Bioimpedance Spectroscopy Processing and Applications

Hershel Caytak[1], Alistair Boyle[2], Andy Adler[3], and Miodrag Bolic[4]

[1]School of Electrical Engineering and Computer Science, 800 King Edward Avenue, University of Ottawa, Ottawa, K1N6N5, Canada. Email: hershel.caytak@gmail.com
[2]School of Electrical Engineering and Computer Science, 800 King Edward Avenue, University of Ottawa, Ottawa, K1N6N5, Canada. Email: aboyle2@uottawa.ca
[3]Systems and Computer Engineering, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6, Canada. Email: adler@sce.carleton.ca
[4]School of Electrical Engineering and Computer Science, 800 King Edward Avenue, University of Ottawa, Ottawa, K1N6N5, Canada. Email: mbolic@site.uottawa.ca

Bioimpedance Spectroscopy (BIS) uses multifrequency impedance measurements of biological tissues to estimate clinically and experimentally relevant parameters. This chapter reviews the steps involved in measurement, data processing, and applications of BIS data, with an emphasis on managing data quality and sources of errors. Based on a description of of error sources, caused by measurement configuration, hardware, and modelling, we describe BIS data denoising. Two classes of modeling, explanatory and descriptive, can be used to reduce data dimensionality to a set of parameters or features. Explanatory models consider the electrical properties of samples and involve fitting data to simplified equivalent electrical circuits. Descriptive models involve reduction of the data to a set of eigenvectors/values which can be studied independently of any assumed electrical characteristics of the sample. Techniques described include fitting and decomposition methods for extraction of explanatory and descriptive model parameters, respectively. Denoising techniques discussed include adjusting measurement configuration, corrective algorithms for removal of artifacts and use of supervised machine learning for identification of features characteristic of noisy impedance spectra. The chapter concludes with a discussion of the use of classifiers for labeling BIS data in a range of applications including discrimination of healthy vs pathological tissues.

KEYWORDS

Artifacts, Bioimpedance Spectroscopy, Classifiers, Denoising, Measurement, Modeling, Multi-frequency, Processing

1 BIOIMPEDANCE SPECTROSCOPY OVERVIEW

Bioimpedance measures the voltage required to drive low-level alternating current through living tissue. Such measurements provide a safe, non-invasive, and relatively inexpensive method to measure tissue properties. Biological tissues typically have much more complex impedance spectra than homogeneous materials, arising from internal membranes and macromolecules. Bioimpedance is comprised of resistance to movement of charged particles and reactance which arises from changing electrical and magnetic field opposing movement of electrical charge. Bioimpedance spectroscopy (BIS) is the acquisition of bioimpedance measurements at multiple frequencies, and the measured BIS impedance spectra can help characterize the structural and chemical composition of biological tissue. Research and clinical applications of BIS include assessment of human body composition, cerebral monitoring and assist in diagnostic classification of various pathologies.

In many cases, BIS measurements are not considered a primary diagnostic tool due to high intra- and inter-patient measurement variability. In this work, sources of error are discussed (Section 2): noise originating from instrumentation design, experimental design (electrode placement, electrode-interface impedance), properties of biological tissue under test (electrical anisotropy, geometric irregularities) and inaccuracies of models used for data interpretation. Since raw BIS measurements are large, multi-variate data sets, the raw impedance spectra usually undergo further processing to calculate a few experimentally or clinically relevant parameters. These parameters can be used as features to classify tissues under test (e.g. pathological vs normal tissue) or may be subject to further processing including statistical analysis and monitoring of changes in the parameters over time (e.g. for hypoxia). BIS parameters are also used as inputs into specialized algorithms for applications such as body composition assessment and analysis of body fluid levels.

Processing steps, including denoising, parameterization, and classification are of key importance in ensuring that the raw measured impedance data is converted to relevant, statistically robust and clinically useful parameters.

Existing surveys on BIS generally focus on instrumentation, measurement design, modelling and applications, while the processing steps are often only briefly described. In this chapter, we aim to provide an overview and framework for processing raw impedance measurements. We also describe factors that affect processing and denoising methods, such as sources of noise, target application, assumed mathematical model and experimental configuration. The goal of this chapter is not to present a comprehensive review of the state of the art in BIS research and applications but rather to summarize the processing and denoising steps for BIS data.

