943-7900.0001572.**

Author keywords: Leak localization; Transient wave; Water hammer; Spectral-based method; Beamforming.

Introduction

The World Bank has estimated the worldwide monetary value of lost water from water supply lines to be about USD 14 billion per year (Kingdom et al. 2006). Leaks also pose health risks because they are potential entry points for contaminants during low-pressure intrusion events (Colombo et al. 2009). Therefore, a fast, accurate, and low-cost leak localization method is desired.

Signal processing in fluid conduits, where a hydraulic transient signal is introduced into the pipe system and the pressure response is measured at specified locations, is a promising general approach for defect detection and pipe-wall condition assessment. The basic premise for the transient (wave) signal sampling and processing approach to defect identification and characterization is that a measured wave signal in the fluid in the conduit is modified by its interaction with the physical system as it propagates and reflects throughout the system. Accordingly, it contains information, a sort of imprint, of the conduit's properties and state. This principle forms the basis of a range of transient-based defect detection methods (TBDDM) (e.g., Liggett and Chen 1994; Nash and Karney 1999; Vtkovskỳ et al. 2000; Stephens 2008; Covas and Ramos 2010; Wang et al. 2002; Nixon et al. 2006; Brunone 1999; Brunone and Ferrante 2010; Covas et al. 2005b; Ferrante et al. 2007; Liou 1998; Beck et al. 2005; Mpesha et al. 2001; Ferrante and Brunone 2003; Covas et al. 2005a; Lee et al. 2005a, b, 2010; Meniconi et al. 2011, 2015; Sattar and Chaudhry 2010; Taghvaei et al. 2010; Rubio Scola et al. 2017).

The transfer matrix method (Chaudhry 2014; Wylie and Streeter 1978) provides a nonlinear relation between measured pressure head and unknown parameters of the leaks. The nonlinearity is due to the multiple scattering of the wave by the leaks, which results in modal interdependence. An analytical investigation of the modal interdependence in the presence of a single leak using both one-dimensional (1D) and two-dimensional (2D) water-hammer waves was performed by Nixon et al. (2006). They concluded that modal interdependence is negligible for realistic leak sizes. Wang and Ghidaoui (2018a) showed that the modal-interdependence terms for multiple leaks involve products of the usually small leak sizes and used this fact to arrive at a linear relation between measured pressure head and unknown parameters of the leaks. Based on this expression, Wang and Ghidaoui (2018b) proposed a matched-field processing (MFP) approach able to identify a leak based on a 1D search of leak location along the pipe independent of the leak size, which was proven to be robust with respect to a high level of noise. However, it was found that the presence of local maxima in the objective function of MFP may interfere with the judgment of leak location. Therefore, methods capable of eliminating these interferences are desirable.

Spectral-based methodology is a category of signal processing techniques that are powerful tools for source localization in wireless communications and acoustics (Krim and Viberg 1996).

[1]Research Associate, Dept. of Civil and Environmental Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Hong Kong, China (corresponding author). ORCID: https://orcid.org/0000-0002-1156-3840. Email: xunwang00@gmail.com

[2]Professor, Dept. of Electronic and Computer Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Hong Kong, China.

[3]Ph.D. Student, Dept. of Electronic and Computer Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Hong Kong, China.

[4]Chinese Estates Professor of Engineering and Chair Professor, Dept. of Electronic and Computer Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Hong Kong, China.

[5]Chair Professor, Dept. of Electronic and Computer Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Hong Kong, China.

Note. This manuscript was submitted on January 18, 2018; approved on September 6, 2018; published online on December 31, 2018. Discussion period open until May 31, 2019; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydraulic Engineering***, © ASCE, ISSN 0733-9429.

These methods essentially construct a spectrum-like objective function versus the concerned parameter where one can expect that this objective function reaches its maximum at the actual values of the parameters being identified. Spectral-based methods can be classified into beamforming methods (Krim and Viberg 1996; Capon 1969; Lagunas et al. 1986) and subspace-based methods (Paulraj et al. 1986; Pisarenko 1973; Bienvenu and Kopp 1980). The former corresponds to different filters as a function of the candidate parameter. The latter utilizes the eigenstructure of the correlation matrix of measurements, which can be decomposed into the subspaces corresponding to the concerned signal and measurement noise.

In this paper, three beamforming techniques (Bartlett's, Capon's, and Lagunas' BFs) and a subspace-based method [multiple signal classification (MUSIC)] are used for pipeline leak localization. Bartlett's BF is equivalent to the MFP method of Wang and Ghidaoui (2018b), which has the drawback that its objective function has high local maxima. The present paper generalizes that method to derive other versions of spectral-based methods, which are able to obtain a unique leak localization.

This paper begins with a model description, followed by the introduction of the proposed four spectral-based methods for leak localization. Numerical simulations are then presented, where cases of a single leak and multiple leaks are both considered, and the proposed methods are evaluated by leak localization error and ability of disturbance suppression. Finally, conclusions are drawn.

Model for Leak Localization in Pipelines

In this section, the model of transient wave propagation in a pipe is introduced. Then, the characteristics of the model are discussed, which is helpful for algorithm design of leak localization.

Model Description

A pipe bounded by an upstream node at $x = x^U = 0$ and a downstream node at $x = x^D$ is considered. Let x^{L*} denote the actual leak location, and x^L is the leak location as a free parameter. The lumped leak size is $s^L = C^d A^L$, where C^d is the discharge coefficient of the leak and A^L is the flow area of the leak opening (orifice). The steady-state discharge of the leak is related to the lumped leak parameter by $Q_0^L = s^L\sqrt{2g(H_0^L - z^L)}$, in which g is gravitational acceleration, z^L denotes the elevation of the pipe at the leak, and Q_0^L and H_0^L are, respectively, the steady-state discharge and head at the leak.

Due to a rapid change in flow setting (e.g., valve operation), the flow discharge and pressure head varying with respect to time and space are denoted by q and h. Given the discharge $q(x^U)$ and head $h(x^U)$ at x^U, the quantities at mth sensor ($m = 1, \ldots, M$) at x_m can be computed via the transfer matrix (after linearization) method (Chaudhry 2014)

$$\begin{pmatrix} q(x_m) \\ h(x_m) \end{pmatrix} = M^{NL}(x_m - x^L)\begin{pmatrix} 1 & -\dfrac{Q_0^L}{2(H_0^L - z^L)} \\ 0 & 1 \end{pmatrix} M^{NL}(x^L) \times \begin{pmatrix} q(x^U) \\ h(x^U) \end{pmatrix} \tag{1}$$

if $x_m > x^L$; and

$$\begin{pmatrix} q(x_m) \\ h(x_m) \end{pmatrix} = M^{NL}(x_m)\begin{pmatrix} q(x^U) \\ h(x^U) \end{pmatrix} \tag{2}$$

if $x_m \le x^L$, in which

$$M^{NL}(x) = \begin{pmatrix} \cosh(\mu x) & -\dfrac{1}{Z}\sinh(\mu x) \\ -Z\sinh(\mu x) & \cosh(\mu x) \end{pmatrix} \tag{3}$$

is the field matrix; $Z = \mu a^2/(\mathrm{i}\omega g A)$ is the characteristic impedance; $\mu = a^{-1}\sqrt{-\omega^2 + \mathrm{i}gA\omega R}$ is the propagation function, where a is wave speed, ω is angular frequency, A is area of pipeline, and R is frictional resistance term. Here, $R = (FQ_0)/(gDA^2)$ for turbulent flows, where F is the Darcy-Weisbach friction factor, Q_0 is the steady-state discharge, and D is the internal pipe diameter. Unsteady-state friction is neglected in the transient model, which is reasonable for the considered low-frequency range (Nixon et al. 2006; Covas et al. 2010c).

By rewriting Eq. (1), the pressure head at x_m for the jth frequency ($j = 1, \ldots, J$) can be obtained (Wang and Ghidaoui 2018b)

$$h(\omega_j, x_m) = h^{NL}(\omega_j, x_m) + s^L G(\omega_j, x^L, x_m) + n_{mj} \tag{4}$$

where

$$h^{NL}(\omega_j, x_m) = -Z\sinh(\mu x_m)q(x^U) + \cosh(\mu x_m)h(x^U) \tag{5}$$

is the theoretical pressure head at $x = x_m$ that does not include the leak terms

$$G(\omega_j, x^L, x_m) = \begin{cases} -\dfrac{\sqrt{g}Z\sinh(\mu(x_m - x^L))}{\sqrt{2(H_0^L - z^L)}}(Z\sinh(\mu x^L)q(x^U) - \cosh(\mu x^L)h(x^U)), & x_m > x^L \\ 0, & x_m \le x^L \end{cases} \tag{6}$$

and n_{mj} = additive independent Gaussian white noise with zero mean and covariance σ^2. For cases where the noise is nonwhite but the noise structure is known, the whitening scheme of Wang and Ghidaoui (2018b) can be used such that the leak localization techniques proposed in the present paper can still be applied.

In Eq. (4), $q(x^U)$ and $h(x^U)$ (boundary conditions at the upstream node) are assumed to be known. If the upstream is connected to a reservoir, $h(x^U)$ can be reasonably assumed to be $h(x^U) = 0$. The discharge $q(x^U)$ can be estimated, denoted as $\hat{q}(x^U)$, if a sensor near the upstream boundary, whose location is denoted by x^{M_0}, is available. Assuming there is no leak between x^U and x^{M_0} and using the pressure-head measurement $h(x^{M_0})$ at x^{M_0}, the transfer matrix for the wave propagation between x^U and x^{M_0} can be written

$$\begin{pmatrix} q(x^U) \\ h(x^U) \end{pmatrix} = M^{NL}(x^U - x^{M_0}) \begin{pmatrix} q(x^{M_0}) \\ h(x^{M_0}) \end{pmatrix}$$
$$= \begin{pmatrix} \cosh(\mu(x^{M_0} - x^U)) & \frac{1}{Z}\sinh(\mu(x^{M_0} - x^U)) \\ Z\sinh(\mu(x^{M_0} - x^U)) & \cosh(\mu(x^{M_0} - x^U)) \end{pmatrix} \times \begin{pmatrix} q(x^{M_0}) \\ h(x^{M_0}) \end{pmatrix} \quad (7)$$

from which the discharge $q(x^U)$ can be estimated (Kashima et al. 2011, 2013) by

$$\hat{q}(x^U) = \frac{\cosh(\mu(x^{M_0} - x^U))h(x^U) - h(x^{M_0})}{Z\sinh(\mu(x^{M_0} - x^U))} = -\frac{h(x^{M_0})}{Z\sinh(\mu(x^{M_0} - x^U))} \quad (8)$$

Letting $\Delta h_{jm} \equiv h(\omega_j, x_m) - h^{NL}(\omega_j, x_m)$ and denoting

$$\Delta\mathbf{h} = (\Delta h_{11}, \ldots, \Delta h_{J1}, \ldots, \Delta h_{1M}, \ldots, \Delta h_{JM})^\top \quad (9)$$

it follows that

$$\Delta\mathbf{h} = s^L\mathbf{G}(x^L) + \mathbf{n} \quad (10)$$

In this equation, $\mathbf{G}(x^L)$ is a JM-dimensional vector

$$\mathbf{G}(x^L) = (G(\omega_1, x^L, x_1), \ldots, G(\omega_J, x^L, x_1), \ldots, G(\omega_1, x^L, x_M), \ldots, G(\omega_J, x^L, x_M))^\top \quad (11)$$

and

$$\mathbf{n} = (n_{11}, \ldots, n_{J1}, \ldots, n_{1M}, \ldots, n_{JM})^\top \quad (12)$$

is the vector of noise. In the next section, $\Delta\mathbf{h}$ is used to localize the leak.

Both the noise and model terms in wireless communications and acoustics for signal source localization, where the spectral-based methods have been applied, are random. This property of wireless communications and acoustics models makes it possible to use a single-source model to localize multiple sources. However, this is not the case for the transient model under investigation. In particular, in the current paper, $\mathbf{n}$ is random, but $s^L\mathbf{G}(x^L)$ is not. Of course the model term $s^L\mathbf{G}(x^L)$ contains modeling errors, but such errors are not random.

Characteristics of the Model of Wave Propagation in Pipes

In this section, the features of the model in Eq. (10) are investigated. First, in the inverse problem, $\mathbf{G}$ has double identities: (1) it is a function of the free parameter x^L of leak location (to be tuned to estimate the actual leak location); and (2) the actual leak location implicitly appears in $\mathbf{G}$ via the estimated boundary condition $\hat{q}(x^U)$ obtained from the head measurements $h(x^{M_0})$ at x^{M_0}. This feature can be clearly seen from Eq. (6) and by distinguishing the actual leak location x^{L*} and the leak location as a free parameter x^L in $\mathbf{G}$, as $\mathbf{G}(x^L; \hat{q}(x^U; x^{L*}))$. In the forward problem where the leak location is known (simulation of wave generation and measurement), a measurement $\Delta\mathbf{h}$ in Eq. (10) with a leak at x^{L*} is obtained when $x^L = x^{L*}$ in $\mathbf{G}$, i.e., as follows:

$$\Delta\mathbf{h}(x^{L*}) = s^L\mathbf{G}(x^{L*}; \hat{q}(x^U; x^{L*})) + \mathbf{n} \quad (13)$$

In the following subsections, the variation of $\Delta\mathbf{h}(x^{L*})$ with x^{L*} and the variation of $\mathbf{G}(x^L; \hat{q}(x^U; x^{L*}))$ with x^L are both investigated. Their coupling, more exactly, the feature of $\mathbf{G}^H\Delta\mathbf{h}$, is then studied. Here, the superscript H denotes the operation of conjugate transpose.

Shape of Head Difference Δh with Actual Leak Location x^{L*}

Here, the average head difference over frequencies, denoted as $\overline{|\Delta h|}$, with respect to varied leak location x^{L*} is plotted in Fig. 1. Here, no noise is added and $\overline{|\Delta h|}$ is actually computed from 3,001 frequencies $\{\omega{:}\alpha\omega_{th}, \alpha = 1, 1.01, 1.02, \ldots, 31\}$. The pipe length is 2,000 m and the pipe diameter is 0.5 m. The measurement location is at $x^M = 2{,}000$ m and $x^M = 1{,}800$ m, and the leak size is $s^L = 10^{-4}$ m^2. Fig. 1 shows that the average spectrum does not significantly vary with respect to different leak locations except near the upstream node and near the measurement point. This virtue is essential for defining a uniform signal-to-noise ratio (SNR) in different cases of leak location, i.e., a given SNR implies the same noise level no matter where the leak is located. This issue is more specifically introduced in a subsequent section.

