CCA External Project Examples

Using dolphin whistles to identify species.

Dolphin whistles form time-frequency contours that are complex, overlapping, and unpredictable [1]. These features make automatic detection, classification and localisation of whistles difficult and labour-intensive. However, all such whistles fall within a fixed range of time and frequency parameters [2]. It has previously been demonstrated that a dolphin whistle detector can be constructed in MATLAB using this information [3, 4] and it is hoped to build on the achievements of that project by identifying which of the thirty-three species of marine dolphins produced the whistle. The project will start with a brief literature search for dolphin whistle classification techniques and a review of the theory behind any that are found. Then it will proceed to collect and analyse a set of the previously recorded dolphin whistles available at various internet sites, to extract the time and frequency statistics and then to determine the differences between a limited selection of species. Given this information, the main challenge in this project will be to develop a pattern recognition algorithm which identifies signals that match the specific set of parameters for a particular species, whilst rejecting interference such as ambient noise and transient sounds. This algorithm could be based on the previous MATLAB whistle detector [3]; algorithms published in the literature, or could take an entirely new approach. Finally, the probability of detection and false alarm rate of the algorithm will be assessed using a different set of previously recorded whistles.

References

[1] Caldwell, M.C. and D.K. Caldwell, 1965. Individualized whistle contours in bottlenose dolphins (*Tursiops truncatus*). *Nature*, 207 434-435.

[2] Bazúa-Durán, C., 2004. Differences in the whistle characteristics and repertoire of Bottlenose and Spinner Dolphins. *Anais da Academia Brasileira de Ciências*, 76(2) 386-392.

[3] Hoffmann, Franca, 2014. CCA mini-project report: *Dolphin Whistle Detector Algorithms: A New Robust Dolphin Detector using Coherence Enhanced Diffusion*

[4] Dobbins, P. and F. Hoffman, A new, robust, dolphin whistle detector algorithm, *Institute of Acoustics 40th Anniversary Conference*, NEC, Birmingham, 15-16 October 2014

Flow noise reduction in sonar systems.

Flow noise in a sonar system is caused by non-acoustic pressure fluctuations in the immediate vicinity of the receiving transducer [1]. These fluctuations only arise if sonar and water are in relative motion, e.g. if the transducer is towed or attached to the hull of a moving vessel. In this situation, a hydrodynamic boundary layer forms between the moving sonar and the stationary water at some distance away. The pressure fluctuations are caused by turbulence in this boundary layer. A characteristic feature of flow noise is that it is inversely proportional to frequency cubed [2]. It has been observed that large reductions in flow noise could be achieved by surrounding the sonar transducer by a streamlined housing, usually called a sonar dome. Domes reduce flow noise by minimising turbulent flow and by transferring the source of flow noise to a distance from the transducer. The main challenge of this project will be to investigate how one might determine the optimum shape for a sonar dome for a given frequency band and transducer size over a range of realistic vessel speeds. The acoustic characteristics of the dome material should be taken into account [e.g. 3], along with the thickness required to resist deformation due to hydrodynamic forces and any internal structure within the dome.

References

[1] Dowling, A.P., 1998. Underwater flow noise. *Theoretical and Computational Fluid Dynamics*, 10(1) 135-153

[2] Urick, R.J., 1983. *Principles of Underwater Sound*, Peninsula Publishing, Los Altos, CA.

[3] Taylor, J.M. Jr., 1952. *Acoustic Measurements on Edo 60 Inch Lucite Sonar Dome*. Navy Underwater Sound Reference Lab., Orlando, Florida

Matheuristic Approaches for Improving Performance of MATLAB Mixed Integer Linear Programming Solver

Abstract

Mixed Integer Linear Programming problem (MILP) involves minimising a linear objective function, subject to linear constraints, where some of the variables have to be integer. Furthermore, if some of the variables in a MILP problem are required to be binary (i.e. either 0 or 1), the resulting problem is called a 0-1 MILP problem [7].