1.1 BIS Processing Steps

The topics in this chapter are organized around major processes which form part of BIS signal analysis. The sequence of BIS data acquisition and processing steps are described as blocks as shown in Figure 1.1; relationships (inputs and outputs) are demonstrated by arrow lines connecting the various blocks. The BIS measurement configuration generally consists of use of two (bipolar) or four (tetrapolar) electrodes for injection of current and measurement of resulting voltage drop across a biological sample/tissue. Noise sources, experimental and instrumentation, affect measurement error and are thus shown as inputs to the measurement configuration block. Denoising strategies include adjustment of measurement configuration and removal of noise and artifacts from raw impedance data and model-extracted features. Impedance data are simplified to a set of parameters using models; the output of the models can fed to classifiers for labeling tasks or for identifying noise artifacts, monitored for changes over time or used as inputs of specialized algorithms to measure various physiological quantities and parameters. In the following sections we first discuss the experimental and instrumentation sources of noise and their relation to measurement configuration. Next we describe major processing steps including common forms of data modelling and associated denoising and processing steps (i.e. fitting methods). We also explain how features extracted from BIS models can be fed into classifiers for detection/discrimination based applications or used for further post processing.

2 Sources of Inaccuracy

Knowledge of significant sources of error is important in determining optimal measurement configuration, data processing steps, model selection, appropriate denoising strategies and for determination of accuracy of the clinical/diagnostic parameters derived from the raw impedance. Errors in the interpretation of estimated features may be induced by non-ideal hardware or by models that do not accurately represent the measurement configuration.

2.1 Non-ideal Hardware

Bioimpedance spectroscopy measurement hardware uses low frequency (Hz to MHz) analog circuits and corresponding sampling methods on strong (mV) common-mode signals (Figure 2.1). An idealized model of BIS hardware can be constructed from ideal amplifiers, digital-to-analog and analog-to-digital convertors, filters, wiring and electrodes. Real hardware introduces non-linear and non-ideal behaviour to each of these components which affect measurement quality. The challenges in implementing reliable hardware that can produce high quality measurements are addressed through appropriate selection of drive, connectivity, and measurement circuits.

The quality of the current drivers is determined by quantization errors, current source matching, output impedance, and frequency range. Connectivity to the body through electrodes and wiring limits system performance by electrode polarization, wiring crosstalk, leakage currents and stray capacitance. Differential measurements (μV) are limited by Common Mode Rejection Ratio (CMRR), filtering, analog-to-digital sampling dynamic range and

