Optimized spectral analysis in magnetic resonance spectroscopy for early tumor diagnostics

Karen Belkić¹-³ and Dževad Belkić⁴
¹ Department of Oncology-Pathology, Karolinska Institute, Stockholm, Sweden
² School of Community and Global Health, Claremont Graduate University, USA
³ Institute for Prevention Research, University of Southern California, USA

Email: Karen.Belkic@ki.se

Abstract. Molecular imaging through magnetic resonance spectroscopy (MRS) can provide information about key metabolites. Conventional applications of MRS are hampered by data analysis via the fast Fourier transform (FFT). Most MRS studies for cancer detection have relied upon estimations of a mere handful or even a single composite metabolite, e.g. total choline. These have yielded incremental improvements in diagnostic accuracy. In vitro studies reveal richer metabolic information for identifying cancer, particularly in closely-overlapping resonances. Among these are phosphocholine, a marker of malignant transformation. The FFT cannot assess these congested spectral components. This can be done by the fast Padé transform (FPT), an advanced, high-resolution, quantification-equipped method, applied to MRS time signals as encoded from patients with breast cancers and other cancers, with benign pathology and with normal tissue, as illustrated here in the latter. With realistic noise levels, the FPT accurately computes the metabolite concentrations, including phosphocholine, which completely underlies phosphoethanolamine. In sharp contrast, the FFT produces a rough envelope spectrum with only a few shortened, broadened peaks, and key metabolites altogether absent. The FPT clearly separates true metabolites from spurious resonances. The efficiency and high resolution of the FPT translates into shortened examination time of the patient. These capabilities strongly suggest that by applying the FPT to time signals encoded in vivo from breast cancer and other malignancies, MRS will fulfill its potential to become a clinically-reliable, cost-effective method for early cancer detection.

1. Introduction
Magnetic resonance-based modalities are promising for early cancer detection, particularly since they do not entail exposure to ionizing radiation. Magnetic resonance imaging (MRI) is usually very sensitive, but is costly and not sufficiently specific. Molecular imaging through magnetic resonance spectroscopy (MRS) can potentially improve the specificity of MRI, by providing information about key metabolites. In this paper, we will focus upon MRS-based diagnostics and how these can be improved through optimization of signal processing methods. The actual problem area addressed will be related to screening and early detection of breast cancer. This problem is chosen in light of the fact that worldwide, breast cancer is the most commonly diagnosed malignancy among women and a major cause of death [1]. Survival is profoundly impacted by the stage at diagnosis [2]. Screening significantly improves survival, since with early detection, adequate and timely care can be provided [3]. However, consensus has not been reached as to the best strategies for early breast cancer detection.
There is widespread recognition that screening programs need to be improved and made more cost effective [1]. We examine these topics from the standpoint of mathematical optimization via advanced signal processing, as this aspect of the problem pertains to MRS for breast cancer diagnostics.

1.1. Screening for breast cancer: The current status of diagnostic imaging

Currently, anatomical imaging through mammography together with ultrasound and magnetic resonance imaging (MRI) is the mainstay for breast cancer screening, with MRI generally showing the best sensitivity [5]. However, MRI has poor specificity and, as such, is associated with high biopsy rates, very often for benign lesions [3]. Expense is another drawback of MRI, such that it is currently considered cost-effective only for women at very high breast cancer risk [6].

Since these anatomic imaging modalities are insufficiently specific, functional and molecular imaging through magnetic resonance (MR) are garnering attention vis-à-vis breast cancer diagnostics [7]. Diffusion-weighted MRI (DWI) helps distinguish benign lesions and normal breast from cancer. Since malignancy is often associated with increased cellularity or intracellular density, the random motion of free water may be restricted, as reflected in a decreased apparent diffusion coefficient (ADC). However, a low ADC can occur with fibrotic tissue or benign proliferative changes, while necrotic tissue in a cancer is associated with increased ADC. Problems remain in selecting optimal cutpoints and recording parameters for DWI. Thus, breast cancer is not detected via DWI with full certainty [8].

Improved identification of breast cancer can potentially be achieved via molecular imaging through MRS or MR spectroscopic imaging (MRSI), whereby the metabolic characteristics of the examined tissue are assessed. When a single MRS voxel is inadequately representative, MRSI provides volumetric coverage. A recent meta-analysis of in vivo MRS and MRSI studies using 1.5T or 3T scanners reveals a pooled sensitivity of 73% and specificity of 88% for detecting over 700 breast cancers versus over 400 scans of non-cancerous breast [9]. These analyses were based on the total choline (tCho) peak, whose resonant frequency is ~3.2 parts per million (ppm). As an indicator of cell membrane turnover, tCho was used as a sign of cancer. Since signal-to-noise ratio (SNR) increases linearly with magnetic field strength, higher field scanners would be expected to improve SNR, to enhance MRS-based breast diagnostics. However, the diagnostic accuracy of MRS was not better with 3T compared to 1.5T scanners [9]. Using higher field (4T or 7T) MR scanners [10-12], tCho was detected in normal breast and benign lesions and was undetectable in several cancers. Overlap was seen among the ranges of tCho for these three categories of breast tissue. The high costs would also preclude high-field scanners from most breast cancer screening protocols. All told, tCho as a single composite spectral entity has not yielded adequate diagnostic accuracy for breast cancer screening. Further difficulties arise due to the various cutpoints of tCho used to define breast cancer.