Various combinatorial optimization problems, including a wide range of practical problems in business, engineering and science can be modelled as 0–1 MILP problems [7]. A number of special cases of the 0–1 MILP problem, such as knapsack, set packing, cutting and packing, network design, protein alignment, travelling salesman, etc., are known to be NP-hard [4] and many examples still remain hard to solve, even with the rapid development of computing resources.

Matheuristics are techniques for solving optimization problems: they are derived by hybridization of metaheuristics and mathematical programming methods. An essential feature of matheuristics is the exploitation of the mathematical model of the problems of interest; therefore they are sometimes also called “model-based” metaheuristics [2, 6]. In the field of MILP problems, matheuristics are known to be able to improve the performance of commercial state-of-the-art solvers. Some examples of methods that were proven to outperform IBM® ILOG® CPLEX® for solving 0-1 MILP problems can be found in [1, 3, 5].

Starting from a review of existing matheuristic methods for solving 0-1 MILP problem, including the latest advances in the field in the last couple of years, the goal of the project will be to compare several matheuristic methods implemented in MATLAB® and report the results, summarising the performance of the methods, with respect to solution quality (as compared to optimal solution - if one exists) and computational time required. Specifically, the aim of the research will be to compare the performance of various matheuristics with that of the built-in MATLAB intlinprog solver and investigate if the performance of intlinprog can be improved in a similar way as that of IBM ILOG CPLEX.

Subject to time permitting, the student will be welcome to suggest further improvements to some of those methods, beyond the current state-of-the-art.

In summary, the project can be divided into the following stages:

1) literature review
2) gathering problem instances
3) implementation of solution method(s) in MATLAB
4) performance comparison of different methods and summary of results
5) suggestions for further research to improve the current state-of-the-art

References

[1] Angelelli, E., Mansini, R., & Speranza, M. G. Kernel search: A general heuristic for the multi-dimensional knapsack problem. *Computers & Operations Research*, 2010:*37*(11), 2017-2026.

[2] Boschetti M.A., Maniezzo V., Roffilli M., Bolufe Rohler, A. Matheuristics: Optimization, simulation and control. In M.J. Blesa, C. Blum, L. Di Gaspero, A. Roli, M. Samples, and A. Schaerf, editors, Hybrid Metaheuristics, volume 5818 of Lecture Notes in Computer Science, pages 171-177, Springer, 2009. ISBN:978-3-642-04918-7.

[3] Fischetti M, Lodi A. Local branching. Mathematical Programming 2003;98(2):23–47. [4]

Garey M., Johnson D., et al. Computers and intractability: a guide to the theory of NP-completeness. San Francisco: WH Freeman; 1979.

[5] Lazić, J., Hanafi, S., Mladenović, N., & Urošević, D. Variable neighbourhood decomposition search for 0–1 mixed integer programs. Computers & Operations Research, 2010:37(6), 1055-1067.

[6] Wikipedia: https://en.wikipedia.org/wiki/Matheuristics

[7] Wolsey, L.A. “Integer Programming,” Wiley-Interscience, 1999.

Subspace Identification Methods for Simulation of 3- phase Rotating Equipment and Fault Detection based on Residual Analysis

Introduction and requirement for this project

Electric motors are used in industry to power a very wide range of industrial equipment. In many of these industrial applications, reliable and predictable performance of the motor and driven equipment system are extremely important. Undetected faults or inappropriate operating conditions can lead to equipment breakdowns and loss of efficiency, and hence there is a requirement for robust techniques to detect these problems. One approach is to analyse the frequency components found in the motor current.

The electric current drawn by the motor depends on the behaviour of the combined motor and driven equipment system, including not only the applied external load, but also internal phenomena such as developing faults. These faults can cause variations in the current drawn by the motor which can be seen as distortions on the current waveform. Since distortions to the current waveform can also be caused by distortions on the voltage waveform, identifying the features associated with equipment faults requires a method for distinguishing the distortions to the current waveform that have NOT been caused by distortions on the voltage waveform.

Established techniques using mathematical modelling via system identification techniques work well but can suffer certain limitations that restrict their usability in some circumstances, particularly where noise is present in the reference signals.

This project is about identifying, testing and demonstrating alternative approaches in order to overcome these issues.