Fig. 1 also illustrates that if a leak is very close to the left end of the pipe (the distance between the leak and the upstream reservoir is shorter than approximately 50 m), it is difficult to detect because the signal $\Delta\mathbf{h}$ is weak in this case. Furthermore, this figure, as well as Eq. (6), shows that a leak located at the downstream side of a sensor ($x^L > x_m$) cannot be identified by the model. Therefore, at least one sensor at the downstream end is needed.

Shape of G with Free Parameter x^L

The shape of $\mathbf{G}$ with respect to free parameter x^L of leak location is studied, which is essential for estimating the leak location in the inverse problem. Here, a numerical example is shown in Fig. 2(a) that plots $|G(x^L; \hat{q}(x^U; x^{L*}))|$ as a function of x^L along the pipe at the measurement location $x^M = 2{,}000$ m for the angular frequencies ω_{th}, $3\omega_{th}$, $5\omega_{th}$, and $7\omega_{th}$ [$\omega_{th} = a\pi/(2l)$ denotes the

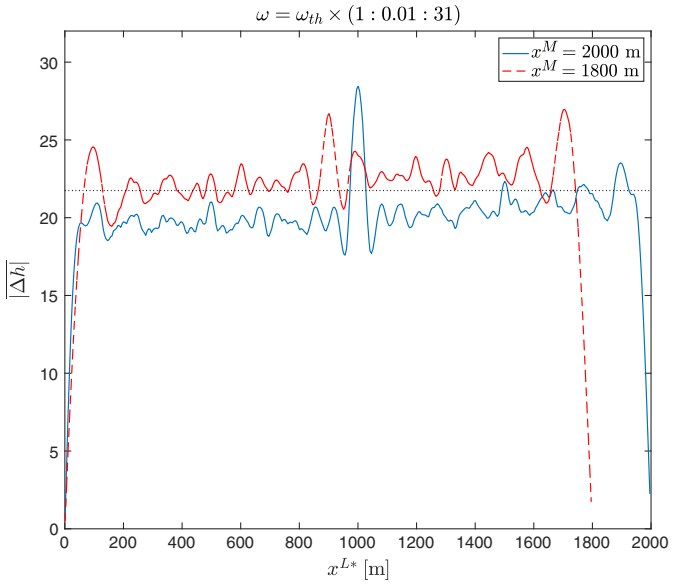

Fig. 1. Plot of $\overline{|\Delta h|}$ (average of $|\Delta h|$ over frequencies) with respect to various leak locations, where $x^M = 2{,}000$ m (solid line) and $x^M = 1{,}800$ m (dashed line). The pipe length is $l = 2{,}000$ m.

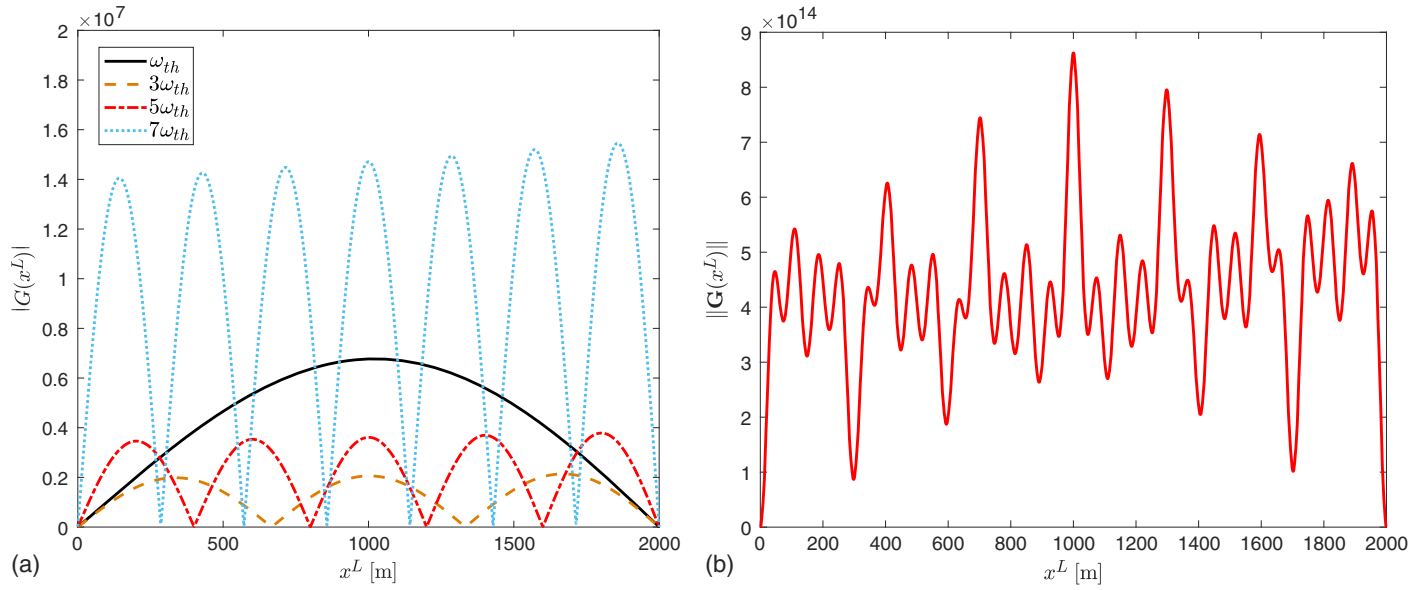

Fig. 2. (a) Plot of $|G(x^L)|$ for resonant frequencies $\omega = \omega_{th}$, $\omega = 3\omega_{th}$, $\omega = 5\omega_{th}$, and $\omega = 7\omega_{th}$, respectively; and (b) plot of $\|\mathbf{G}(x^L)\|$ in which the frequencies are $\{\alpha\omega_{th}:\alpha = 1, 2, \ldots, 31\}$. The measurement location is $x^M = 2{,}000$ m, the actual leak location is $x^{L*} = 600$ m, and the pipe length is $l = 2{,}000$ m.

fundamental angular frequency]. Here, the actual leak location is $x^{L*} = 600$ m and the pipe length is $l = 2{,}000$ m. It is clear that these functions have different zeros but all are equal to zero at $x^L = 0$ and x^M. Therefore, the vector $\mathbf{G}(x^L)$ becomes the zero vector only when $x^L = 0$ and x^M, as shown in Fig. 2(b). In this figure, $\|\mathbf{G}(x^L)\|$ is plotted where the used frequencies are $\{\alpha\omega_{th}: \alpha = 1, 2, \ldots, 31\}$.

When the boundary condition $h(x^U) = 0$ is enforced (i.e., the upstream is a reservoir) and $q(x^U)$ is estimated, each element of $\mathbf{G}(x^L)$ in Eq. (6) becomes

$$G(\omega, x^L, x_m) = -\frac{\sqrt{g}Z}{\sqrt{2(H_0^L - z^L)}} Z \sinh(\mu(x_m - x^L)) \sinh(\mu x^L) \hat{q}(x^U), \quad x^L \in [0, x_m] \tag{14}$$

This equation implies that $\mathbf{G}(x^L)$ is symmetric with respect to $x^M/2$ if the pipe is in a horizontal plane ($z^L = 0$) and the variation of H_0^L along the pipe is neglected. Here, a normalized ambiguity function (following the terminology in radar systems) f_A is plotted to show the similarity of $\mathbf{G}$ between two different locations

$$f_A(x_1, x_2) = \frac{|\mathbf{G}^H(x_1)\mathbf{G}(x_2)|}{\|\mathbf{G}(x_1)\|\|\mathbf{G}(x_2)\|} = \frac{|\mathbf{G}^H(x_1; \hat{q}(x^U; x^{L*}))\mathbf{G}(x_2; \hat{q}(x^U; x^{L*}))|}{\|\mathbf{G}(x_1; \hat{q}(x^U; x^{L*}))\|\|\mathbf{G}(x_2; \hat{q}(x^U; x^{L*}))\|} \tag{15}$$

The corresponding results are shown in Fig. 3, and the chosen frequencies are $\{\omega:\alpha\omega_{th}, \alpha = 1, 2 \ldots, 31\}$. Fig. 3(a) shows the result wherein the actual leak is located at $x^{L*} = 600$ m [implicitly appears in $\mathbf{G}$ via $\hat{q}(x^U; x^{L*})$] and $x^M = 2{,}000$ m; the bottom-left to top-right strong diagonal is the expected property of similarity at the same location, whereas the top-left to bottom-right strong diagonal is the undesired symmetric property (which leads to confusion when it comes to identifying the location of a leak). Similar results are displayed in Fig. 3(b), where the measurement location is $x^M = 1{,}800$ m; here, in the region $\{(x_1, x_2): x_1 > x^M \text{ or } x_2 > x^M\}$ (upper and right margins) the function f_A is not defined. Apart from the two strong diagonals, this case also contains many undesired structures, which might be translated into potential confusion in terms of leak localization.

The undesired strong diagonal line in Figs. 3(a and b) can be largely weakened by using more than one sensor because the symmetric point $x^M/2$ is different for different measurement locations. Fig. 3(c) shows the normalized ambiguity function with two sensors at 1,800 and 2,000 m, from which indeed the undesired top-left to bottom-right diagonal has significantly vanished. Again, in the upper and right margins, the values of f_A are lower than other regions because only one sensor contributes to f_A in that region.

In Figs. 3(a–c), a feature with a characteristic length 600 m can be observed. Actually, this length equals to the leak location x^L where reflections of transient waves happen. This conclusion is confirmed via Fig. 3(d), where $x^{L*} = 300$ m [again, it implicitly appears in $\mathbf{G}$ by affecting $\hat{q}(x^U; x^{L*})$]; in this case, the characteristic length becomes 300 m.

Coupling of Δh and G: Spatial Aliasing Effect

The product of data and the corresponding model appears in many parameter estimation methods. Therefore, the feature of $\Delta\mathbf{h}^H\mathbf{G}$ is studied in the case of no noise. Here, a two-dimensional function of actual leak location x^{L*} and varying free parameter x^L is defined, which stands for the normalized cross-correlation function between the head difference $\Delta\mathbf{h}(x^{L*})$ and the function $\mathbf{G}(x^L; \hat{q}(x^U; x^{L*}))$

$$f_C(x^{L*}, x^L) = \frac{|\mathbf{G}^H(x^L; \hat{q}(x^U; x^{L*}))\Delta\mathbf{h}(x^{L*})|}{\|\mathbf{G}(x^L; \hat{q}(x^U; x^{L*}))\|\|\Delta\mathbf{h}(x^{L*})\|} = \frac{|\mathbf{G}^H(x^L; \hat{q}(x^U; x^{L*}))\mathbf{G}(x^{L*}; \hat{q}(x^U; x^{L*}))|}{\|\mathbf{G}(x^L; \hat{q}(x^U; x^{L*}))\|\|\mathbf{G}^H(x^{L*}; \hat{q}(x^U; x^{L*}))\|} \tag{16}$$

A similar two-dimensional function of x^{L*} and x^L is defined

$$f_B(x^{L*}, x^L) = \frac{|\mathbf{G}^H(x^L; \hat{q}(x^U; x^{L*}))\Delta\mathbf{h}(x^{L*})|}{\|\mathbf{G}(x^L; \hat{q}(x^U; x^{L*}))\|^2} = \frac{s^L|\mathbf{G}^H(x^L; \hat{q}(x^U; x^{L*}))\mathbf{G}(x^{L*}; \hat{q}(x^U; x^{L*}))|}{\|\mathbf{G}(x^L; \hat{q}(x^U; x^{L*}))\|^2} \tag{17}$$

J. Hydraul. Eng., 2019, 145(3): 04018089

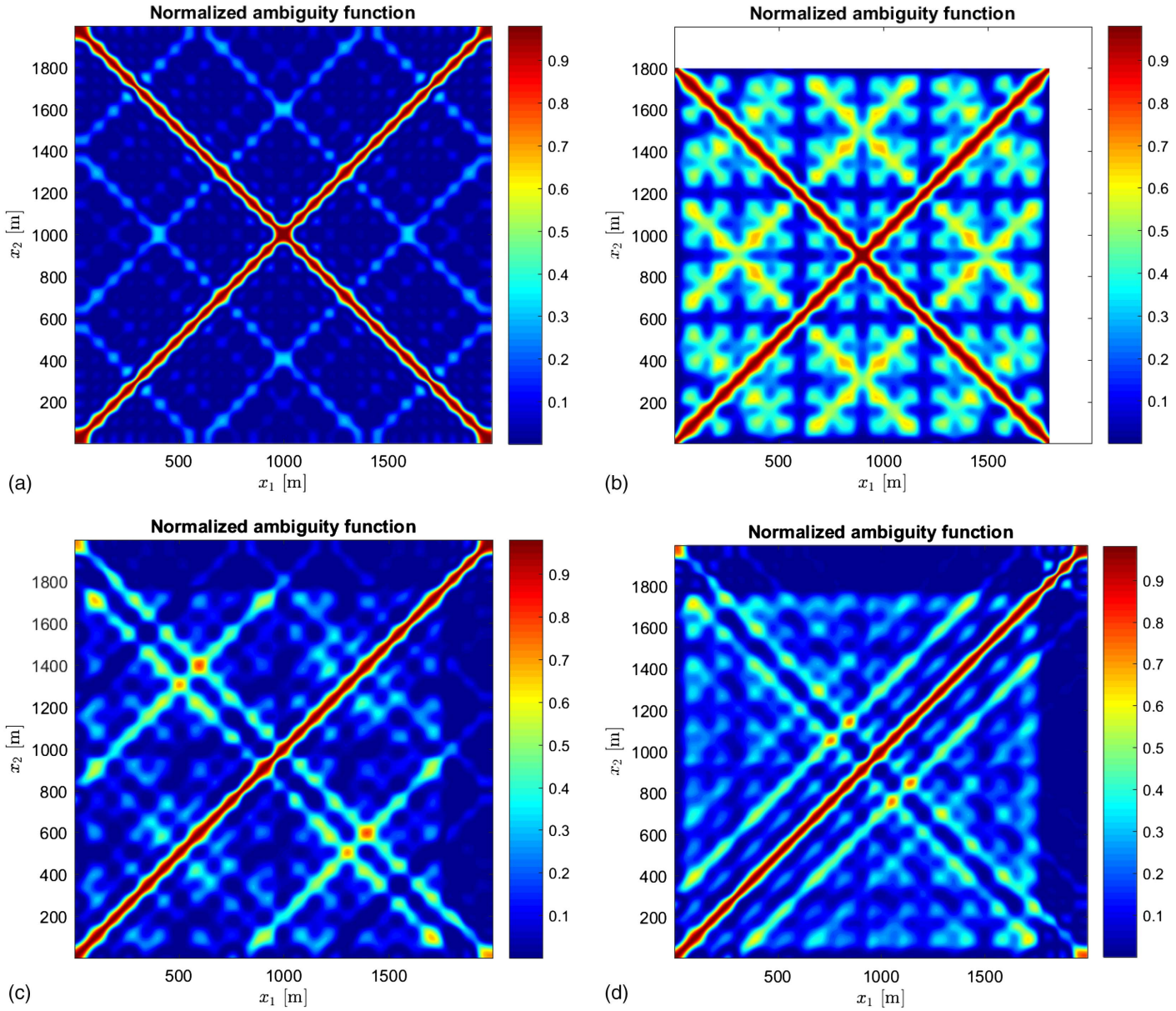

Fig. 3. Normalized ambiguity function of $\mathbf{G}$. Sensor locations are (a) $x^M = 2{,}000$ m; (b) $x^M = 1{,}800$ m; and (c and d) $x^{M_1} = 2{,}000$ m and $x^{M_2} = 1{,}800$ m. The actual leak is (a–c) $x^{L*} = 600$ m; and (d) $x^{L*} = 300$ m).