1.2. Inadequacies due to reliance upon conventional data processing in MRS and MRSI

Whereas isolated resonances are better resolved with higher field scanners, widths are broadened when MRS time signals are encoded with stronger magnets. This adversely impacts on overlapping resonances that are abundant in MR spectra of the breast. Choline detection can be hampered with higher field magnets due to lipid-induced sidebands [13]. These lipids can be suppressed by prolonging the echo time (TE). However, potentially informative metabolites with short $T_2$ relaxation time will also have decayed with long TE. Furthermore, the lipid itself may be part of the actual disease process, such that diagnostic information could be lost with this encoding strategy [14].

Nuclear magnetic resonance performed in vitro on extracted breast specimens reveals more abundant metabolic information [15]. Components of tCho: phosphocholine (PC) resonating at ~ 3.22 ppm, glycerophosphocholine (GPC) at 3.23 ppm and free choline (Cho) at 3.21 ppm can better characterize breast tissue [16]. With malignant transformation in the breast, PC is increased in association with a “glycerophosphocholine to phosphocholine switch” due to the over-expressed enzyme choline kinase responsible for PC synthesis [16-18]. Besides the resonant frequencies of PC,
GPC and free choline being extremely close to one another, the far more abundant phosphoethanolamine (PE) also resonating at ~ 3.22 ppm, and which is not a cancer biomarker, obscures the PC peak on total shape spectra.

Encoded MRS time signals require mathematical transformation into their spectral representation in the frequency domain. Currently, this is done by the fast Fourier transform (FFT), built into all clinical MR scanners. The FFT is a low-resolution, linear, non-parametric processor. As a single polynomial, the FFT passes all the noise from the time domain into the frequency domain, with no way to suppress this noise corruption. The FFT can only produce a total shape spectrum; the underlying components cannot be generated via Fourier analysis. Thus, the overlapping resonances contained within the tCho peak cannot be disentangled by the FFT. No trustworthy information can be gleaned regarding the abundance of phosphocholine, PC, nor about the concentration of PE which, as mentioned, is completely superimposed on PC. Via the FFT, not even the number of overlapping resonances contained within a given peak can be determined. Post-processing via fitting is employed to surmise how many components underlie a given peak. Consequently, it is likely that false information will be generated by overfitting (overmodelling) and/or that physical resonances are missed because of underfitting (undermodelling). Fitting procedures cannot generate reliable quantitative information even about predetermined resonances, due to the likelihood of bias. Only by serendipity will a correct guess be made [19, 20]. It has therefore been concluded that the inability to autonomously reconstruct the spectral parameters, i.e. the fundamental complex frequencies and amplitudes, completely eliminates the FFT as a signal processor for solving the quantification problem, the “raison d’être” of MRS [19, 20].

1.3. The fast Padé transform: An advanced signal processor well-suited to MRS and MRSI

The fast Padé transform (FPT) is a high-resolution, non-linear signal processor with parametric and non-parametric forms. It is well-suited to MRS and MRSI [19-22], and can solve the above-described difficulties in breast cancer diagnostics. The number of genuine resonances contained in an encoded MRS time signal is handled as a parameter by the FPT, and is ascertained exactly. The complex-valued fundamental frequencies and amplitudes are precisely reconstructed for each resonance. With these parametric data, the key clinical information, the metabolite concentrations are accurately computed, even when resonances closely overlap [19, 21-22]. The spectrum generated by the FPT is the quotient of two polynomials \( P/Q \). There are two equivalent variants: FPT\(^{(+)}\) and FPT\(^{(-)}\) inside and outside the unit circle for the causal and anti-causal representation, respectively [19-22]. These two variants provide a fully self-contained checking procedure. The FPT algorithmically separates the genuine and spurious resonances by identifying pole-zero confluences. This is vital when there are many spurious resonances as in most MRS quantification problems, since over-determined systems of linear equations are encountered. The FPT-generated spectrum as the quotient of two polynomials \( P/Q \) produces a unique set of genuine spectral poles and zeros. The denominator polynomial \( Q \) is associated with the system poles and represents the positions (chemical shifts) and widths of peaks in a spectrum. The numerator polynomial \( P \) is associated with the valleys between adjacent peaks, and the system zeros are described thereby. Coincident poles and zeros, termed Froissart doublets [23], are unstable, roaming stochastically in the complex frequency plane with the slightest perturbation, and they never converge. Consequently, these spurious resonances are akin to noise which must be identified and discarded from the final results of the analysis. The two polynomials \( P \) and \( Q \) provide two degrees of freedom which facilitate noise suppression, as follows. In the Padé-reconstructed spectrum \( P_{k'}/Q_{k'} \), the total number of resonances \( K' \) is larger than its genuine counterpart \( K \) in the unknown ratio \( P_k/Q_k \). The overestimation \( \Delta K = K' - K \) is equi-partitioned between \( P_{k'} \) and \( Q_{k'} \), such that \( P_{k'} \) and \( Q_{k'} \) will have an equal number of the same spurious resonances in \( P_{k'}/Q_{k'} \). This excess, spurious content cancels out in \( P_{k'}/Q_{k'} \). After this cancellation, only the true information \( P_k/Q_k \) remains in the reconstructed data. Further details are given Refs. [19, 23-28].
1.4. Applications of the FPT to MRS time signals from breast and rationale for the present study
We have applied the FPT to noise-free MRS time signals as encoded from extracted breast specimens from normal breast, fibroadenoma and breast cancer [15]. In all three cases, the FPT resolved and precisely quantified the nine resonances, including PC and PE that were almost completely overlapping [4, 19, 26]. Many spurious resonances were also produced; at convergence 750 resonances were retrieved, i.e. 741 spurious plus the 9 physical resonances. Although fewer than 2% of the resonances were genuine, the FPT delineated these two categories with certainty in all cases. The FPT was then applied to MRS time signals associated with breast cancer and benign breast pathology (fibroadenoma) with added noise (σ = 0.0289 RMS, where RMS is the root-mean-square of the noise-free time signal, RMS = (N−1∑n=0N−1|cn|2)1/2 and σ is the standard deviation). Convergence was achieved in both cases at Np = 1700, with accurate reconstruction of the spectral parameters for all 9 genuine metabolites. Also, 841 spurious resonances were generated and were recognized as Froissart doublets with pole-zero confluences and zero-valued amplitudes, together with marked instability in the face of changes in noise level or signal length [27, 28]. By this three-fold “signature” the non-physical content is distinguished from physical resonances [7, 23, 27, 28].