Project briefing

The subspace identification techniques (4SID) constitute a good alternative to classical identification methods (prediction error and least squares, see [3]), and especially for multiple-input and multiple-output (MIMO) linear systems. Three main basic subspace-based approaches are considered for solving system identification problems: Multivariable Output Error state Space (MOESP) [4], Numerical algorithm for Subspace State Space System Identification (N4SID) [5], and Canonical Variate Analysis (CVA) [4]. There is no optimal criterion for selecting one of the techniques amongst them, and characteristics like ill- conditioning of the problem needs to be taken into consideration (see [6]).

Relevant Mathematical aspects

Consider a Linear Time-Invariant (LTI) discrete-time state space model, described by

$$x_{k+1} = A\,x_k + B u_k + w_k \qquad (0)$$

$$y_k = C x_k + D u_k + v_k \qquad (0)$$

where $x_k \in R^n$ is the n-dimensional state vector at time k, $u_k \in R^m$ is the input vector, $y_k \in R^l$ is the output vector, A, B, C, and D are real matrices, and $\{w_k\}$, $\{v_k\}$ are zero mean, stationary ergodic state and output disturbance or noise sequences, uncorrelated with $\{u_k\}$ and the initial state of (1), with covariances satisfying the relation

$$\left\{ u_p \quad {}^{q} \quad {}^{q} \right\} = \begin{bmatrix} Q & S \\ S^T & R_v \end{bmatrix} \delta_{pq} \geq 0 \qquad (0)$$

where ε denotes the expected value operator and δ_{pq} is the Kronecker delta symbol. The matrix pair (A , C) is assumed observable, and (A ,(B $Q^{1/2}$)) controllable. A particular model, in *innovation form*, is

$$x_{k+1} = A x_k + B u_k + K e_k \qquad (0)$$

$$y_k = C x_k + D u_k + e_k \qquad (0)$$

where $\{e_k\}$ is a white noise sequence, and K is the *Kalman gain matrix* (see [1] for detailed explanation).

In system identification problems, the system order, n, and the quadruple of system matrices (A, B, C, D) have to be determined (up to a system similarity transformation) using the input and output sequences, $\{u_k\}$ and $\{y_k\}$, $k = 1:t$. In addition, and depending on the purpose of the model, the Kalman gain matrix K in (3), as well as the state and output noise covariance matrices in (2) may also need to be identified.

In the framework of this project, the identified model is used for simulation purposes (i.e. by setting $K = 0$, see [2]) and fault detection. For instance, given the initial estimate $\hat{x}_1$, and the trajectory $\{u_k\}$, the simulated output can be computed recursively using the formulas

$$\hat{y}_k = C \hat{x}_k + \hat{D} u_k \qquad (0)$$

$$\hat{x}_{k+1} = \hat{A} \hat{x}_k + \hat{B} u_k \qquad (0)$$

where the estimated quantities have been marked by hat signs. The difference between the estimated and the real output(s) - referred to as *residual(s)* - is then used for fault detection by analysis of the *Power Density Spectrum* of the residual(s).

Background and Project Structure

Starting from a review of the existing 4SID methods for simulation purposes, the goal of the project will be to compare several of those methods implemented in SCILAB and VISUAL BASIC (VB), summarizing the performance of the methods with respect to:

- Convergence and stability of the results for different datasets of the same motor-equipment,
- Quality of the residual analysis,
- Optimal user-define parameters, and
- Computational time required.

The student may decide to implement one or more 4SID techniques, or combine them to improve the performance of the existing ones.

In summary, the project can be divided into the following stages:

a) Literature review,
b) Gathering problem instances,
c) Implementation of solution method(s) into SCILAB and/or VB, d)
Performance comparison of different methods,
e) Optimization of user-define parameters for the selected method(s), and f)
Suggestions for further research for improving the current state-of-art.

References

[1] A. K. Tangirala. Principles of system Identification, Theory and Practice. ISBN:13: 978-1-4398-9602-0, pp. 656-718.

[2] A. K. Tangirala. Principles of system Identification, Theory and Practice. ISBN:13: 978-1-4398-9602-0, pp. 702-703.