This is actually the objective function of MFP of Wang and Ghidaoui (2018b) (or, equivalently, Bartlett's BF, which is introduced in the next section). Eqs. (16) and (17) are plotted in Fig. 4 for the three different cases of measurement locations as previously. Similar to the results in Figs. 3(a and b), Figs. 4(a–d) show that a single measurement station results in an undesired strong diagonal line due to the symmetric feature of $\mathbf{G}$. Moreover, a substrong line from (0, 0) to $(L/2, L)$, as well as gradually weaker lines from (0, 0) to $(L/\beta, L)$ $(\beta = 3, 4, \ldots)$, can be found, which implies that for each actual leak x^{L*} the functions f_C and f_B have high local maxima at βx^{L*} $(\beta = 2, 3, \ldots)$ if $\beta x^{L*} < L$. This phenomenon is known as the spatial aliasing effect, which is undesired for obtaining a clear leak localization. Figs. 4(d–f) signify that the MFP method proposed by Wang and Ghidaoui (2018b) (equivalent to the Bartlett's BF, which is detailed in the next section) cannot cope with this problem, and thus new techniques are desired, which is a main motivation of the present paper.

Spectral-Based Methods for Leak Localization

In this section, spectral-based methods are used to solve the leak localization problem. These methods are based on measuring the spatial power spectrum corresponding to a potential leak location and then sweeping this spectrum over all possible locations. Hence, a 1D search is done that essentially plots the power versus leak location where one can identify peaks with actual leaks. A review of these methods for source localization problems has been given by Krim and Viberg (1996).

Spectral-based methods can be classified into beamforming methods and subspace-based methods. Different beamforming methods correspond to different designs of the weighting vector (filter in both spatial coordinate and frequency) as a function of the candidate leak location x^L, denoted as $\mathbf{w}(x^L)$, then compute the power

$$P(x^L) = P(\mathbf{w}(x^L)) = \mathbf{w}^H(x^L)\hat{\mathbf{R}}\mathbf{w}(x^L) \tag{18}$$

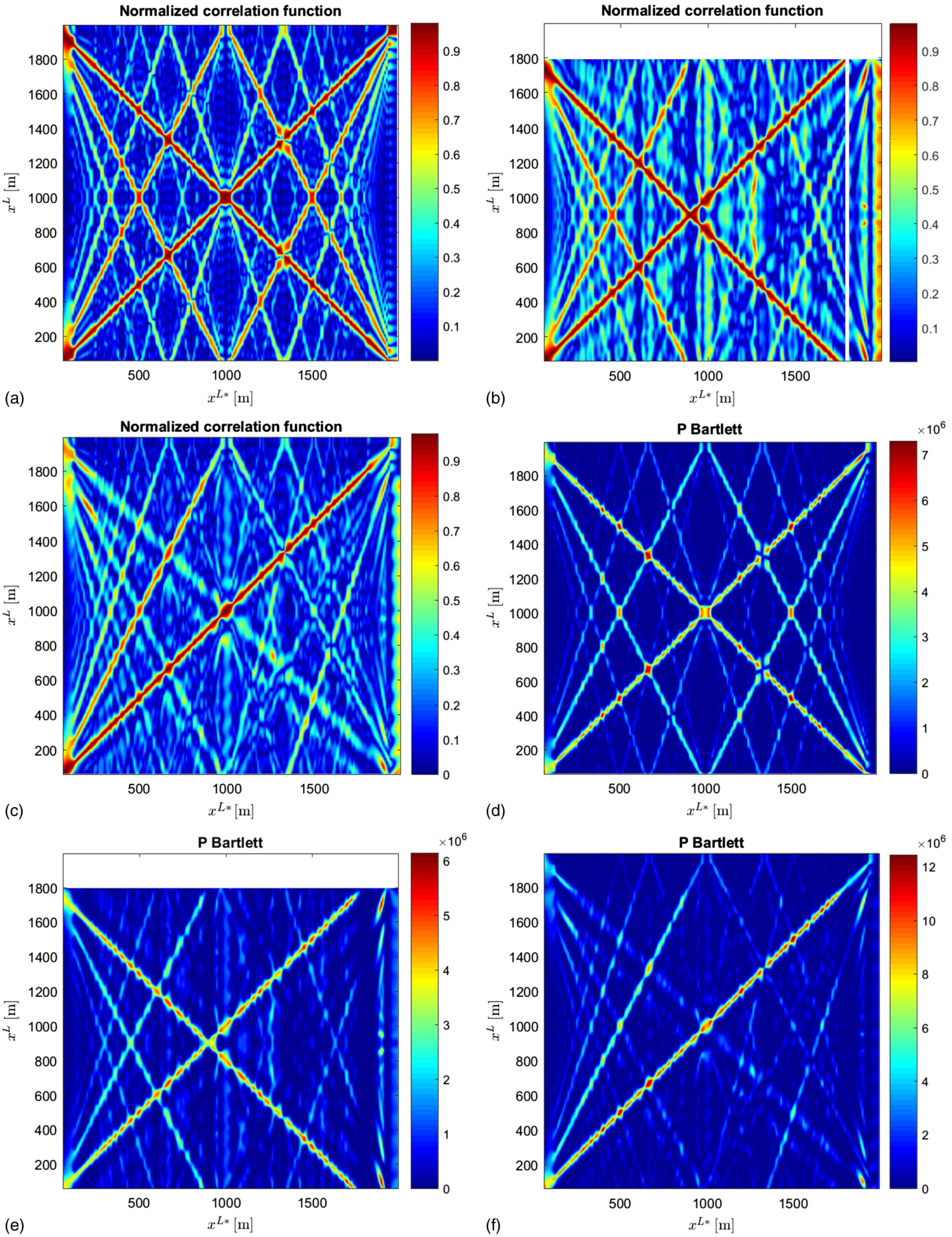

Fig. 4. (a–c) Normalized correlation function between $\mathbf{G}$ and $\Delta\mathbf{h}$; and (d–f) objective function of MFP. Sensor locations are (a and d) $x^M = 2{,}000$ m; (b and e) $x^M = 1{,}800$ m; and (c and f) $x^{M_1} = 2{,}000$ m and $x^{M_2} = 1{,}800$ m.

where $\hat{\mathbf{R}}$ is estimate of the correlation matrix, whereas the subspace-based method involves the eigendecomposition of $\hat{\mathbf{R}}$.

In the following subsections, estimation methods for the correlation matrix are introduced first. Then, three beamforming techniques (Bartlett's, Capon's, and Lagunas' BFs) and a subspace-based method (MUSIC) are presented.

Estimation of Correlation Matrix

Assume that N measurements of $\Delta\mathbf{h}$ are obtained, denoted by $\Delta\mathbf{h}_n$, $n = 1, \ldots, N$. The sample correlation matrix (SCM) is computed by

$$\hat{\mathbf{R}}_{\text{SCM}} = \frac{1}{N}\sum_{n=1}^{N} \Delta\mathbf{h}_n \Delta\mathbf{h}_n^H \tag{19}$$

However, when the sample size N is lower than the dimension MJ of head measurement, Eq. (19) is not full rank. Alternatively, the sample correlation matrix can be regularized using a diagonal loading (DL) method to guarantee the full rank

$$\hat{\mathbf{R}}_{\text{DL}} = (1-\epsilon)\frac{1}{N}\sum_{n=1}^{N} \Delta\mathbf{h}_n \Delta\mathbf{h}_n^H + \epsilon\mathbf{I} \tag{20}$$

where $\mathbf{I}$ is the identity matrix; and ϵ = regularization parameter, which can be determined by an optimal approach (in the sense of minimum Frobenius error of the correlation matrix) introduced by Ledoit and Wolf (2004).

Another estimate of correlation matrix can be obtained by realizing that the model in Eq. (10) has a strong low-rank component plus noise (commonly termed factor model in machine learning and finance). The estimation of the correlation matrix under such model corresponds to the popular principal component analysis (PCA) (Jolliffe 2002). In particular, the estimation of the correlation matrix for the rank-one plus noise model in Eq. (10) is [compare with Eqs. (54) and (57) of Sun et al. (2016)]

$$\hat{\mathbf{R}}_{\text{PCA}} = (\lambda_1 - \widehat{\sigma^2})\mathbf{U}_1\mathbf{U}_1^H + \widehat{\sigma^2}\mathbf{I} \tag{21}$$

where λ_1 = largest eigenvalue of the sample correlation matrix $\hat{\mathbf{R}}_{\text{SCM}}$; $\mathbf{U}_1$ = eigenvector associated with λ_1; and

$$\widehat{\sigma^2} = \frac{1}{MJ-1}\sum_{i=2}^{MJ} \lambda_i \tag{22}$$

is the average of the noise eigenvalues $\lambda_2, \ldots, \lambda_{MJ}$ (from the second largest eigenvalue to the smallest) of $\hat{\mathbf{R}}_{\text{SCM}}$.

Beamforming Methods

Here, Bartlett's, Capon's, and Lagunas' BF methods are introduced for leak localization.

Bartlett's Beamforming

Bartlett's BF for the leak detection problem was already introduced by Wang and Ghidaoui (2018b). Here, only the main results are recalled. The weighting vector is obtained by maximizing the output power and imposing the constraint of unit length ($\|\mathbf{w}\| = 1$) as follows:

$$\begin{aligned}\mathbf{w}_B &= \arg\max_{\mathbf{w}} \mathbb{E}(\mathbf{w}^H \Delta\mathbf{h}\Delta\mathbf{h}^H\mathbf{w}) \\ &= \arg\max_{\mathbf{w}}((s^L)^2\mathbf{w}^H\mathbf{G}\mathbf{G}^H\mathbf{w} + \sigma^2) = \frac{\mathbf{G}}{\|\mathbf{G}\|}\end{aligned} \tag{23}$$

Inserting Eq. (23) into Eq. (18), the spatial power spectrum is obtained

$$P_B(x^L) = \frac{\mathbf{G}^H(x^L)\hat{\mathbf{R}}\mathbf{G}(x^L)}{\mathbf{G}^H(x^L)\mathbf{G}(x^L)} \tag{24}$$

where $\hat{\mathbf{R}}$ is the estimated correlation matrix obtained from Eqs. (19), (20), or (21). The leak position is estimated by finding the peaks in the 1D plot of Eq. (24). Under the assumption of independent Gaussian noise, the leak estimate using Bartlett's BF is a maximum likelihood estimate (MLE). It is also equivalent to a matched-filter approach, which maximizes the SNR (Wang and Ghidaoui 2018b). It is important to recognize that Eq. (24) corresponds to a slice of Eq. (17) where each x^{L*} is fixed [or a vertical slice of Figs. 3(b, d, and f)]. Therefore, this approach has the disadvantage of the presence of local maxima. In the following subsections, other methods are introduced that try to compensate for the local peaks.

Capon's Beamforming

The Capon's BF (Capon 1969), also known as the minimum variance distortionless response (MVDR) filter, is derived by minimizing the power of interference from noise while maintaining a fixed contribution from the leak at x^L. This is equivalent to

$$\min_{\mathbf{w}} P(\mathbf{w}) = \min_{\mathbf{w}} \mathbf{w}^H\hat{\mathbf{R}}\mathbf{w}, \qquad \text{subject to } \mathbf{w}^H\mathbf{G}(x^L) = 1 \tag{25}$$

By using Lagrange multipliers, the optimal $\mathbf{w}$ can be obtained

$$\mathbf{w}_C = \frac{\hat{\mathbf{R}}^{-1}\mathbf{G}(x^L)}{\mathbf{G}^H(x^L)\hat{\mathbf{R}}^{-1}\mathbf{G}(x^L)} \tag{26}$$

Inserting Eq. (26) into Eq. (18), the spatial power spectrum is obtained as follows:

$$P_C(x^L) = \frac{1}{\mathbf{G}^H(x^L)\hat{\mathbf{R}}^{-1}\mathbf{G}(x^L)} \tag{27}$$

Eq. (25) implies that the Capon's BF concentrates on the contribution from each leak location to the measurements. For this reason, a plot of Eq. (27) shows a narrow peak at the actual leak and it has a strong ability to suppress side lobes.