Based upon these results, we herein apply the FPT to noise-corrupted time signals from breast without pathology to provide normative data. Of particular interest is the capability of the FPT to resolve the resonances between 3.21 ppm and 3.23 ppm, where, as noted, elevated phosphocholine is a marker of malignant transformation [16, 18]. We assess the performance of the FPT in handling noise-corrupted MRS time signals in this controlled setting, as a key step towards clinical application of Padé-optimized MRS for breast cancer diagnostics.

2. Methods

2.1. The input data: an MRS time signal as encoded from normal, non-infiltrated breast
The time signal of the typical quantum-mechanical form was generated according to:

\[ c_n = \sum_{i=1}^{K} d_i e^{i\omega_i t} \quad \text{Im} (\omega_k) > 0 \quad (0 \leq n \leq N - 1), \quad (1) \]

where N is the total signal length. These input data are based upon encoded time signals from tissue samples of normal, non-infiltrated breast from twelve patients, as per Ref. [15]. Each \( c_n \) is a sum of \( K \) = 9 damped complex exponentials \( \exp (i\omega_n t) \) (\( 1 \leq k \leq 9 \) with complex amplitudes \( d_k \). Herein, \( \omega_k \) and \( d_k \) are the fundamental angular frequencies and amplitudes; \( \omega_k = 2\pi v_k \), where \( v_k \) is the linear frequency. The time signals of the form (1) were quantified via the FPT, as per Refs. [19-22].

Since \( \omega_k \) is complex-valued, such that for \( \text{Im}(\omega_k) > 0 \), as in (1), quantity \( c_n \) decreases exponentially over time \( n \tau (n = 0, 1, 2, \ldots, N-1) \). Thus, \( \exp (i\omega_n t) = \exp (-\eta \tau \lambda_k + i n \tau \mu_k) \), where \( \lambda_k = \text{Im} (\omega_k) \) and \( \mu_k = \text{Re} (\omega_k) \). The rhs of Eq. (1) is then a linear combination of \( K \) exponentially attenuated frequency-dependent cosinusoids and sinusoids each of which is premultiplied by the constant, real intensity \( |d_k| \):

\[ d_k e^{i\omega_k t} = |d_k| e^{i(\varphi_k + i\omega_k t)} = |d_k| \left[ e^{-\eta_0 \lambda_k t} \left[ \cos(n\mu_k t + \varphi_k) + i \sin(n\mu_k t + \varphi_k) \right] \right], \quad (2) \]

where \( \varphi_k \) is the phase of the complex amplitude \( d_k \).

Herein, for conciseness, we present explicitly only the reconstructions of the FPT\(^{-}\). The converged findings are identical to those obtained from the FPT\(^{+}\) [19-22]. The diagonal (\( M = K \)) and non-diagonal (\( M \neq K \)) complex-valued spectra in the FPT\(^{-}\) are:

\[ G_{M,K}(z^{-1}) = \frac{G_M(z^{-1})}{Q_K(z^{-1})} \quad P_M(z^{-1}) = \sum_{r=0}^{M} p_r z^{-r}, \quad Q_K(z^{-1}) = \sum_{s=0}^{K} q_s z^{-s}. \quad (3) \]
Therein, \( z^{-1} \) is the harmonic variable \( z^{-1} = e^{-\text{lo}x} \) and \( \{p_r^-, q_s^\pm\} \) are the expansion coefficients. The unique Padé spectrum \( G_{M,K}(z^{-1}) \) from Eq. (3) exists for the given (input) exact finite-rank Green’s function \( G^{-1}(z^{-1}) \):

\[
G(z^{-1}) = G_{M,K}(z^{-1}) + O(z^{-M-K-1}); \quad G(z^{-1}) = \sum_{n=0}^{N-1} c_n \, z^{-n}.
\]  

Here, the symbol \( O(z^{-M-K-1}) \) signifies the error which is a series in powers \( z^{-m}(M + K + 1 \leq m \leq \infty) \). Thus, the approximation \( G(z^{-1}) \approx G_{M,K}(z^{-1}) \) in the general FPT\(^{-1}\) includes exactly \( M + K \) signal points \( \{c_n\} \) from the input data \( G(z^{-1}) = \sum_{n=0}^{N-1} c_n \, z^{-n} \) with \( M + K \leq N - 1 \).