[3] High-performance numerical algorithms and software for Subspace-based linear systems. *Journal of Computational and Applied Mathematics,* 170 (2004)371–397. DOI: 10.1016/j.cam.2003.12.046

[4] P. Trnka. Subspace Identification Methods. *Technical Report*, Czech Technical University in Prague, Faculty of Electrical and Engineering, Department of Control Engineering. Sep 2005.

[5] P. Van Overschee and B. De Moor. N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems. *Automatica*, Vol.30, No. 1, pp. 73-93, 1994. 0005-1098/94.

[6] S. Hachicha, M. Karrat and A. Chaari. N4SID and MOESP Algorithms to Highlight the Ill-conditioning Problem into Subspace Identification. *International Journal of Automation and Computing,* 11(1), February 2014, 30-38. DOI:10.1007/s11633-014-0763-z.

Twinkling in sonar systems.

Turbulence and other inhomogeneities in the water column cause fluctuations in propagating acoustic signals in the same way that turbulence in the interstellar medium causes stars to twinkle [1,2]. The amplitude fluctuations bring about signal fading and failure to detect targets well within the theoretical range of the system. Phase fluctuations, however, cause loss of directivity and angular resolution in receiving arrays, spreading of transmitted beams, variations in the apparent arrival direction of signals and fluctuations in their arrival time. Fluctuations also result in an occasional high peak in the signal amplitude, allowing sources or targets to be detected at ranges much greater than predicted by the conventional sonar equation [3]. The aim of this project will be to model the underwater medium as a random phase- changing screen [4] and predict the detection range that might be obtained using these occasional high peaks.

References

[1] Hewish, A., 1974. *Pulsars and High Density Physics*, Nobel Lecture.

[2] Uscinski, B. J., 1977. *The Elements of Wave Propagation in Random Media*, McGraw-Hill.

[3] Uscinski, B. J. and M. Spivack. 2005. High Intensity Ribbons in Multiply Scattering Media. *Waves in Random Media* 15(3) 339-352.

[4] Buckley, R., 1975. Diffraction by a Random Phase-Changing Screen: a Numerical Experiment. *J. Atmos. Terr. Phys*. 37(11)1431-1457

Optimization techniques for neural nets [numerical]

In the last decade artificial neural nets have become practically useful, being used commercially in handwriting recognition (automatic letter sorting) and speech recognition. Applications are now being found in many other classification or recognition tasks (e.g. image labelling and fraud detection). In the near future it is likely that computer games and robotics will include more and larger neural nets: the problem then being the number of parameters.

The resurgence of neural nets is due to a number of reasons: the internet is a large and readily available source of training and test data; algorithms for many of the difficult tasks such as learning algorithms and how to implement short term memory have been found; and the computing power available is sufficient to train large nets (millions of weights) in a reasonable time (i.e. days). If the nets could be trained more rapidly or more accurately then progress in the field could be accelerated- it is this problem that will be explored in this project.

A neural net consists of nodes that take signals from many other nodes or the input, sum them according to a set of weights and apply a monotonic, usually non-linear function to the sum to produce an output. Training a neural net means finding the weights. During training, the information available is in the form of weight derivatives. There are standard batch optimization techniques that can be used if the full training set is used (e.g. conjugate gradients, LBFGS- see [1]) but if the data sets are large and highly redundant- which is often the case- it is quicker to train on a subset of the training data, a "mini-batch", and then use a different training set on the next mini-batch. This leads to a stochastic gradient which breaks the assumptions of the batch optimization techniques. A number of stochastic gradient techniques have been used [2,3]: back propagation (gradient descent), back propagation with (Nesterov) momentum, etc.; and also a range of techniques for overcoming problems with widely differing component magnitudes: adaptive step sizes and scaling strategies (e.g. rprop, rmsprop).

The project will compare a new algorithm which only uses the gradient direction with existing algorithms to determine its strengths and weaknesses. The model problem will be to identifying hand-written digits in the MNIST database using single and multi-layer classifiers. The new method is based on a particular discretization of the following updates:

$$\frac{ds}{dt} = a|s|\frac{g}{|g|}$$

$$\frac{dp}{dt} = |s|\left(\frac{s}{|s|} + \beta\frac{g}{|g|}\right)$$

where g is the gradient, s an internal search direction, and p the current position. The precise discretization will be supplied.