However, the power spectrum in Eq. (27) has a singularity at $x^L = 0$ due to the fact that the function $\mathbf{G}(x^L)$ is the zero vector when $x^L = 0$. This singularity also occurs at x^M if only one sensor is used, as introduced in the previous section and depicted in Fig. 2. Therefore, the Capon's power spectrum $P_C(x^L) \to \infty$ as $x^L \to 0$ because $\mathbf{w}$ has to become large to satisfy the constraint $\mathbf{w}^H\mathbf{G}(x^L) = 1$. Instead, this paper proposes a normalized constraint $\mathbf{w}^H\mathbf{G}(x^L)/\|\mathbf{G}(x^L)\| = 1$, which leads to a normalized version of Capon's BF whose corresponding $\mathbf{w}$ is

$$\mathbf{w}_{\text{NC}} = \frac{\|\mathbf{G}(x^L)\|\hat{\mathbf{R}}^{-1}\mathbf{G}(x^L)}{\mathbf{G}^H(x^L)\mathbf{R}^{-1}\mathbf{G}(x^L)} \tag{28}$$

and the normalized spatial power spectrum is

$$P_{\text{NC}}(x^L) = \frac{\mathbf{G}^H(x^L)\mathbf{G}(x^L)}{\mathbf{G}^H(x^L)\hat{\mathbf{R}}^{-1}\mathbf{G}(x^L)} \tag{29}$$

Lagunas' Beamforming

Unlike Bartlett's BF, Capon's BF does not maintain a constant norm $\|\mathbf{w}\|$. As a consequence, the power expression $\mathbf{w}^H\hat{\mathbf{R}}\mathbf{w}$ is subject to artificial scaling factors of $\mathbf{w}$. The previous section addressed this problem of normalization by fixing the scaling factor so that the Capon's spatial power spectrum in Eq. (27) becomes Eq. (29). An alternative approach, called Lagunas' BF (or bandwidth-normalized Capon), was proposed by Lagunas et al. (1986). The idea is to make the definition of power spectrum invariant with respect to any scaling factor of $\mathbf{w}$. In particular, instead of using the expression $\mathbf{w}_C^H\hat{\mathbf{R}}\mathbf{w}_C$, one can use the normalized version $\mathbf{w}_C^H\hat{\mathbf{R}}\mathbf{w}_C/(\mathbf{w}_C^H\mathbf{w}_C)$, where $\mathbf{w}_C$ is the Capon's weighting vector in Eq. (26). This leads to the normalized spatial power spectrum

$$P_L(\mathbf{x}^L) = \frac{\mathbf{G}^H(x^L)\hat{\mathbf{R}}^{-1}\mathbf{G}(x^L)}{\mathbf{G}^H(x^L)\hat{\mathbf{R}}^{-2}\mathbf{G}(x^L)} \tag{30}$$

where $\hat{\mathbf{R}}^{-2} = \hat{\mathbf{R}}^{-1}\hat{\mathbf{R}}^{-1}$.

Subspace-Based Leak Localization: MUSIC

The correlation matrix of the head difference $\Delta\mathbf{h}$ can be written

$$\mathbf{R} = \mathbb{E}(\Delta\mathbf{h}^H\Delta\mathbf{h}) = (s^L)^2\mathbf{G}(x^L)\mathbf{G}^H(x^L) + \sigma^2\mathbf{I} \tag{31}$$

and can be eigendecomposed as follows:

$$\mathbf{R} = \mathbf{U}\Lambda\mathbf{U}^H = \lambda_1\mathbf{U}_1\mathbf{U}_1^H + \sigma^2\mathbf{U}_2\mathbf{U}_2^H \tag{32}$$

in which Λ is a diagonal matrix whose diagonal elements are ordered eigenvalues $(\lambda_1, \sigma^2, \ldots, \sigma^2)$; $\mathbf{U}$ is the matrix composed by the corresponding eigenvectors; $\mathbf{U}_1$ and $\mathbf{U}_2$ = first column of $\mathbf{U}$ (because just one leak is assumed whose signature spans one dimension) and second through the last column of $\mathbf{U}$, respectively. In addition, $\mathbf{U}_1 \propto \mathbf{G}(x^L)$ and thus the subspace $\mathbf{U}_2$ is orthogonal to $\mathbf{G}(x^L)$. Therefore, a way to localize the leak is to adjust x^L such that $\mathbf{G}(x^L)$ is orthogonal to $\mathbf{U}_2$, i.e., $\mathbf{G}^H(x^L)\mathbf{U}_2 = \mathbf{0}^\top$.

By eigendecomposing the estimate of correlation matrix $\hat{\mathbf{R}}$ and letting $\hat{\mathbf{U}}_2$ denote the matrix composed by the eigenvectors of $\hat{\mathbf{R}}$ corresponding to the second largest eigenvalue to the smallest eigenvalue, the leak position can be estimated by maximizing

$$P_M(x^L) = \frac{\mathbf{G}^H(x^L)\mathbf{G}(x^L)}{\|\mathbf{G}^H(x^L)\hat{\mathbf{U}}_2\|^2} \tag{33}$$

which is called the MUSIC method (Bienvenu and Kopp 1980; Pisarenko 1973; Krim and Viberg 1996). As previously discussed, the denominator of Eq. (33) is close to zero when x^L is equal to the actual leak location, and the numerator of this equation normalizes the spectra as the normalized Capon's BF.

Simulation Results

In this section, numerical examples are introduced to study and compare the various leak localization methods. The cases of a single leak and multiple leaks are discussed, respectively.

Numerical Setup and Definition of SNR

Transient wave propagation in a single pipeline is simulated; the setup is shown in Fig. 5. The pipe is connected to two reservoirs; the steady-state heads at the upstream and downstream reservoirs are $H_1 = 25$ m and $H_2 = 20$ m, respectively. The pipe length is $l = 2{,}000$ m and the diameter is $D = 0.5$ m. The Darcy-Weisbach friction factor of the pipe is $F = 0.02$ and the steady-state discharge is $Q_0 = 0.0153$ m^3/s. The wave speed is $a = 1{,}200$ m/s. A valve is located just upstream of the downstream reservoir and its role is to generate the desired transient wave. It is assumed that an impulse wave is generated by rapidly closing and opening the valve; thus, the boundary conditions $h(x^U) = 0$ and $q(x^D) = 1$ are applied. There are two pressure sensors located at 1,800 m and 2,000 m, respectively. Another sensor located at $x^{M_0} = 50$ m is used to estimate $q(x^U)$ via Eq. (8). The location x^{M_0} of this sensor cannot be very close to the upstream node; otherwise, the estimate $\hat{q}(x^U)$ of $q(x^U)$ is very sensitive to noise (a small perturbation of measurement) because the denominator of Eq. (8) is very close to zero. This is because the upstream in this case is a reservoir where pressure fluctuation is zero. That is, the closer the measurement is to this boundary, the smaller is the measured pressure.

In this section, different noise levels are considered to study the performance of the proposed leak localization methods. Zero-mean independent and identically distributed Gaussian white noise is added to all three sensors. The noise level is quantified by SNR, which is defined

$$\text{SNR} = 20\log_{10}(\overline{|\mathbb{E}(\Delta\mathbf{h})|}/\sigma) = 20\log_{10}\left(\frac{s^L}{\sigma}\overline{|\mathbf{G}|}\right) \tag{34}$$

where $\overline{|\mathbb{E}(\Delta\mathbf{h})|}$ = average head difference; and σ = standard deviation of Gaussian white noise. Three issues about this definition of SNR in this paper are clarified as follows:

- The head difference $\Delta\mathbf{h}$, instead of the measured head, is considered for defining SNR. The reason is that $\Delta\mathbf{h}$ offers the main information for leak detection. Furthermore, Eq. (34) shows that increasing SNR is equivalent to decreasing the noise level σ or proportionally increasing the leak size s^L. This implies that the influence of leak size can be equivalently quantified by varying SNR.
- The average of head difference in Eq. (34) is the average over all the frequencies (whole spectrum), rather than the average of the selected frequencies for leak localization. Actually, different choice of frequencies implies totally different average head difference. For example, if only resonant frequencies were used, the average would be very high. In order to have a fair comparison with other cases that may select different frequencies, the average head difference is considered as the average over the

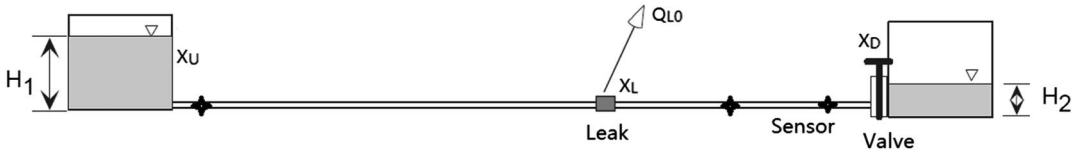

Fig. 5. Setup of the simulation experiment.

whole spectrum (numerically computed by a very fine mesh of frequency) for the reference of SNR, no matter what frequencies are chosen for leak localization.

- Fig. 1 shows that $\overline{|\Delta h|}$ remains almost at the same level for different leak locations x^L, except when x^L is very close to the upstream node of the pipe or the sensor location. Therefore, the proposed definition of SNR in Eq. (34) is almost independent of leak location such that for a given SNR, the leak localization errors for different actual leak locations are comparable.

Single Leak

In this section, the localization of a single leak is considered. The leak is assumed to be located at $x^{L*} = 600$ m, and its size is $s^L = C^d A^L = 1 \times 10^{-4}$ m^2. In the estimation, the resonant and antiresonant frequencies $\{\alpha\omega_{th}{:}\alpha = 1, 2, \ldots, 31\}$ are used, i.e., $J = 31$, and the dimension of $\Delta\mathbf{h}$ and $\mathbf{G}$ is $MJ = 62$. Here, the sample size $N = 10MJ = 620$. Fig. 6 shows the localization results with SNR $= 0$ dB using the four spectral-based methods, in which the correlation matrix is estimated using the PCA method from Eq. (21). In this case, the objective function reaches maximum at the actual leak position for all the four methods. However, the Bartlett's BF has a relatively wide main lobe and some side lobes, particularly a very high secondary lobe at around 1,200 m. The appearance of the high secondary lobe has been shown in Fig. 4 and it disturbs the leak localization because it may be wrongly identified as a second leak, particularly if the leak number is unknown.

In contrast, the other three methods all successfully suppress side lobes and clearly show a narrow main lobe, especially Lagunas' BF, which totally removes all the side lobes, whereas Capon's BF [throughout this section, Capon's BF signifies the normalized version, Eq. (29)] and MUSIC both have a small local maximum at 1,200 m. The cases with SNR $= -30$ dB and SNR $= -40$ dB are displayed in Figs. 7 and 8. For the higher noise levels, all the four methods can still accurately estimate the leak. Bartlett's BF still returns a result with a high secondary peak as the case SNR $= 0$ dB. Capon's BF, Lagunas' BF, and MUSIC all have side lobes in the latter two cases of SNR whose levels increase as SNR decreases. However, Lagunas' BF has lowest side lobes, and the side lobe height of all the three methods is lower than Bartlett's BF.

In order to observe the statistical properties of the leak localization results, the simulation is repeated (from data generation to leak localization, with different realizations of random noise), and the root-mean square error (RMSE) of leak location estimate with various SNR is plotted in Fig. 9(a). Here, the correlation matrix is still estimated using the PCA method by Eq. (21). The Cramé-Rao

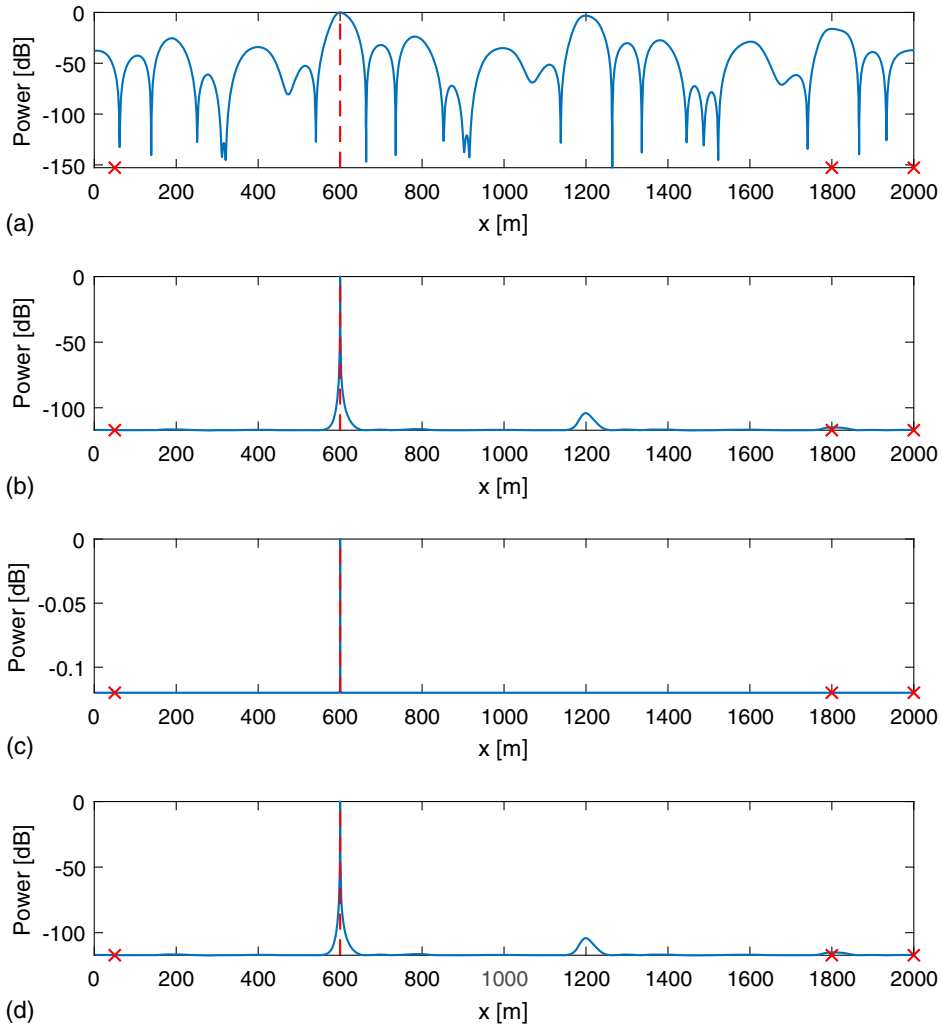

Fig. 6. Localization of a single leak using spectral-based methods by plotting the power spectra: (a) Barlett BF; (b) Capon BF; (c) Lagunas BF; and (d) MUSIC. Dashed line and crosses represent leak and sensor locations, respectively. The SNR of measurement noise is 0 dB.