The MRS time signals from Ref. [15] of length \( N = 65536 \) were recorded at a Larmor frequency \( (\nu_l) \) of 600 MHz and with a static magnetic field strength \( B_0 \approx 14.1 \text{T} \). A bandwidth (BW) of 6 MHz was used, where the inverse of this BW is the sampling time \( \tau \). The nine resonances identified in Ref. [15] were grouped into two frequency bands. The first was from 1.3 ppm to 1.5 ppm and the second from 3.2 ppm to 3.3 ppm. Within the second band were seven resonances, including two which very closely overlapped at \( \approx 3.22 \text{ ppm} \): phosphocholine, PC resonance \#4 and phosphoethanolamine, PE resonance \#5 separated by 0.001 ppm. In the present study, we used a 32-fold shorter time signal of total length \( N = 2048 \). The amplitudes \( d_i \) were computed based upon the reported concentrations of metabolites \( C_{\text{met}} \) for normal non-infiltrated breast tissue from twelve patients [15], with \( |d_i| = 2C_{\text{met}}/C_{\text{ref}} \), where \( C_{\text{ref}} = 0.05 \text{ mM/g} \). The internal reference used in Ref. [15] was TSP (3-(trimethylsilyl)-3,3,2,2-tetradeutero-propionic acid), such that \( |d_i| = C_{\text{met}}/(25 \mu \text{M/g}) \) of wet weight (ww) of the tissue. In Ref. [15], the \( T_2^* \) relaxation times were not reported. We set the line widths (full-widths at half-maximum, FWHMax) to be from 0.0008 to 0.0009 ppm. The peaks are represented by Lorentzians, consistent with the time signal from Eq. (1). In the absorption mode, we have used mainly the diagonal, \( \text{Re}(p_{K}^+/Q_{K}^+) \) and paradiagonal, \( \text{Re}(p_{K-1}^-/Q_{K}^-) \) forms, since these two variants for \( t = 0 \) and \( l = 1 \) are empirically observed to be more stable compared to those using other values of \( l( \geq 2, 3, \ldots) \) in \( p_{K-l}^+/Q_{K-l}^+ \). The line widths are proportional to \( \text{Im}(\nu_l) \). The smallest chemical shift difference of 0.001 ppm is only 1.11 to 1.25 times larger than the line widths of \( 8 \times 10^{-4} \) to \( 9 \times 10^{-4} \) ppm. The phases \( \phi_t(l \leq k \leq 9) \) from complex-valued \( d_i \) were all set to zero. Thus, all the \( d_i \)’s are real, \( d_i = |d_i| \). The input data are displayed in the top panel (1) of table 1.

Through the parametric version of the diagonal FPT\(^{-1}\) to analyze the MRS time signal data from normal, non-infiltrated breast tissue, the coefficients \( \{p_r^-, q_s^-\} \) of the polynomials \( P_{K}^-(z^{-1}) \) and \( Q_{K}^-(z^{-1}) \) from Eq. (3) were computed from the defining product \( Q_{K}^-(z^{-1})G(z^{-1}) \approx P_{K}^-(z^{-1}) \). When herein the coefficients of the same powers of \( z^{-1} \) are equated, two systems of linear equations are generated, one for \( \{q_s^-\} \) and the other for \( \{p_r^-\} \):

\[
\sum_{j=0}^{K} c_{s-j} \, q_j^- = 0 \quad (0 \leq s \leq K, q_0^- = 1), \quad p_r^- = \sum_{j=0}^{r} c_{r-j} \, q_j^- \quad (0 \leq r \leq K).
\]  

After the set \( \{q_s^-\} \) is extracted from the input data \( \{c_n\} \) \( (0 \leq n \leq N-1) \), the other set \( \{p_r^-\} \) becomes immediately available via the analytical expression \( p_r^- = \sum_{j=0}^{r} c_{r-j} \, q_j^- \) \( (0 \leq r \leq K) \) from (5). Thus, the set \( \{p_r^-, q_s^-\} \) is obtained by solving only one system of linear equations (that for \( \{q_s^-\} \)). Pole-zero coincidences come from correlation between the numerator \( (P_{K}) \) and the denominator \( (Q_{K}) \) polynomials in the Padé quotient \( P_{K}/Q_{K} \). This correlation is evidenced by Eq. (5) where the expansion coefficients \( \{p_r^-\} \) of \( P_{K} \) are derived from the expansion coefficients \( \{q_s^-\} \) of \( Q_{K} \). To extract the peak parameters, the characteristic equation \( Q_{K}^-(z^{-1}) = 0 \) is solved, leading to \( K \) unique roots \( z_{K,0}^{-1} = e^{-i\nu_{K,0}}(1 \leq k \leq K) \).
Thus, the sought eigen-frequency \( \omega_{k,Q} \) is deduced via \( \omega_{k,Q} = (i/\tau) \ln(z_{k,Q}^{-1}) \). Hereafter, \( z_{k,Q}^{-1} \) and \( z_{k,Q}^{-1} \) denote the zeros of \( P_K(z^{-1}) \) and \( Q_K(z^{-1}) \), respectively. The zeros \( z_{k,Q}^{-1} \) are required for quantification. In addition, for signal-noise separation via Froissart doublets the zeros \( z_{k,P}^{-1} \) are also needed. Through the FPT\(^{-1}\), the parameters \( \{ \omega_{k,Q}', d_k' \} (1 \leq k \leq K) \) of each physical resonance are obtained directly from the examined MRS time signal. The \( k \)th metabolite concentration is computed from the reconstructed amplitudes \( d_k' \) via \( C_{\text{met}} = (25 \mu M/g) |d_k'| \), as mentioned.

2.2. Addition of noise corruption to the input data
We performed the reconstructions using the input data corrupted by noise. The latter were generated by adding complex-valued random zero-mean Gauss-distributed white noise of preassigned levels to the noiseless MRS time signal. The noise levels were \( \sigma = 0.00289 \) RMS and \( \sigma = 0.0289 \) RMS with, as noted, RMS being the root-mean-square of the noise-free time signal.