The project would suit a candidate with some programming experience.

[1] Press WH, Teukolsky SA, Vetterling WT, Flannery BP, "Numerical recipes", CUP

[2] Sutskever I, Martens J, Dahl G, Hinton G, "On the importance of initialization and momentum in deep learning", Proceedings of the 30th Conference on Machine Learning, Atlanta, USA, 2013.

[3] Schaul T, Zhiang S, LeCun Y, "No More Pesky Learning Rates", arXiv:1206:1106

Modelling the drug supply chain in clinical trials

Clinical trials are run across many sites, and it is necessary to supply each site with the drugs that are needed for the patients that it will treat. It is important that whenever a patient turns up at a site we can supply them with the drugs they need; failure to do so is bad for the patient and can compromise the whole trial. The manufacture, transport, and storage of all these novel compounds is slow and difficult, so in order to make rapid progress and develop new medicines quickly and efficiently we want to only produce and store just enough drug, whilst controlling the risk of being unable to supply a patient with the drugs that they need.

For our purposes we can model the supply chain as a tree, with a factory as the root node. This may supply one or more depots, which may supply other depots or sites directly (and so on). So depots and sites each stock some level of drug, which is decreased as they supply patients or child depots and sites. But as the level of drug decreases they are free to order from their upstream.

Patients arrive at a site randomly (say, according to a Poisson process with known --or assumed ---rate) and are then randomized to a treatment arm. Within each treatment arm patients may — randomly — undergo dose escalations or reductions that unexpectedly change their need for drug from that which is expected.

A number of questions may be explored within this project, for example:

1. Optimal strategy: How and when should sites and depots order from their upstream? Can anything analytical be clone? We want to minimize the costs of manufacture, shipping, and storage, but also control the risk of being unable to supply a patient with the drugs that they need. Is there a good numerical approach to finding the optimum strategy?

2. Constrained supply, few patients, many sites: if patients that qualify for the trial are uncommon it's necessary to open many sites in order to recruit enough patients. Most sites won't recruit anyone; some will recruit many more than average. If supply is so constrained that it's not possible to start the trial with drug at every site how should we proceed?

3. High cost comparator: if one of the drugs in the trial has disproportionately high costs how does this affect supply chain strategies?

4. Anything else: anything that the candidate deems to be both useful and interesting.

Optimal rational approximation and the double interval Zolotarev problem in the complex plane

Problem statement

We consider the near best L^∞ approximation of $1/\sqrt{z}$ (or closely related functions) by a rational function, over certain curves in the complex plane.

Zolotarev's fourth problem [1], [6] provides the best rational approximation $r_{n-1,n}(z)$ of type $(n-1, n)$ to $\gamma = 1/\sqrt{z}$ over a real interval $I = [a, b]$, where $0 < a < b$ in the sense that

$$r_{n-1,n} = \arg\min_{r_{n-1,n}} \left\| \frac{1/\sqrt{z} - r_{n-1,n}(z)}{1/\sqrt{z} + r_{n-1,n}(z)} \right\|_{L^\infty(I)} . \tag{1}$$

Here, $r_{n-1,n}$ is a real rational function of type $(n-1, n)$ with a monic denominator, and I is the real interval $[a, b]$. Zolotarev's explicit solution involves Jacobi elliptic functions to determine the interpolation points. To make use of this in applications, we require the rational approximation $r_{n-1,n}(z)$ should be expressed as a Stieltjes continued fraction (or S-fraction) [2] with coefficients $\hat{h}_i$ $(0 \le i \le n-1)$ and h_i $(1 \le i \le n)$:

$$r_{n-1,n}(z) = \cfrac{1}{\hat{h}_0 z + \cfrac{1}{h_1 + \cfrac{1}{\hat{h}_1 z + \ldots \cfrac{1}{h_{n-1} + \cfrac{1}{\hat{h}_{n-1} z + \frac{1}{h_n}}}}}} . \tag{2}$$