J. Hydraul. Eng., 2019, 145(3): 04018089

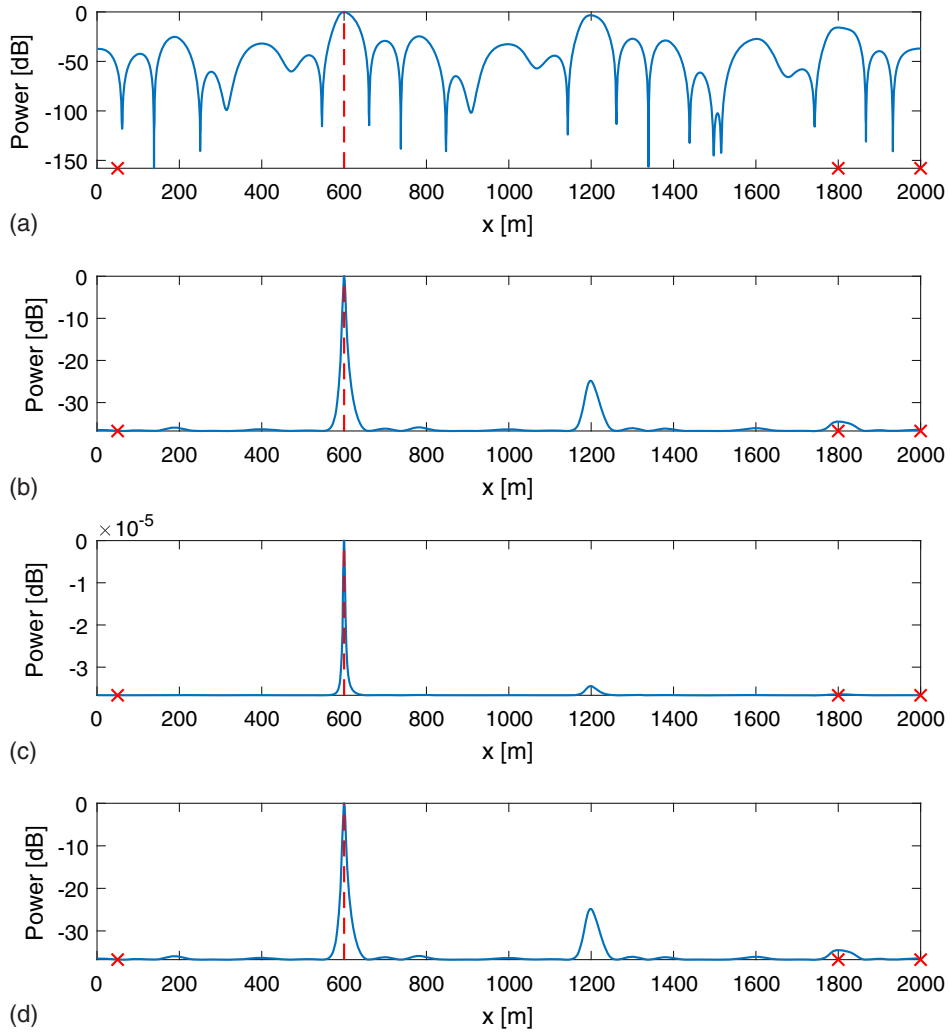

Fig. 7. Localization of a single leak using spectral-based methods by plotting the power spectra: (a) Barlett BF; (b) Capon BF; (c) Lagunas BF; and (d) MUSIC. Dashed line and crosses represent leak and sensor locations, respectively. The SNR of measurement noise is −30 dB.

lower bound (CRLB) of the leak location estimator is also shown [actually, it is the square root of the CRLB of MSE in Eq. (41)]. The CRLB implies the lower limit of RMSE among all possible estimators; its derivation (Wang and Ghidaoui 2018a; Zhou et al. 2018) can be found in the Appendix. The SNR is varied from −40 to −10 dB, and each result shown in Fig. 9 is obtained from 100 simulations. The average head difference is 21.7 m, which is used as the reference in the definition of SNR and is shown by the dotted line in Fig. 1. It is clear that as SNR increases, the localization error (RMSE from the 100 results) of all the four methods decreases, and they have almost the same results. In this SNR range (from −40 to −10 dB), the average localization error of all the four methods is less than 1 m.

Furthermore, the RMSE of all the proposed methods is close to the corresponding CRLB, i.e., the limit of RMSE of any method using the current measurements, which justifies the accuracy of these estimators. In the plot of the spatial power spectrum, the height of global maximum compared with other local maxima is also very important for leak localization because the latter may be wrongly identified as another leak, especially for practical cases where the number of leaks in a pipe is unknown. Here, the ratio of heights between the secondary lobe and main lobe for the four methods with respect to SNR is computed and plotted in Fig. 9(b). It shows that the side lobe height of Bartlett's BF is always high: the ratio of lobe heights is close to 1 for all the SNRs. The other three methods have obviously much stronger ability to suppress the side lobes, particularly Lagunas' BF, which always has the lowest side lobe and almost always has only a single peak.

Figs. 9(c–f) shows the results using the other two covariance matrix estimation methods, SCM in Eq. (19) and DL in Eq. (20). It is clear that in these cases, the accuracy of Lagunas' BF decreases, whereas the other three methods are almost as good as the previous case. The reason might be that the correlation matrices obtained from SCM and DL are less accurate and Lagunas' BF is more sensitive to the correlation matrix estimation accuracy. However, the results of Lagunas' BF are still acceptable when SNR is not lower than −40 dB for SCM and SNR is not lower than −25 dB for DL.

In practice, it is generally not advisable to repeat transient experiment many times to avoid structural fatigue due to multiple loading and to minimize operational disturbance of the system. Therefore, the performance of the proposed methods with fewer measurements is also investigated. Here, 10 measurements are simulated and used for leak localization, i.e., $N = 10$. In order to guarantee the full rank of the correlation matrix estimation, SCM cannot be used; thus, the correlation matrix is computed from DL

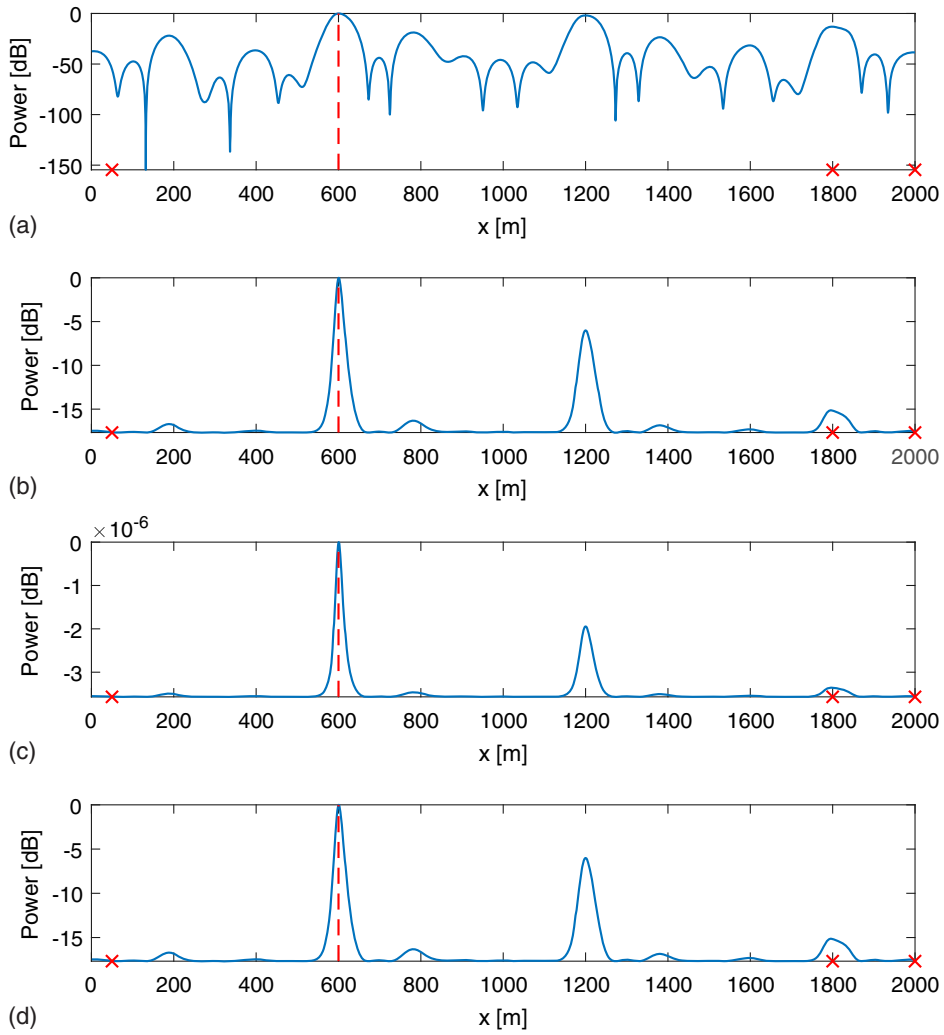

Fig. 8. Localization of a single leak using spectral-based methods by plotting the power spectra: (a) Barlett BF; (b) Capon BF; (c) Lagunas BF; and (d) MUSIC. Dashed line and crosses represent leak and sensor locations, respectively. The SNR of measurement noise is −40 dB.

and PCA via Eqs. (20) and (21), respectively. Fig. 10(a) shows the RMSE of the four methods with PCA along with the corresponding CRLB, the range of SNR is from −10 to 20 dB. For SNR from −5 to 20 dB, the three BF methods return similar results that are accurate enough, whereas MUSIC has a slightly higher error. For a low SNR being −10 dB, all four methods do not work.

Fig. 10(b) displays the height ratio between the secondary and main lobes. Similarly to the previous case, the leak localization using Bartlett's BF is disturbed by the secondary lobe, which is even as high as the main lobe. Again, Lagunas' BF have the best ability to suppress the local maxima. Figs. 10(c and d) show the corresponding results with DL correlation matrix estimation. In this case, due to the imprecision of correlation matrix, Lagunas' BF returns bad results for all SNRs. Again, the other three methods are less sensitive to this imprecision and thus the leak localization result is acceptable.

By comparing Figs. 9(a, c, and f) and 10(a and c), it can be found that for a larger sample size $N = 640$, the estimation error is closer to the CRLB than the case of smaller size $N = 10$, which can be explained by the theory of MLE. Bartlett's BF is equivalent to MLE, which has the property that its MSE converges in probability to CRLB as the sample size tends to infinity.

Finally, leak localization with a smaller sample size $N = 1$–8 is tested and shown in Fig. 11; SNR is, respectively, 20 and 0 dB. Here, PCA is used to estimate the correlation matrix except Bartlett's BF for $N = 1$, where SCM is used. In the case of SNR = 20 dB, Capon's BF, Lagunas' BF, and MUSIC can accurately estimate the leak with minimum $N = 2$ (error is below 1 m on average), whereas Bartlett's BF works even for $N = 1$. In terms of suppressing side lobes, Lagunas' BF again performs best among the four methods. For a lower SNR, more data are required for a precise leak localization. However, Fig. 11(c) shows that $N = 6$, which is fully affordable and applicable in real pipe systems, is enough to guarantee an average error lower than 1 m for all the beamforming methods, even for a very low SNR being 0 dB.

Here, a few concluding remarks for single leak localization using the spectral methods are given as follows:

- PCA most accurately estimates the correlation matrix in the sense of leak localization accuracy. With this method, all the other four methods have almost the same leak localization accuracy, but Lagunas' BF has the best ability to suppress side lobes and is thus preferable.
- Lagunas' BF is more sensitive to the accuracy of correlation matrix estimation than the other three methods. Therefore, if the correlation matrix is badly estimated, Capon's BF and MUSIC are good choices that have lower side lobes than Bartlett's BF.

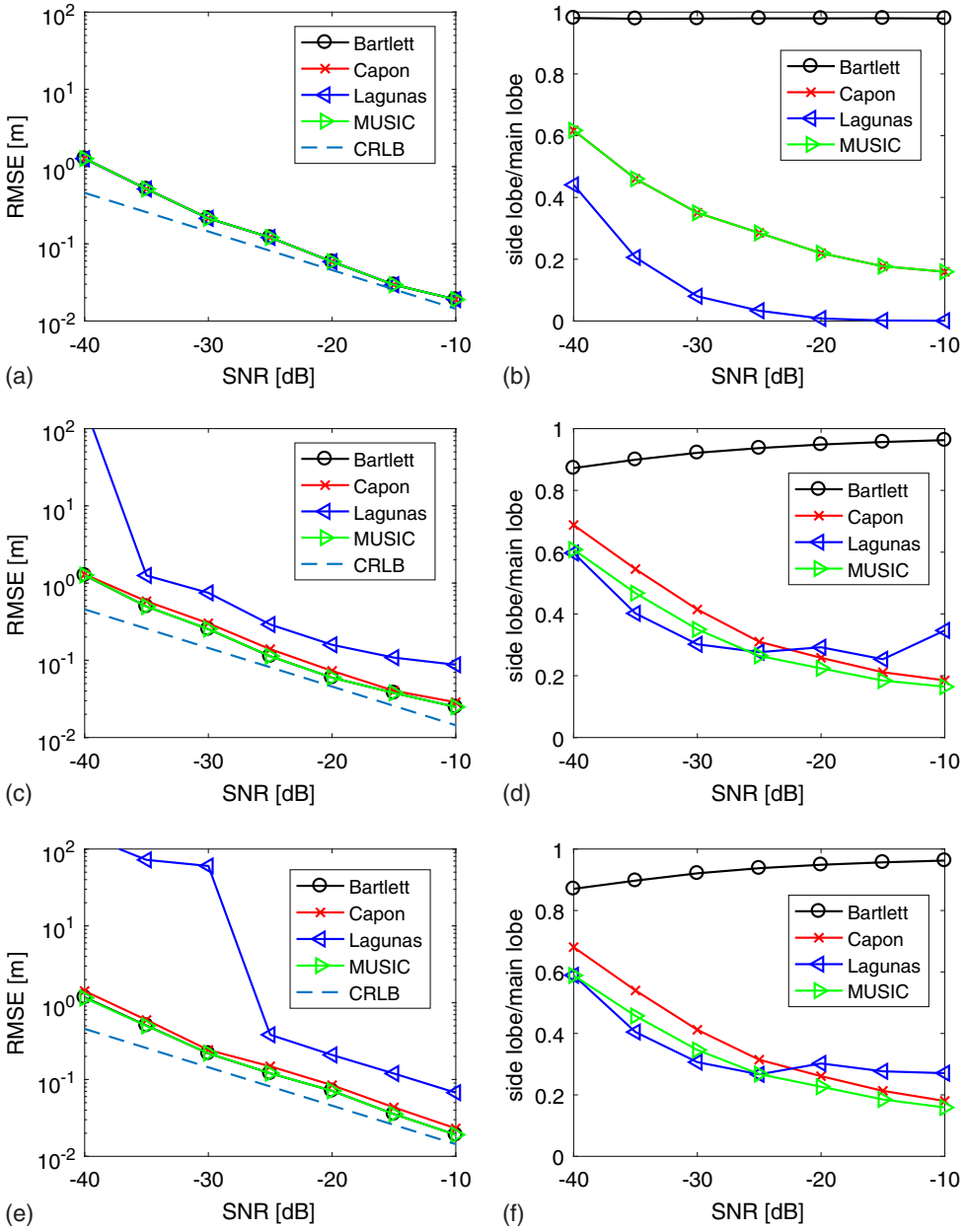

Fig. 9. (a, c, and e) RMSE; and (b, d, and f) ratio between the heights of secondary and main lobes of leak localization using spectral-based methods. The correlation matrix is estimated from (a and b) principal component analysis method; (c and d) sample correlation matrix; and (e and f) diagonal loading method. The sample size used for sample correlation matrix is $N = 640$.