2.3. Comparisons of the FPT and the FFT in non-parametric signal processing
It is only for total shape spectra that the Padé and Fourier spectra can directly be compared, because as a non-parametric estimator, that is all the FFT can compute. The Fourier spectrum is given by:

\[
F_k = \frac{1}{N} \sum_{n=0}^{N-1} c_n \exp(-2i\pi kn/N) \quad 1 \leq k \leq N-1, \quad (6)
\]

where \( 2\pi k/T \) are fixed Fourier grid frequencies that are unrelated to \( \omega_k \) from \( c_n \). Here, \( T \) is the total duration or acquisition time of the signal, \( T = N\tau \). The variable \( \exp(-2i\pi kn/N) \) in (6) is an undamped harmonic. The non-parametric data are provided by the FPT when the Padé polynomials \( P_K \) and \( Q_K \) are extracted from the time signal \( \{c_n\} \). The ratio \( P_K/Q_K \) is the complex-valued total shape spectrum whose real part is the absorption spectrum. The FPT generates the total shape spectrum \( P_K/Q_K \) at any given set of sweep frequencies that need not correspond to the preassigned Fourier grid [19, 20]. Thus, the FPT has interpolation capabilities relating directly to the features of the analyzed time signal.

2.4. Quantification of MRS time signals from normal, non-infiltrated breast by the parametric FPT
The MRS data are quantified by the FPT via polynomial rooting, whereby the roots of the characteristic equations of the numerator \( (P_K) \) and the denominator \( (Q_K) \) polynomials generate the zeros and poles of the Padé spectrum \( P_K/Q_K \), respectively. Since the rational function \( P_K/Q_K \) is meromorphic, the zeros of \( Q_K \) are the poles of \( P_K/Q_K \). Meromorphic functions have poles as the only singularities. Roots \( z_{k,Q}^{-1} \) of equation \( Q_K(z^{-1}) = 0 \) reconstruct the fundamental or eigen-frequencies \( \omega_{k,Q} \). The amplitude \( d_k' \) is given via the analytical expression for the Cauchy residue of \( P_K(z^{-1})/Q_K(z^{-1}) \) taken at the \( k \)th pole \( z^{-1} = z_{k,Q}^{-1} \). Although the quantification problem entails finding the eigen-set \( \{ \omega_{k,Q}, d_k' \} (1 \leq k \leq K) \), it is not the concluding part of spectral analysis, since this does not generate the eigenset \( \{ \omega_{k,Q}, d_k' \} (1 \leq k \leq K) \) with \( K \) being the true number of resonances. Rather, because an over-determined system of linear equations is set up, a larger collection is produced \( \{ \omega_{k,Q}', d_k' \} (1 \leq k \leq K') \), where \( K' \) is the sum of \( K \) genuine and \( K'' \) spurious resonances \( (K' = K + K'') \). Thus, we obtain \( P_{K'}/Q_{K'} \) instead of \( P_K/Q_K \). To omit the \( K'' \) false resonances that are not the constituents of the \( K \) input eigen-frequencies and the \( K \) associated eigen-amplitudes, we root the numerator polynomial via \( P_{K'}(z^{-1}) = 0 \) to obtain the zeros \( \{ z_{k,K'}^{-1} \} (1 \leq k \leq K') \) of the spectrum \( P_{K'}(z^{-1})/Q_{K'}(z^{-1}) \). A subsequent binning is carried out according to stable versus unstable resonances as a function of the varying degree \( K' \) of polynomials \( P_{K'} \) and \( Q_{K'} \). Finally, the exact number \( K \) of genuine resonances is reconstructed from \( K = K' - K'' \) after discarding the \( K'' \) unstable resonances as spurious.
3. Results
We systematically augmented the signal length for the same bandwidth (i.e. 3 acquisition times) to check the constancy of the spectral parameters. The reconstructed spectral parameters and metabolite concentrations are shown in panels (2) to (4) of table 1 at total orders \( K = 350, 850 \) and 1020. The partial signal length \( N_p \) is given by \( N_p = 2K \) for total signal length \( N = 2048 \). In panel (2) of table 1, at \( N_p = 700 \), only 8 of the 9 genuine resonances were resolved. Between 3.220 ppm and 3.221 ppm only 1 resonance was generated at 3.221 ppm, whose computed concentration is the sum of PC + PE (“4 + 5”). At \( N_p = 1700 \) convergence was attained for all the resonances, such that both PC and PE are fully resolved with their correct amplitudes and computed concentrations (panel (3)). At all higher \( N_p \) and at full signal length \( N \) convergence remained stable, as illustrated in the bottom panel (4) of table 1 at \( N_p = 2040 \).