The Zolotarev rational approximation to $1/\sqrt{z}$ over the real interval $[a, b] \in \mathrm{R}^+$, where $0 < a < b$, is readily extended to the real interval $[c, d] \in \mathrm{R}^-$ where $c < d < 0$. Recently [5] showed how to construct a nearly-optimal rational approximation to $1/\sqrt{z}$ on the union of two real intervals

$$I = [a, b] \cup [c, d] \text{ , where } c < d < 0 < a < b . \tag{3}$$

For practical use in numerical analysis (see Appendix), we actually require the best approximation to a modified square root $\gamma_1 = 1/\sqrt{z + \frac{z^2}{4}}$ rather than $\gamma = 1/\sqrt{z}$ over the union of *two* disjoint finite intervals in the *complex* plane:

$$I = I_{pr} \cup I_{ev} , \tag{4}$$

$$\Im(z) < 0 \text{ and } \Re(z) \le 0 \quad \forall z \in \Gamma_{pr}, \tag{5}$$
$$\Im(z) > 0 \text{ and } \Re(z) \le 0 \quad \forall z \in \Gamma_{ev}. \tag{6}$$

We use the same approach from [5] to construct a near-best rational approximation over the union of complex intervals

$$\Gamma_{pr} = [c, d] - iE, \tag{7}$$
$$\Gamma_{ev} = [a, b] - iE, E \in \mathrm{R}^+. \tag{8}$$

The approach we use to construct a near best approximation, akin to (1), follows [5], modified for complex intervals:

- Construct the $2n_1$ Zolotarev interpolation points for $\sqrt{z}$ on the interval $[a/b, 1]$ and the $2n_2$ interpolation points for $\sqrt{z}$ on the interval $[-1, -d/c]$ separately. Here we require $n_1 + n_2 = n$.
- Scale these interpolation points to $[a, b]$ and $[c, d]$ and shift them by $-iE$.
- Calculate the unique complex rational interpolant $r_{n-1,n}(z)$ of type $(n-1, n)$ which interpolates $\gamma_1(z)$ at these $2n$ points.
- Calculate the coefficients of the S-fraction (2).

For practical reasons, the value of n used in the rational approximation $r_{n-1,n}(z)$ of type $(n-1, n)$ should be small, say $5 < n \le 20$. We specify n and ask how to construct $r_{n-1,n}(z)$ which is the near best rational approximation to $\gamma_1(z)$ in order to minimize relative L^∞ error in the sense of (1). Since we solved a rational interpolation problem for $\gamma_1(z)$ rather than $\gamma(z)$, the Zolotarev points are no longer optimal, but are nearly optimal for minimising the relative L^∞ approximation error.

The proposed project has two aims:

- How well does the theory from [5], [2] for real intervals extend to complex intervals for approximating $1/\sqrt{z}$, and how good is the near-best approximation to $\gamma_1(z)$ in this case ?
- What is the best numerical method for obtaining the rational approximation of $\gamma_1(z)$ over complex intervals and to obtain its S-fraction representation ? Is the problem always well-posed ?

A major part of the project would be to review the literature: for example how far do the results from [3] provide an answer to this problem ? A related problem of practical importance is when Γ_{pr} and Γ_{ev} are not simply intervals shifted parallel to the real axis, but are smooth, non-intersecting, finite curves in the 3^{rd} and 4^{th} quadrants of the complex plane.

For a single interval $\Gamma = [a, b] \in \mathrm{R}^+$, it is known that all the coefficients $\hat{h}_i$ and h_i in (2) are real, positive and increasing if $r_{n-1,n}$ has negative simple poles and positive residues. Is it possible to say anything about the coefficients when there are two complex intervals ?

Appendix: practical application

This appendix has been added to motivate a practical application of near-best complex rational approximation. When constructing numerical boundary conditions for the wave equation, we require to approximate the so-called Neumann-to-Dirichlet (NtD) map on the boundary of a domain. Amazingly, in simple but representative cases, this can be reduced to the problem of finding the near best rational approximation of $1/\sqrt{z}$ which minimizes the relative L^∞ error, over certain intervals in the complex plane. Another practical application is in numerical linear algebra, when designing preconditioners for iterative solution of the Helmholtz equation [4]. The use of complex intervals I_{pr} and I_{ev} is required in this case.