- In the case of a low noise level, only Bartlett's BF can estimate the leak with a single sample, whereas the other methods require at least $N = 2$; for a high noise level, all the methods need more samples and the BF methods are more robust than MUSIC.

Multiple Leaks

In this section, the proposed methods are tested for the case of double leaks. The simulated measurements are obtained from the model of double leaks, but the 1D search is done for leak detection via Eqs. (24), (29), (30), and (33), which implies that just one leak is assumed in the model for the inverse problem. When two leaks are not close to each other, the output function may have two local maxima that correspond to the two leak positions. This property for Bartlett's BF has been discussed by Wang and Ghidaoui (2018b). By this property, one can expect to use the single-leak model to detect multiple leaks.

The leak distance is important for leak localization problem. As illustrated by Wang and Ghidaoui (2018b) and justified by the Nyquist-Shannon sampling theorem, when the distance between two leaks is less than half minimum wavelength, they cannot be separately identified. Here, three cases of long leak distance (much larger than half minimum wavelength which is approximately 129 m in the cases under investigation) are considered first. Fig. 12 shows the localization results using the four spectral-based methods where two leaks are located at $x^{L_1} = 800$ m and $x^{L_2} = 1{,}600$ m where the sizes of the leaks are $s^{L_1} = 10^{-4}$ m^2 and $s^{L_2} = 1.2 \times 10^{-4}$ m^2. The SNR of measurement noise is 0 dB, and the sample size is $N = 10MJ = 620$. The correlation matrix is

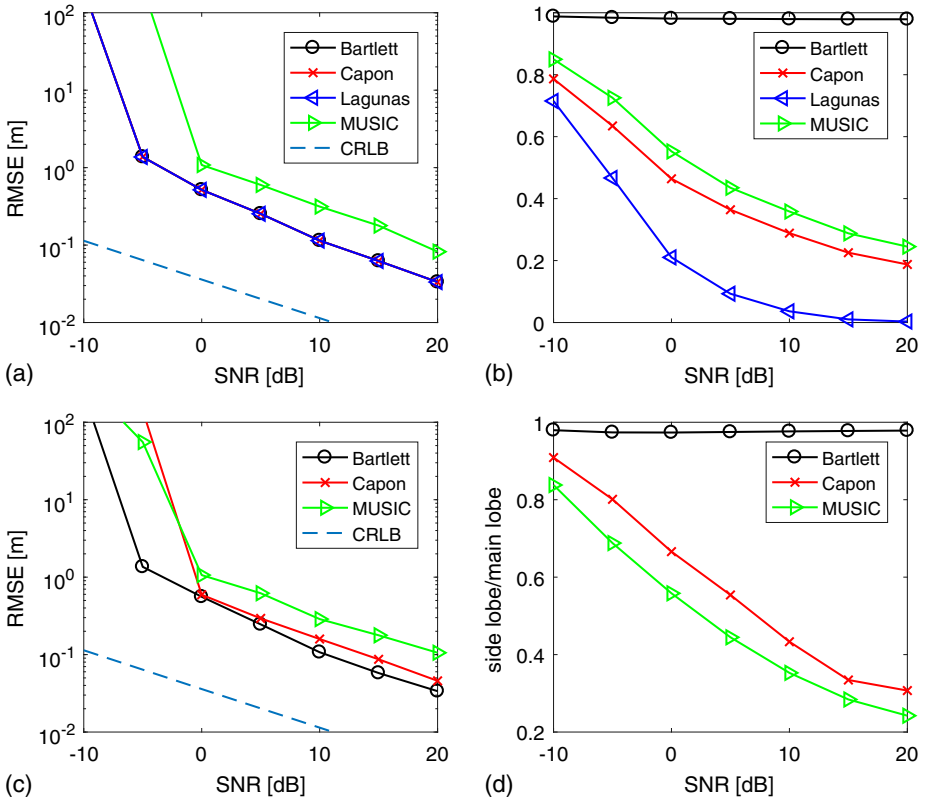

Fig. 10. (a and c) RMSE; and (b and d) ratio between the heights of secondary and main lobes of leak localization using the spectral-based methods. The correlation matrix is estimated from (a and b) principal component analysis method; and (c and d) sample correlation matrix. The sample size used for sample correlation matrix is $N = 10$.

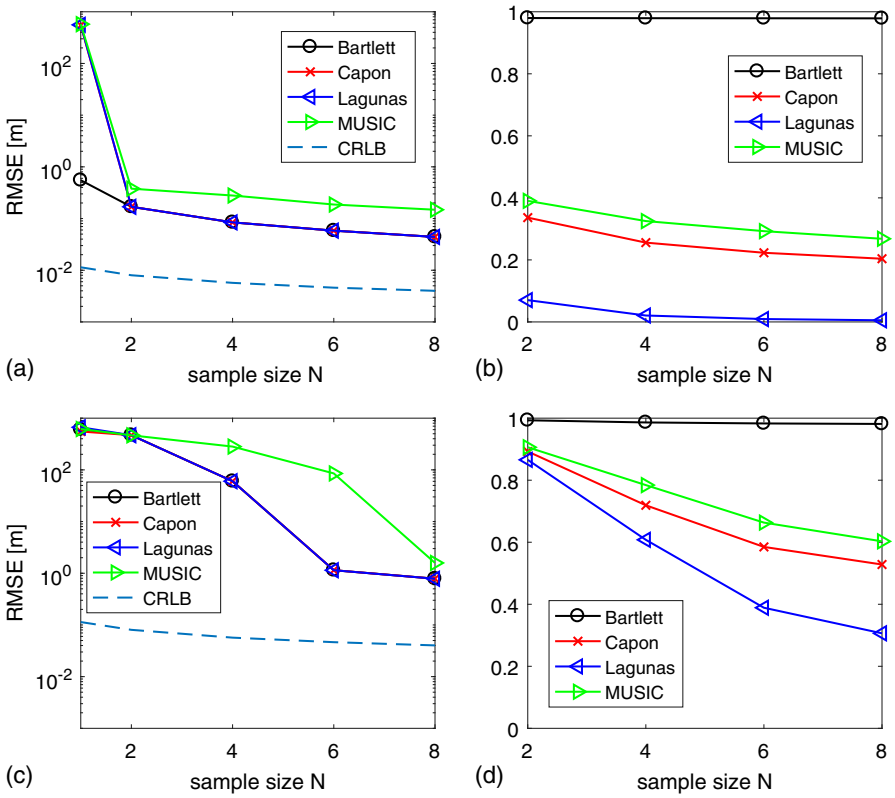

Fig. 11. (a and c) RMSE; and (b and d) ratio between the heights of secondary and main lobes of leak localization using spectral-based methods. The SNR is (a and b) 20 dB; and (c and d) 0 dB. The sample size used for sample correlation matrix is $N = 1, 2, 4, 6, 8$.

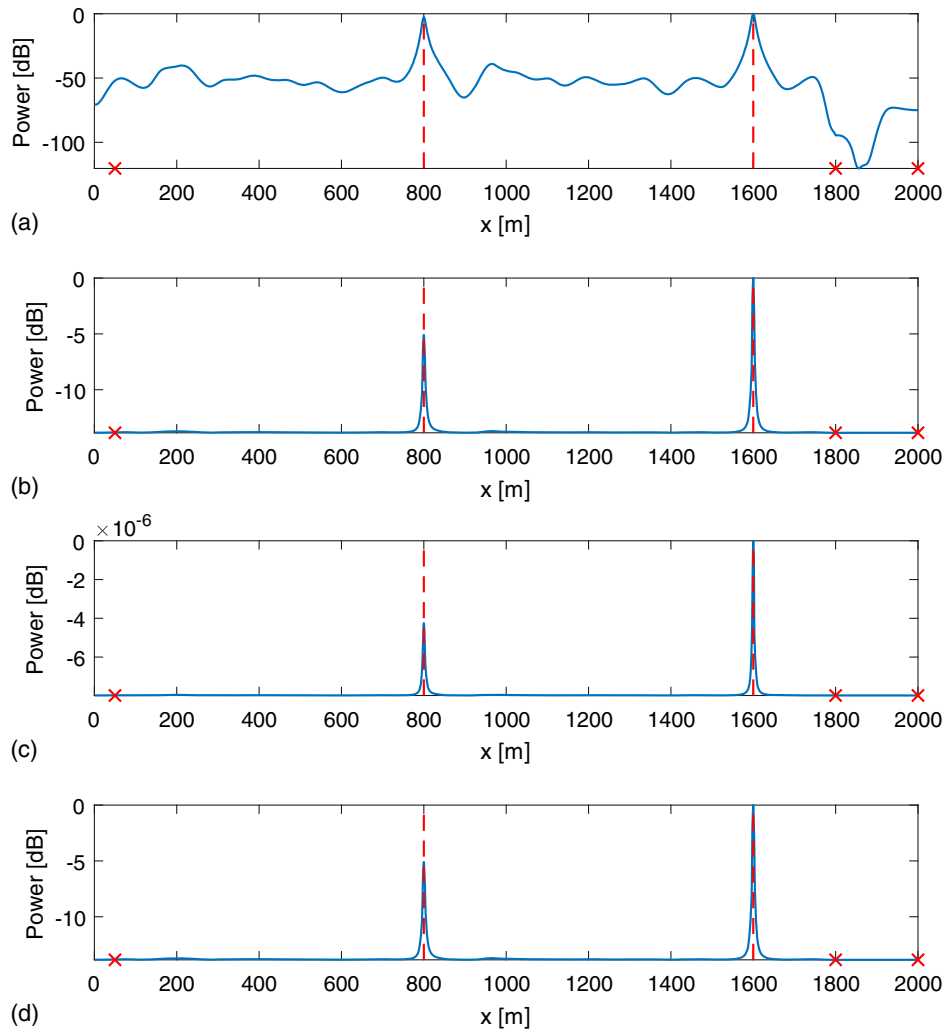

Fig. 12. Localization of double leaks using the spectral-based methods by plotting the power spectra: (a) Barlett BF; (b) Capon BF; (c) Lagunas BF; and (d) MUSIC. Leak locations are $x^{L_1} = 800$ m and $x^{L_2} = 1{,}600$ m. Leak sizes are $s^{L_1} = 10^{-4}$ m^2 and $s^{L_2} = 1.2 \times 10^{-4}$ m^2. The SNR of measurement noise is 0 dB.

estimated via PCA. In this case, all the four methods can accurately localize the two leaks.

Similarly, Figs. 13 and 14 plot the power spectra for two other cases of leak locations, $x^{L_1} = 700$ m and $x^{L_2} = 1{,}600$ m with $x^{L_1} = 600$ m and $x^{L_2} = 1{,}600$ m, respectively. In the former case, the two largest local maxima in each figure are near the actual leaks, but other side lobes appear. In the latter case, each local maximum can be found around each actual leak, but other higher lobes may disturb the determination of leak locations. From the preceding results, one can conclude that the localization efficiency of multiple leaks using the 1D searching methods depends on the leak locations. Furthermore, Capon's BF, Lagunas' BF, and MUSIC are better than Bartlett's BF in the sense of suppressing interferences.

Finally, a case where two leaks are close to each other is considered, where $x^{L_1} = 400$ m and $x^{L_2} = 460$ m. The localization results are shown in Fig. 15. In this case, each power spectrum reaches maximum between the two leaks, which implies that the range of the leaks can be indicated but the two leaks cannot be separately identified. In addition, Lagunas' BF better suppresses the local maxima where no leak exists.

According to the aforementioned results, it is concluded that 1D searching using spectral-based methods is not robust for multiple-leak detection problems that depend on the leak locations and the distance between the leaks. However, these methods are able to roughly estimate the leaks and are thus still valuable. More exactly, at each actual leak location, the spatial power spectrum of each method has a local maximum nearby. By virtue of the low computational cost of the 1D search, these results can be a fast but rough estimate. They can be used as an initial estimate for more complicated methods, for example a method that models the multiple leaks and searches these leaks in the multidimensional parameter space (Wang and Ghidaoui 2018a, 2019). Those methods are able to more accurately localize multiple leaks, albeit at a higher computational cost.