Figure 1 summarizes key clinical MRS data for normal, non-infiltrated breast tissue. In panel (1) the input data, the real part (Re) of the complex time signal is shown, as per the encoded data from normal non-infiltrated breast reported in Ref. [15]. To the noiseless time signal, we added Gaussian-distributed noise with \( \sigma = 0.0289 \) RMS. The imaginary part (Im) of the time signal is similar (omitted). The output data are shown either for the interval between 3.2 ppm and 3.3 ppm or [3.16, 3.34] ppm. These are presented as spectra in panels (2), (3), (4) and as metabolite concentrations in panel (5). The absorption total shape spectrum reconstructed via the FFT in panel (2) at full signal length \( (N = 2048) \) is totally inaccurate and non-convergent, with only a few rough peaks, while the FPT has converged at \( N_p = 1700 \) on panel (3) (total shape spectrum) and panel (4) (component shape spectra) where PC is seen to completely underlie PE. The converged Padé-generated concentration map at \( N_p = 1700 \) is presented on panel (5), wherein the concentrations of all the metabolites, including PC are accurately computed. Signal-noise separation is presented in panels (6) and (7) where FWHMax (full width at half maximum: spectral poles) equals FWHMin (full width at half minimum: spectral zeroes) for non-physical data symbolized as \( \bullet \) (Froissart doublets). These show marked instability at the two noise levels: panel (6) \( (\sigma = 0.00289 \) RMS) and panel (7) \( (\sigma = 0.0289 \) RMS). The zero peak heights in (8) (the abscissa crosses the centers of the empty circles) are the 3rd “signature” of spurious resonances. In contradistinction, for physical resonances \( \text{FWHMax} \neq \text{FWHMin} \). For such results, the reconstructed and input poles (O and \( \times \) ) coincide, as indicated by \( \bigcirc \). These are stable against the two noise levels differing by a factor of 10 in panels (6) and (7). Also, the related peak heights are non-zero (panel (8)). Although such true structures may show pole-zero near-coincidences and some of the peak heights (GPC, PC and CHO) are close to zero, they show no change for different noise levels, as seen in the dash-lined box in panels (6) and (7).

4. Discussion and Conclusions
The algorithmic success of the FPT is a direct result of its quantum mechanical underpinnings, whereby the frequency spectrum is represented as the unique ratio of two polynomials as in the finite-rank Green’s function [19, 20]. Its superior resolution compared to the FFT demonstrated herein and elsewhere [19, 20, 22, 27, 28] has several reasons. One is the interpolation capability of the FPT. This occurs because a spectrum in the FPT\(^{(-)}\), given by a polynomial quotient \( P^K (z^{-1})/Q^K (z^{-1}) \) where \( z^{-1} = e^{-\omega t} \), can be taken at any sweep frequency \( \omega \). Consequently, frequency \( \omega \) can be selected independently of the employed number of signal points \( \{c_n\} \). Thus, the FPT achieves interpolation in spectra, based on the actual analyzed time signal, whereas the minimal separation \( \omega_{\text{min}} \) in an FFT spectrum is determined by the total acquisition time \( T \). The FFT spectrum is defined only on the Fourier grid points \( 2\pi k/T \) \( \{k = 0,1,2,3,\ldots,N - 1\} \), where \( N \) is the total signal length. Thus, the FFT cannot interpolate. The superfluous data (zeros) added in the encoding domain do not improve resolution in the Fourier spectrum. Further, the FPT can extrapolate beyond the \( N \) input signal points because the implicitly present inverse of \( Q^K \) is an infinite sum (Maclaurin series). The FPT has an extra degree of freedom to cancel out noise from the numerator \( P^K \) and the denominator \( Q^K \), whereas
CONVERGENCE of QUANTIFICATION in FPT\(^{(-)}\); PARTIAL SIGNAL LENGTHS: \(N_P = 700, 1700, 2040\)

(1) Input Data: Spectral Parameters, Concentrations and Metabolite Assignments

| \(N_k\) or \(k\) | \(\text{Re}(\nu_k)\) (ppm) | \(\text{Im}(\nu_k)\) (ppm) | \(|d_k|\) (au) | \(C_k\) (\(\mu\)M/g) | \(M_k\) |
|-----------------|-----------------|-----------------|----------|-----------------|------|
| 1               | 1.332           | 0.0008          | 0.020    | 0.500           | LAC  |
| 2               | 1.471           | 0.0008          | 0.004    | 0.100           | ALA  |
| 3               | 3.212           | 0.0008          | 0.001    | 0.025           | CHO  |
| 4               | 3.220           | 0.0008          | 0.001    | 0.025           | PC   |
| 5               | 3.221           | 0.0009          | 0.007    | 0.175           | PE   |
| 6               | 3.232           | 0.0008          | 0.001    | 0.025           | GPC  |
| 7               | 3.251           | 0.0008          | 0.018    | 0.450           | \(\beta\)-GLC |
| 8               | 3.273           | 0.0008          | 0.006    | 0.150           | TAU  |
| 9               | 3.281           | 0.0008          | 0.011    | 0.275           | M-INS|

(2) Pade–Reconstructed Data: \(N_p = 700\), \(N = 2048\) (Partial Convergence)

| \(N_k\) or \(k\) | \(\text{Re}(\nu_k)\) (ppm) | \(\text{Im}(\nu_k)\) (ppm) | \(|d_k|\) (au) | \(C_k\) (\(\mu\)M/g) | \(M_k\) |
|-----------------|-----------------|-----------------|----------|-----------------|------|
| 1               | 1.332           | 0.0008          | 0.020    | 0.500           | LAC  |
| 2               | 1.471           | 0.0008          | 0.004    | 0.100           | ALA  |
| 3               | 3.212           | 0.0008          | 0.001    | 0.025           | CHO  |
| 4               | 3.221           | 0.0009          | 0.008    | 0.200           | PC+PE|
| 6               | 3.232           | 0.0008          | 0.001    | 0.025           | GPC  |
| 7               | 3.251           | 0.0008          | 0.018    | 0.450           | \(\beta\)-GLC |
| 8               | 3.273           | 0.0008          | 0.006    | 0.150           | TAU  |
| 9               | 3.281           | 0.0008          | 0.011    | 0.275           | M-INS|