The coefficients in (2) are used to construct numerical boundary conditions and represent the Neumann- to-Dirichlet map for the scalar wave equation for $u(\mathbf{x})$

$$k^2(\mathbf{x})u + \nabla^2 u = 0, k^2 \in \mathbb{C} \quad , \Im(k^2) > 0 \quad , \Re(k^2) \geq 0. \tag{9}$$

For $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, consider the half space $x_1 \leq 0$. Assuming k to be constant, and taking the Fourier transform in the x_2-direction with dual variable ξ_2 leads to an equation for $u(x_1, \xi_2)$:

$$k^2 u + \frac{\partial^2 u}{\partial x_1^2} - \xi_2^2 = 0, \tag{10}$$

which has fundamental solutions

$$u(x_1, \xi_2) = e^{\pm \sqrt{(\xi_2^2 - k^2)}x_1}. \tag{11}$$

Well-posedness requires the solution to be finite as $x_1 \to -\infty$. Hence, for $\xi_2 \in \mathbb{R}$, the symbol of the Neumann-to-Dirichlet (NtD) operator at $x_1 = 0$, for given ξ_2, is $-1/\sqrt{\xi_2^2 - k^2}$. The practical application of the Zolotarev approximation to $1/\sqrt{z}$ appears when (10) is discretized by a finite difference scheme - see [2], [5] and [4] for details. Consider the discrete form of (10),

$$(k^2 - \xi_2^2)u_j + \frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} = 0, \tag{12}$$

for $-\infty < j \leq 0$, with mesh size h. The dispersion relation corresponding to the well-posed solution leads to a one-sided numerical approximation of the inverse of the NtD (i.e. the Dirichlet-to-Neumann map or DtN),

$$-\frac{u_0 - u_{-1}}{u_0} = -\frac{1}{2}z + \sqrt{z + \frac{z^2}{4}}, \tag{13}$$

where $z = -(k^2 - \xi_2^2)h^2$. This motivates the rational approximation for $\gamma_1(z)$. The S-fraction (2) corresponds to a finite-volume discretization of (10),

$$(k^2 - \xi_2^2)u_j + \frac{1}{\hat{h}_j}\left(\frac{u_{j+1} - u_j}{h_j} - \frac{u_j - u_{j-1}}{h_{j-1}}\right) = 0, \tag{14}$$

for $j = 1, \ldots, n-1$, and with the additional constraint $u_n = 0$. The coefficients h_j and $\hat{h}_j$ are step lengths on primary and dual grids - see [2] and [5] for details.

Only a very brief motivation for the problem has been given here in the interest of space. This relates the optimal numerical approximation of certain rational approximations in the complex plane to practical applications.

References

[1] N. I. Achieser: *Theory of Approximation*, Frederick Ungar, New York, 1956.

[2] Sergey Asvadurov, Vladimir Druskin, Murthy Guddati and Leonid Knizhnermann: *On optimal finite-difference approximation of PML*, SIAM J. Numer. Anal., 41(1), pp.287-305, 2003.

[3] B. Le Bailly and J. P. Thiran: *Optimal Rational Functions for the Generalized Zolotarev Problem in the Complex Plane*, SIAM J. Numer. Anal., 38(5), pp.1409-1424, 2000.

[4] P. N. Childs, V. Druskin, L. Knizhnerman: *Preconditioning the Helmholtz Equation using approximate Dirichlet-to-Neumann operators on optimal grids*, IMA Conference on Numerical Linear Algebra and Optimisation, University of Birmingham, 3rd-5th September 2014.

[5] Vladimir Druskin, Stefan Guttel and Leonid Knizhnerman: *Near-optimal perfectly matched layers for indefinite Helmholtz problems*, MIMS EPrint: 2013.53, University of Manchester, 2013.

[6] A. A. Gončar: *Zolotarev Problems connected with rational functions*, Mat. Sbornik, Tom 78(210), no. 4, 1669.