Conclusion

In this paper, four spectral-based methods, specifically three beamforming (BF) methods (Bartlett's BF, Capon's BF, and Lagunas' BF) and a subspace-based method (MUSIC), were employed to detect leaks in a pipeline via transient wave. These approaches involve a low computational cost because they only search a 1D space of leak location along the pipe, independent of the leak size.

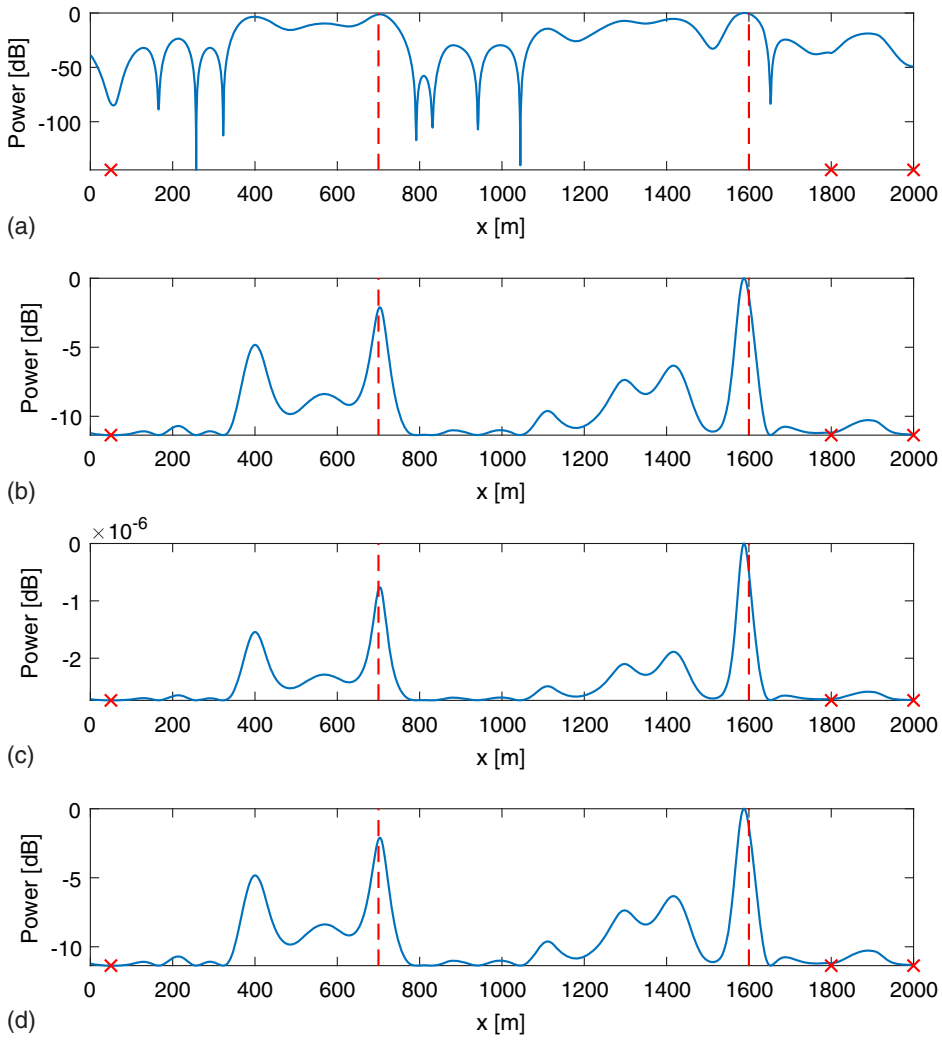

Fig. 13. Localization of double leaks using spectral-based methods by plotting the power spectra: (a) Barlett BF; (b) Capon BF; (c) Lagunas BF; and (d) MUSIC. Leak locations are $x^{L_1} = 700$ m and $x^{L_2} = 1{,}600$ m. Leak sizes are $s^{L_1} = 10^{-4}$ m^2 and $s^{L_2} = 1.2 \times 10^{-4}$ m^2. The SNR of measurement noise is 0 dB.

These approaches were evaluated by leak localization accuracy and disturbance suppression ability. Sample correlation method, diagonal loading, and principal component analysis methods were used for estimating correlation matrices. The numerical simulation showed that all four spectral-based methods for leak detection were able to accurately estimate a single leak, and the PCA method best estimated the correlation matrix. With the latter method, all the leak detection methods were accurate and robust with respect to a high level of noise, among which Lagunas' BF was best at suppressing the local maxima of the objective function corresponding to a wrong location of leak.

These approaches can also be used to localize multiple leaks based on the 1D search. The performance in this case is not so robust as the case of single leak and is limited by leak distance. However, these fast computations from the spectral-based methods can be used as a rough guess or prior information for more precise but computationally expensive multileak detection methods based on a multidimensional search.

In order to experimentally verify the proposed methods, it is necessary to repeat the classical pipeline transient tests. This approach is currently under development by the authors. Once ready, this system will be used to experimentally test the techniques proposed in this paper. Furthermore, pipe-wall viscoelasticity effect, as well as unsteady friction, can be considered in the transient model, and the proposed spectral-based model can be generalized to these scenarios. This problem is also under investigation.

Appendix. CRLB for Leak Location Estimator

Fisher information is a measure of information brought from data. More specifically, for a random variable X with probability density function $p(x|\theta)$, the information of a data sample provided about the unknown parameter θ is quantified by its Fisher information, which is defined by

$$I(\theta) = -\mathbb{E}\left(\frac{d^2 \log p(x|\theta)}{d\theta^2}\right) = -\int \frac{d^2 \log p(x|\theta)}{d\theta^2} p(x|\theta) dx \quad (35)$$

Furthermore, this information determines the lower bound of the variance of an estimator of θ. If $\hat{\theta}$ is an unbiased estimator of θ, i.e., $\mathbb{E}(\hat{\theta}) = \theta$, then

J. Hydraul. Eng., 2019, 145(3): 04018089

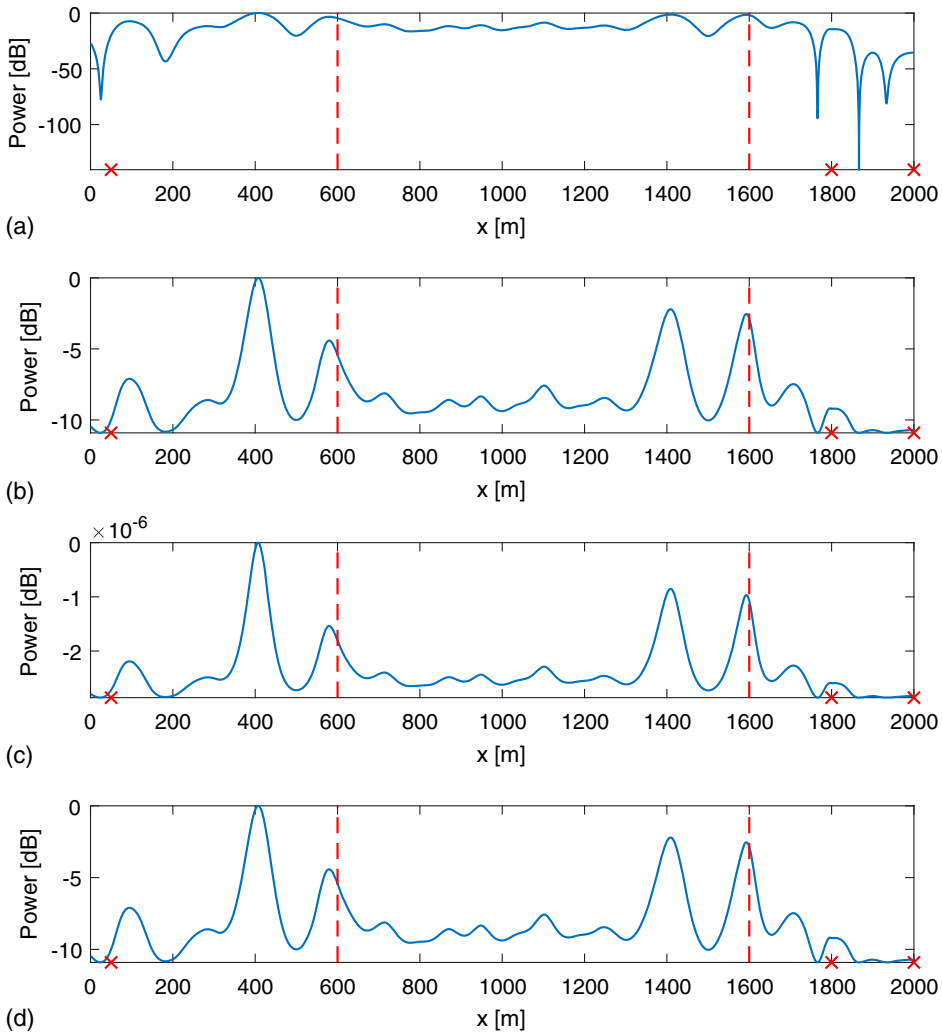

Fig. 14. Localization of double leaks using spectral-based methods by plotting the power spectra: (a) Barlett BF; (b) Capon BF; (c) Lagunas BF; and (d) MUSIC. Leak locations are $x^{L_1} = 600$ m and $x^{L_2} = 1{,}600$ m. Leak sizes are $s^{L_1} = 10^{-4}$ m^2 and $s^{L_2} = 1.2 \times 10^{-4}$ m^2. The SNR of measurement noise is 0 dB.

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{NI(\theta)} \tag{36}$$

The right-hand side of the preceding equation is known as CRLB.

Here, the CRLB of the leak location x^L is computed. The log-likelihood function of $\Delta\mathbf{h}$ is

$$\begin{aligned}\log L(x^L, s^L; \Delta\mathbf{h}) &= -JM\log(\pi\sigma^2) - \frac{1}{\sigma^2}\|\Delta\mathbf{h} - \mathbf{G}(x^L)s^L\|^2 \\ &= -JM\log(\pi\sigma^2) \\ &\quad - \frac{1}{\sigma^2}(\Delta\mathbf{h}^H\Delta\mathbf{h} - s^L\Delta\mathbf{h}^H\mathbf{G} \\ &\quad - s^L\mathbf{G}^H\Delta\mathbf{h} + (s^L)^2\mathbf{G}^H\mathbf{G})\end{aligned} \tag{37}$$

Next, the second-order derivative is computed

$$\begin{aligned}\frac{d^2\mathbf{G}^H\mathbf{G}}{d(x^L)^2} &= ((\mathbf{G}')^H\mathbf{G} + \mathbf{G}^H\mathbf{G}') \\ &= (\mathbf{G}'')^H\mathbf{G} + \mathbf{G}^H\mathbf{G}'' + 2(\mathbf{G}')^H\mathbf{G}'\end{aligned} \tag{38}$$

Therefore

$$\begin{aligned}\mathbb{E}\left(\frac{\partial^2 \log L}{\partial(x^L)^2}\right) &= -\frac{1}{\sigma^2}\mathbb{E}\left(-\Delta\mathbf{h}^H\mathbf{G}''s^L - (\mathbf{G}''s^L)^H\Delta\mathbf{h}\right. \\ &\quad \left. + (s^L)^2\frac{d^2\mathbf{G}^H\mathbf{G}}{d(x^L)^2}\right) \\ &= -\frac{2}{\sigma^2}(s^L)^2\|\mathbf{G}'\|^2\end{aligned} \tag{39}$$

and the Fisher information is

$$I(x^L) = \frac{2}{\sigma^2}(s^L)^2\|\mathbf{G}'(x^L)\|^2 \tag{40}$$

Finally, the CRLB of x^L estimator is

$$\mathrm{CRLB}(x^{L_n}) = \frac{\sigma^2}{2N(s^L)^2\|\mathbf{G}'(x^L)\|^2} \tag{41}$$

where $\mathbf{G}'(x^L)$ is the derivative of Eq. (11). If the pipe is in the same horizontal plane (i.e., $z^L = 0$) and $H_0^L = H_1 + (H_2 - H_1)x^L/l$ (H_1

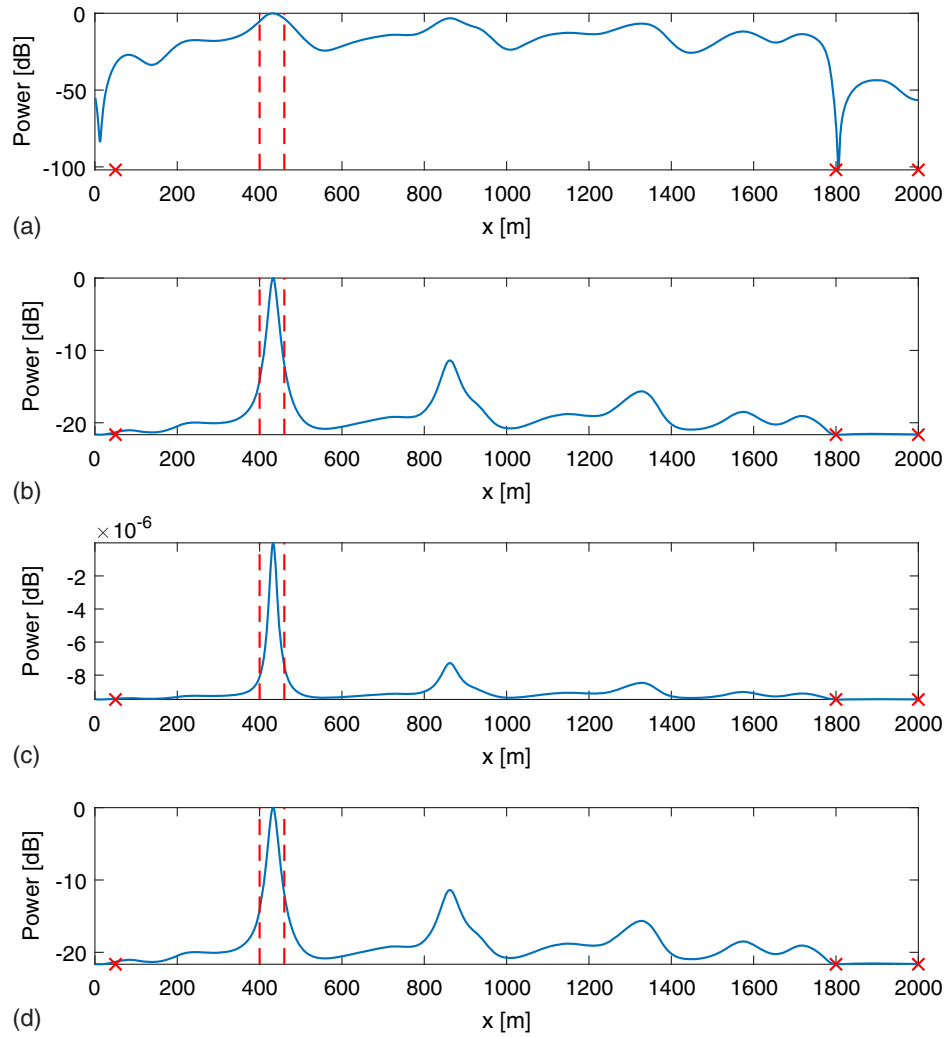

Fig. 15. Localization of double leaks using the spectral-based methods by plotting the power spectra: (a) Barlett BF; (b) Capon BF; (c) Lagunas BF; and (d) MUSIC. Leak locations are $x^{L_1} = 400$ m and $x^{L_2} = 460$ m. Leak sizes are $s^{L_1} = 10^{-4}$ m^2 and $s^{L_2} = 1.2 \times 10^{-4}$ m^2. The SNR of measurement noise is 0 dB.