(3) Pade–Reconstructed Data: \(N_p = 1700\), \(N = 2048\) (Full Convergence)

| \(N_k\) or \(k\) | \(\text{Re}(\nu_k)\) (ppm) | \(\text{Im}(\nu_k)\) (ppm) | \(|d_k|\) (au) | \(C_k\) (\(\mu\)M/g) | \(M_k\) |
|-----------------|-----------------|-----------------|----------|-----------------|------|
| 1               | 1.332           | 0.0008          | 0.020    | 0.500           | LAC  |
| 2               | 1.471           | 0.0008          | 0.004    | 0.100           | ALA  |
| 3               | 3.212           | 0.0008          | 0.001    | 0.025           | CHO  |
| 4               | 3.221           | 0.0009          | 0.007    | 0.175           | PE   |
| 6               | 3.232           | 0.0008          | 1.001    | 0.025           | GPC  |
| 7               | 3.251           | 0.0008          | 0.018    | 0.450           | \(\beta\)-GLC |
| 8               | 3.273           | 0.0008          | 0.006    | 0.150           | TAU  |
| 9               | 3.281           | 0.0008          | 0.011    | 0.275           | M-INS|

(4) Pade–Reconstructed Data: \(N_p = 2040\), \(N = 2048\) (Full Convergence)

| \(N_k\) or \(k\) | \(\text{Re}(\nu_k)\) (ppm) | \(\text{Im}(\nu_k)\) (ppm) | \(|d_k|\) (au) | \(C_k\) (\(\mu\)M/g) | \(M_k\) |
|-----------------|-----------------|-----------------|----------|-----------------|------|
| 1               | 1.332           | 0.0008          | 0.020    | 0.500           | LAC  |
| 2               | 1.471           | 0.0008          | 0.004    | 0.100           | ALA  |
| 3               | 3.212           | 0.0008          | 0.001    | 0.025           | CHO  |
| 4               | 3.221           | 0.0009          | 0.007    | 0.175           | PE   |
| 6               | 3.232           | 0.0008          | 0.001    | 0.025           | GPC  |
| 7               | 3.251           | 0.0008          | 0.018    | 0.450           | \(\beta\)-GLC |
| 8               | 3.273           | 0.0008          | 0.006    | 0.150           | TAU  |
| 9               | 3.281           | 0.0008          | 0.011    | 0.275           | M-INS|

Table 1: Quantification of MRS time signals typical of normal breast [15]. Shown are spectral parameters, concentrations and assignments of metabolites. The real and imaginary frequencies are in the 2nd and 3rd column. The amplitudes, concentrations and metabolite assignments are in the 4th, 5th and 6th column. The input data are in panel (1). The reconstructions by the fast Padé transform, FPT\(^{(-)}\), are in panels (2), (3) and (4) at partial signal lengths \(N_P = 700, 1700\) and 2040. The total signal length is \(N = 2048\). Full convergence by the FPT\(^{(-)}\) is seen in panels (3) and (4) at \(N_p = 1700\) and 2040. For more details, see the main text.
MRS for NORMAL BREAST by FAST PADE TRANSFORM, FPT\(^{-1}\); panels (3)–(8), FAST FOURIER TRANSFORM, FFT: panel (2)

Figure 1: (Color online) The fast Padé transform, FPT\(^{-1}\), applied to MRS time signals typical of normal breast \([15]\). Input data: time signal in panel (1). Output data: absorption spectra in panels (2), (3) and (4); concentrations of metabolites in panel (5); signal-noise separation via Froissart doublets in panels (6), (7) and (8). Symbols: input data (\(\times\)) and output data (\(O\), \(\bullet\)). To avoid clutter in panels (6) and (7), the full width at the half maximum, FWHM, is doubled at frequency 3.220 ppm (as PC\(^*\)). For more details, see the main text.
in the FFT, as a single polynomial and being linear, noise is imported intact from the measured time
domain to the computed frequency spectrum [19, 20]. Due to the superior resolution performance of
the FPT, fewer and shorter MRS time signals could be encoded, thereby diminishing examination
time. This translates into greater comfort for the patient and lower cost.

For normal, non-infiltrated breast, the concentration of PC was ~15% that of the overlying PE,
whereas for the other two cases (fibroadenoma and breast cancer), the relative concentrations were
~30% [15]. Moreover, the PC concentration was less than 20% of that for fibroadenoma [19, 26].
Nevertheless, the PC concentration, albeit very low, was computed precisely by the FPT in the
presence of moderate levels of noise. This genuine metabolite was clearly distinguished from the
numerous spurious resonances by its stability in the face of different noise levels, as well as at
different partial signal lengths after convergence. All the other genuine metabolites were also
confidently identified, with their concentrations precisely computed by the FPT, even though several
(notably the other components of total choline) were likewise only sparsely present. Precisely the
capabilities of the FPT demonstrated here help pave the way for developing urgently-needed
normative data bases, to which comparisons could be confidently made with various benign breast
pathologies, as well as with breast cancer of various types.

The results presented herein provide further evidence that mathematical optimization through the
FPT will be of vital importance for realizing the potential of magnetic resonance spectroscopy, MRS,
and spectroscopic imaging, MRSI, for early cancer diagnostics, in particular for breast cancer. As
long as MRS and MRSI rely upon the FFT, plus post-processing via fitting, breast cancer cannot be
distinguished from normal, non-infiltrated breast nor from benign breast lesions with sufficient
confidence. Moreover, the low resolution of the FFT renders MRS and MRSI of the breast extremely
time-consuming and therefore uncomfortable for the patient and costly. Encoding for MRS entails
poor SNR as a major problem. Exclusive reliance upon the FFT requires encoding many transients
(~ 200) to be averaged for improving SNR. This keeps the patient longer in the scanner, resulting in
increased inhomogeneity of the static magnetic field. During this extended encoding time, the patient
moves, so that more and more transients are likely not to be associated with the same intended voxel.
This leads to incoherent averaging, when adding together these different time signals. Hence, this
illustrates the need to reformulate the entire concept of encoding in MRS by relying upon another
theory which requires fewer encoded transients (~ 20) and yet improves SNR and resolution. The FPT
is optimally poised to serve this purpose critical to the use of MRS in daily clinical practice.

The closely-overlapping resonances, notably the components of total choline, within the MR
spectra of the breast appear to be of crucial diagnostic importance for early cancer diagnostics. The
FFT plus non-unique fitting procedures cannot provide any certainty about these overlapping
resonances. Thus, even at high magnetic field strengths, proton MRS has thus far only generated estimates of total choline. With these, only incremental improvement are seen compared to anatomical
imaging with MRI, in the accuracy with which breast cancer is distinguished from non-malignant
breast tissue.

The high-resolution, quantification-equipped FPT represents a paradigm shift for signal processing.
This has been confirmed in interdisciplinary research fields and applications, where generic time
signals comprised of complex damped harmonics are routinely encountered. These capacities of the
FPT are also shown herein. Namely, in the presence of realistic noise levels, the FPT clearly
distinguishes spurious resonances from genuine metabolites and exactly computes the concentrations
of the latter, including phosphocholine, PC, albeit only minimally present for normal, non-infiltrated
breast. The high resolution of the FPT together with its enhancement of SNR yields shorter
examination time, thus increasing patient comfort plus lowering costs. Particularly with closely-
overlapping resonances in the MR spectra, convergence is attained by a greatly over-determined system of linear equations. The signal-noise separation procedure uniquely linked to the FPT is the
means by which genuine i.e. physical resonances are identified and the metabolite concentrations
precisely computed. Simultaneously, the same procedure identifies the much more numerous spurious
resonances via coincident poles and zeros (Froissart doublets), accompanied by zero or near zero amplitudes and marked instability against various levels of truncation and/or noise. The results presented in this paper in which normative data for the breast are generated by the FPT in the presence of noise, bring us one step closer to applying the fast Padé transform to time signals encoded in vivo from normal, benign and cancerous breast in order for MRS and MRSI to realize their potential to, at last, become a reliable, cost-effective method for breast cancer diagnostics.

Acknowledgements
This work was supported by Radiumhemmets Research Fund, Cancerfonden and FoUU.

References
[1] Eccles S, Aboagye E, Ali S, Anderson A, Armes J and Berditchevski F 2013 Breast Cancer Res. 15 R92
[2] Sant M et al 2003 Int. J. Cancer 106 416
[3] Saslow D et al 2007 CA. Cancer J. Clin. 57 75
[4] Belkić K and Belkić Dž 2011 Isr. Med. Assoc. J. 13 236
[5] Sardanelli F et al 2011 Invest. Radiol. 46, 94
[6] Feig S 2010 Radial. Clin. N. Am. 48 879
[7] Belkić Dž and Belkić K 2013 Technol. Cancer Res. Treat. DOI: 10.7785/tcrt.2012.500401
[8] Malayeri A et al 2011 Radiology 267 735
[9] Balzter P and Dietzel M 2013 Radiology 267 735
[10] Haddadin I et al 2009 NMR Biomed. 22 65
[11] Bolan P et al 2003 Magn. Reson. Med. 50 1134
[12] Klomp D et al 2011 NMR Biomed. 24 1337
[13] Bolan P et al 2002 Magn. Reson. Med. 48 215
[14] Stanwell P and Mountford C 2007 Radiographics 27 S253
[15] Gribbestad I, Sitter B, Lundgren S, Krane J and Axelsson D 1999 Anticancer Res. 19 1737
[16] Katz-Brull R, Seger D, Rivenson-Segal D, Rushkin E and Degani H 2002 Cancer Res. 62 1966
[17] Eliyahu G, Kreizman T and Degani H 2007 Int. J. Cancer 120, 1721
[18] Glunde K, Jiang J, Moestue S and Gribbestad I 2011 NMR Biomed. 24, 673
[19] Belkić Dž and Belkić K 2010 Signal Processing in Magnetic Resonance Spectroscopy with Biomedical Applications (London, Taylor & Francis)
[20] Belkić Dž 2005 Quantum Mechanical Signal Processing and Spectral Analysis (Bristol, Institute of Physics Publishing)
[21] Belkić Dž 2006 Phys. Med. Biol. 51 2633
[22] Belkić Dž 2006 Phys. Med. Biol. 51 6483
[23] Belkić Dž 2009 Adv. Quantum Chem. 56 95
[24] Ojo A 2010 The Analysis and Automatic Classification of Nuclear Magnetic Resonance Signals. PhD Thesis, The University of Edinburgh Research Archive (http://hdl.handle.net/1842/4109)
[25] Zhang J 2013 Human Brain Glutamate, Glutamine, γ-aminobutyric Acid: Proton Magnetic Resonance Spectral Quantification with The Fast Padé Transform. PhD Thesis University of California Los Angeles (escholarship.org/uc/item/1np12339)
[26] Belkić Dž and Belkić K 2009 J. Math. Chem. 45 790
[27] Belkić Dž and Belkić K 2014 J. Math. Chem. 52 2680
[28] Belkić Dž and Belkić K 2014 Technol. Cancer Res. Treat. DOI: 10.1177/1533034614547446