and H_2 are steady-state heads at the upstream and downstream reservoirs), then

$$G'_{mj}(x^L) = -\frac{\sqrt{g}Z^2\mu q(x^U)}{\sqrt{2H_0^L}}\sinh(\mu(x_m - 2x^L)) + \frac{\sqrt{g}Z^2 q(x^U)(H_2 - H_1)}{2l(2H_0^L)^{3/2}}\sinh(\mu(x_m - x^L))\sinh(\mu x^L) = \frac{\sqrt{g}Z^2 q(x^U)}{\sqrt{2H_0^L}}\sinh(\mu(x_m - 2x^L)) \times \left[-\mu + \frac{H_2 - H_1}{4lH_0^L}\sinh(\mu x^L)\right] \quad (42)$$

Acknowledgments

This work has been supported by research grants from the Research Grant Council of the Hong Kong SAR, China (Project No. T21-602/15R, 16203417, and 16208618) and from Chinese Estates Professorship in Engineering (No. R8031). The authors would like to thank Prof. Miguel Ángel Lagunas for his helpful comments and suggestions.

Notation

The following symbols are used in this paper:

A = pipe area;
a = wave speed;
F = Darcy-Weisbach friction factor;
$\mathbf{G}$ = model of $\Delta\mathbf{h}$ as a function of x^L in Eq. (10);
g = gravitational acceleration;
h = complex head oscillation;
J = frequency number selected for leak detection;
M = sensor number;
N = sample size;
$\mathbf{n}$ = measurement noise;
Q_0^L, H_0^L = steady-state discharge and head of leak;
q = complex discharge oscillation;
$\mathbf{R}$ = correlation matrix;
R = frictional resistance;
s^L = leak size;
x^L = leak location;
x_m = mth measurement station coordinate;
Z = characteristic impedance;

z^L = pipe elevation at leak;
$\Delta \mathbf{h}$ = head difference;
μ = propagation function;
ω = angular frequency; and
ω_{th} = fundamental frequency.

References

Beck, S. B., M. D. Curren, N. D. Sims, and R. Stanway. 2005. "Pipeline network features and leak detection by cross-correlation analysis of reflected waves." *J. Hydraul. Eng.* 131 (8): 715–723. https://doi.org/10.1061/(ASCE)0733-9429(2005)131:8(715).

Bienvenu, G., and L. Kopp. 1980. "Adaptivity to background noise spatial coherence for high resolution passive methods." In Vol. 5 of *Proc., IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, 307–310. New York: IEEE.

Brunone, B. 1999. "Transient test-based technique for leak detection in outfall pipes." *J. Water Resour. Plann. Manage.* 125 (5): 302–306. https://doi.org/10.1061/(ASCE)0733-9496(1999)125:5(302).

Brunone, B., and M. Ferrante. 2010. "Detecting leaks in pressurised pipes by means of transients." *J. Hydraul. Res.* 39 (5): 539–547. https://doi.org/10.1080/00221686.2001.9628278.

Capon, J. 1969. "High-resolution frequency-wavenumber spectrum analysis." *Proc. IEEE* 57 (8): 1408–1418. https://doi.org/10.1109/PROC.1969.7278.

Chaudhry, M. H. 2014. *Applied hydraulic transients*, 3rd ed. New York: Springer.

Colombo, A. F., P. Lee, and B. W. Karney. 2009. "A selective literature review of transient-based leak detection methods." *J. Hydro-Environ. Res.* 2 (4): 212–227. https://doi.org/10.1016/j.jher.2009.02.003.

Covas, D., and H. Ramos. 2010. "Case studies of leak detection and location in water pipe systems by inverse transient analysis." *J. Water Resour. Plann. Manage.* 136 (2): 248–257. https://doi.org/10.1061/(ASCE)0733-9496(2010)136:2(248).

Covas, D., H. Ramos, and A. B. De Almeida. 2005a. "Standing wave difference method for leak detection in pipeline systems." *J. Hydraul. Eng.* 131 (12): 1106–1116. https://doi.org/10.1061/(ASCE)0733-9429(2005)131:12(1106).

Covas, D., H. Ramos, N. Graham, and C. Maksimovic. 2005b. "Application of hydraulic transients for leak detection in water supply systems." *Water Sci. Technol. Water Supply* 4 (5–6): 365–374. https://doi.org/10.2166/ws.2004.0127.

Covas, D., I. Stoianov, J. F. Mano, H. Ramos, N. Graham, and C. Maksimovic. 2010. "The dynamic effect of pipe-wall viscoelasticity in hydraulic transients. Part II—Model development, calibration and verification." *J. Hydraul. Res.* 43 (1): 56–70. https://doi.org/10.1080/00221680509500111.

Ferrante, M., and B. Brunone. 2003. "Pipe system diagnosis and leak detection by unsteady-state tests. Part I: Harmonic analysis." *Adv. Water Resour.* 26 (1): 95–105. https://doi.org/10.1016/S0309-1708(02)00101-X.

Ferrante, M., B. Brunone, and S. Meniconi. 2007. "Wavelets for the analysis of transient pressure signals for leak detection." *J. Hydraul. Eng.* 133 (11): 1274–1282. https://doi.org/10.1061/(ASCE)0733-9429(2007)133:11(1274).

Jolliffe, I. T. 2002. *Principal component analysis*. New York: Springer.

Kashima, A., P. J. Lee, M. S. Ghidaoui, and M. Davidson. 2013. "Experimental verification of the kinetic differential pressure method for flow measurements." *J. Hydraul. Res.* 51 (6): 634–644. https://doi.org/10.1080/00221686.2013.818583.

Kashima, A., P. J. Lee, and R. Nokes. 2011. "Numerical errors in discharge measurements using the KDP method." *J. Hydraul. Res.* 50 (1): 98–104. https://doi.org/10.1080/00221686.2011.638211.

Kingdom, B., R. Liemberger, and P. Marin. 2006. *The challenge of reducing non-revenue water (NRW) in developing countries*. Technical Rep. Washington, DC: World Bank.

Krim, H., and M. Viberg. 1996. "Two decades of array signal processing research." *IEEE Sig. Process. Mag.* 13 (4): 67–94. https://doi.org/10.1109/79.526899.

Lagunas, M. A., M. E. Santamaria, A. Gasull, and A. Moreno. 1986. "Maximum likelihood filters in spectral estimation problems." *Sig. Process.* 10 (1): 19–34. https://doi.org/10.1016/0165-1684(86)90062-9.

Ledoit, O., and M. Wolf. 2004. "A well-conditioned estimator for large-dimensional covariance matrices." *J. Multivariate Anal.* 88 (2): 365–411. https://doi.org/10.1016/S0047-259X(03)00096-4.

Lee, P. J., M. F. Lambert, A. R. Simpson, J. P. Vtkovskỳ, and J. Liggett. 2010. "Experimental verification of the frequency response method for pipeline leak detection." *J. Hydraul. Res.* 44 (5): 693–707. https://doi.org/10.1080/00221686.2006.9521718.

Lee, P. J., J. P. Vtkovskỳ, M. F. Lambert, A. R. Simpson, and J. A. Liggett. 2005a. "Frequency domain analysis for detecting pipeline leaks." *J. Hydraul. Eng.* 131 (7): 596–604. https://doi.org/10.1061/(ASCE)0733-9429(2005)131:7(596).

Lee, P. J., J. P. Vtkovskỳ, M. F. Lambert, A. R. Simpson, and J. A. Liggett. 2005b. "Leak location using the pattern of the frequency response diagram in pipelines: A numerical study." *J. Sound Vib.* 284 (3): 1051–1073. https://doi.org/10.1016/j.jsv.2004.07.023.

Liggett, J. A., and L.-C. Chen. 1994. "Inverse transient analysis in pipe networks." *J. Hydraul. Eng.* 120 (8): 934–955. https://doi.org/10.1061/(ASCE)0733-9429(1994)120:8(934).

Liou, J. C. 1998. "Pipeline leak detection by impulse response extraction." *J. Fluids Eng.* 120 (4): 833–838. https://doi.org/10.1115/1.2820746.

Meniconi, S., B. Brunone, M. Ferrante, C. Capponi, C. A. Carrettini, C. Chiesa, D. Segalini, and E. A. Lanfranchi. 2015. "Anomaly pre-localization in distribution-transmission mains by pump trip: Preliminary field tests in the Milan pipe system." *J. Hydroinf.* 17 (3): 377–389. https://doi.org/10.2166/hydro.2014.038.

Meniconi, S., B. Brunone, M. Ferrante, and C. Massari. 2011. "Potential of transient tests to diagnose real supply pipe systems: What can be done with a single extemporary test." *J. Water Resour. Plann. Manage.* 137 (2): 238–241. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000098.

Mpesha, W., S. L. Gassman, and M. H. Chaudhry. 2001. "Leak detection in pipes by frequency response method." *J. Hydraul. Eng.* 127 (2): 134–147. https://doi.org/10.1061/(ASCE)0733-9429(2001)127:2(134).

Nash, G. A., and B. W. Karney. 1999. "Efficient inverse transient analysis in series pipe systems." *J. Hydraul. Eng.* 125 (7): 761–764. https://doi.org/10.1061/(ASCE)0733-9429(1999)125:7(761).

Nixon, W., M. S. Ghidaoui, and A. A. Kolyshkin. 2006. "Range of validity of the transient damping leakage detection method." *J. Hydraul. Eng.* 132 (9): 944–957. https://doi.org/10.1061/(ASCE)0733-9429(2006)132:9(944).

Paulraj, A., R. Roy, and T. Kailath. 1986. "A subspace rotation approach to signal parameter estimation." *Proc. IEEE* 74 (7): 1044–1046. https://doi.org/10.1109/PROC.1986.13583.

Pisarenko, V. F. 1973. "The retrieval of harmonics from a covariance function." *Geophys. J. Int.* 33 (3): 347–366. https://doi.org/10.1111/j.1365-246X.1973.tb03424.x.

Rubio Scola, I., G. Besançon, and D. Georges. 2017. "Blockage and leak detection and location in pipelines using frequency response optimization." *J. Hydraul. Eng.* 143 (1): 04016074. https://doi.org/10.1061/(ASCE)HY.1943-7900.0001222.

Sattar, A. M., and M. H. Chaudhry. 2010. "Leak detection in pipelines by frequency response method." *J. Hydraul. Res.* 46 (EI1): 138–151. https://doi.org/10.1080/00221686.2008.9521948.

Stephens, M. L. 2008. "Transient response analysis for fault detection and pipeline wall condition assessment in field water transmission and distribution pipelines and networks." Ph.D. thesis, School of Civil and Environmental Engineering, Univ. of Adelaide.

Sun, Y., P. Babu, and D. P. Palomar. 2016. "Robust estimation of structured covariance matrix for heavy-tailed elliptical distributions." *IEEE Trans. Sig. Process.* 64 (14): 3576–3590. https://doi.org/10.1109/TSP.2016.2546222.

Taghvaei, M., S. B. M. Beck, and J. B. Boxall. 2010. "Leak detection in pipes using induced water hammer pulses and cepstrum analysis." *Int. J. COMADEM* 13 (1): 19.

Vtkovskỳ, J. P., A. R. Simpson, and M. F. Lambert. 2000. "Leak detection and calibration using transients and genetic algorithms." *J. Water Resour. Plann. Manage.* 126 (4): 262–265. https://doi.org/10.1061/(ASCE)0733-9496(2000)126:4(262).

Wang, X., and M. S. Ghidaoui. 2018a. "Identification of multiple leaks in pipeline: Linearized model, maximum likelihood, and super-resolution localization." *Mech. Syst. Sig. Process.* 107: 529–548. https://doi.org/10.1016/j.ymssp.2018.01.042.

Wang, X., and M. S. Ghidaoui. 2018b. "Matched-field processing for leak detection in a pipe." *J. Hydraul. Eng.* 144 (6): 04018030. https://doi.org/10.1061/(ASCE)HY.1943-7900.0001476.

Wang, X., and M. S. Ghidaoui. 2019. "Identification of multiple leaks in pipeline. Part II: Iterative beamforming and leak number estimation." *Mech. Syst. Sig. Process.* 119: 346–362. https://doi.org/10.1016/j.ymssp.2018.09.020.

Wang, X.-J., M. F. Lambert, A. R. Simpson, J. A. Liggett, and J. P. Vtkovskỳ. 2002. "Leak detection in pipelines using the damping of fluid transients." *J. Hydraul. Eng.* 128 (7): 697–711. https://doi.org/10.1061/(ASCE)0733-9429(2002)128:7(697).

Wylie, E. B., and V. L. Streeter. 1978. *Fluid transients*. New York: McGraw-Hill.

Zhou, B., A. Liu, X. Wang, Y. She, and V. Lau. 2018. "Compressive sensing-based multiple-leak identification for smart water supply systems." *IEEE Internet Things J.* 5 (2): 1228–1241. https://doi.org/10.1109/JIOT.2018.2812163.