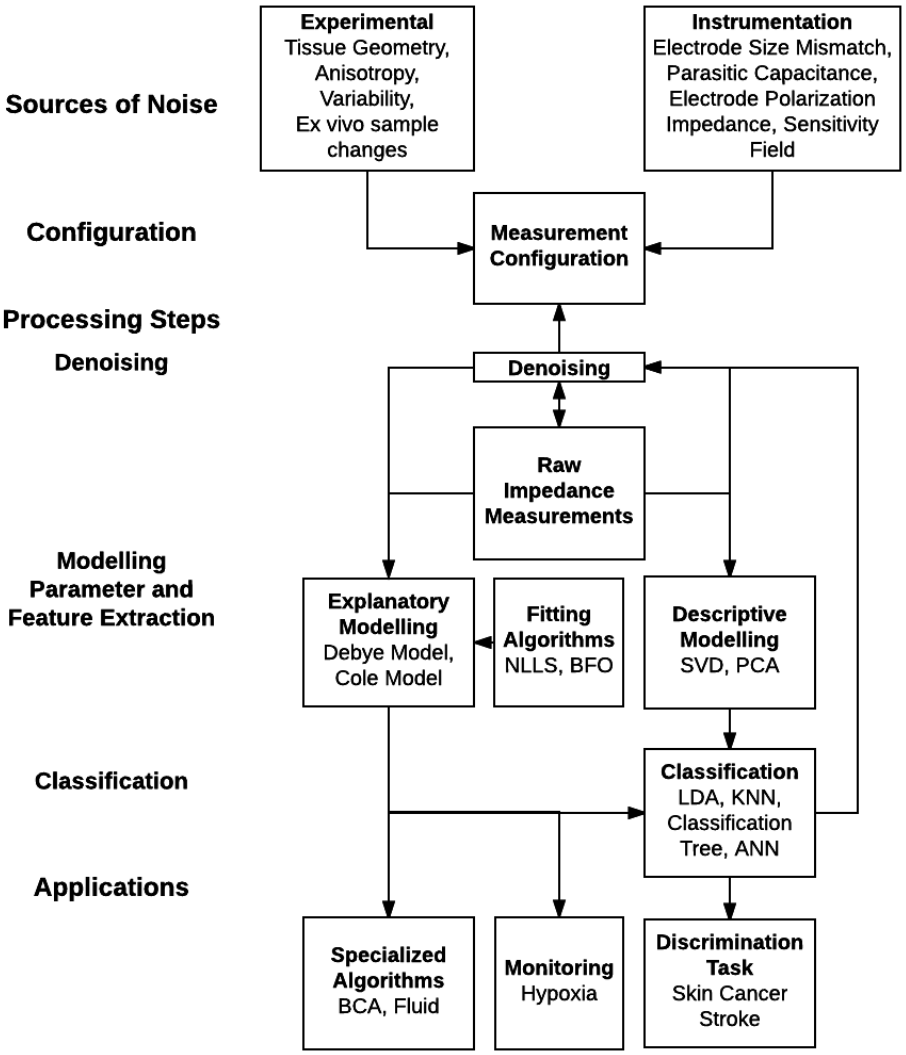

Figure 1.1: Block diagram of major steps of BIS data processing

quantization errors, time source jitter, and phase accuracy.

Many systematic errors may be calibrated out, but calibration cannot be universally accurate across all frequency and measurement configurations. High quality calibration is typically limited to well characterized and repeatable environments. Electrically noisy and relatively uncontrolled environments, such as operational hospital wards, can make it challenging to accurately calibrate an instrument. Other errors limit system performance as intrinsic characteristics of the circuit design choices necessary to achieve multi-frequency measurements. Bioimpedance spectroscopy is, therefore, particularly challenging due to the breadth of the frequencies across which calibration must achieve acceptable accuracy and repeatability.

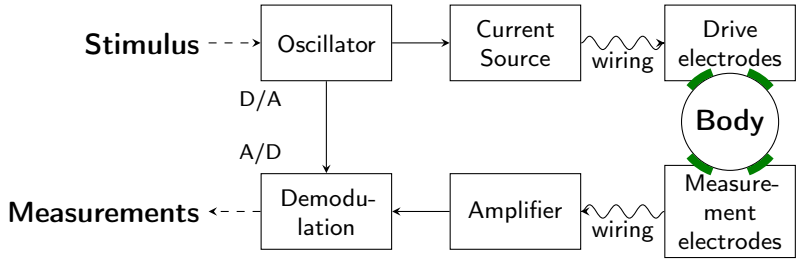

Figure 2.1: Block diagram of typical bioimpedance hardware; stimulus selects a frequency for an oscillator which controls current sources that drive current through a body via wiring and electrodes; voltage differences at the measurement electrodes are amplified and then demodulated to determine the magnitude and phase of measurements

The quality of high frequency measurements are generally limited by CMRR, which drops off beyond a certain circuit dependent frequency (10-100 kHz for off-the-shelf circuits).

The drive and measurement circuits are designed to support a particular range of impedance on the body. Supporting a wide impedance range (ohm to gigaohm) though configurable drive and measurement circuits comes at the expense of circuit complexity, exacerbating calibration challenges or introducing further sources of error in the measurements.

2.2 Experimental Variability

When non-ideal hardware is connected to biological materials, additional sources of measurement variance are introduced: contact impedance, mismatch between electrodes, inhomogeneous materials, electrode polarization and stray capacitance.

One major source of error is the high impedance at the electrode-skin contact boundary. This impedance, commonly noted as Z_{ep}, is frequency dependent and tends to result in high impedance values with large variance at low frequencies. The high variability of Z_{ep} may result in impedance mismatches that affect the measurement accuracy. Electrode impedance mismatch is defined as a large difference of Z_{ep} between any electrode pair. Electrode mismatch effect on impedance measurements was quantified through analysis of a set of tetrapolar (separate voltage and current electrode pairs) right ankle and wrist measurements on three subjects. The device (SFB7 spectrometer) was set to perform a frequency sweep from around 3 to 999 kHz. Electrode mismatch artifact was introduced by replacing a full electrode with an electrode cut in half thus reducing contact surface. Separate experiments were implemented where half electrodes were provided in turn for current carrying and voltage electrodes. Impedance spectrum comparisons for the different test cases showed that current carrying electrode mismatch resulted in very slight changes to measured values. Voltage electrode mismatches resulted in larger artifacts. Specifically a significant increase in resistance (real component of impedance) and impedance modulus was measured, although this effect decreased as frequency was increased. Control of voltage electrode-skin contact is thus critical to ensure reduction of this artifact.

The non-invasive two electrode measurement is often dominated by the high impedance

of the outermost layer of the epidermis, referred to as the stratum corneum (SC). The SC is comprised of closely packed keratinised cells that act as barrier against moisture, radiation, microbes and chemicals. This high impedance layer generates impedance spectra that often can not be interpreted by the standard Cole model (see Section 3.1.2); in addition regions of interest are usually in deeper tissue layers which provide minimal contribution to the total impedance measurement.

An additional factor that contributes to noise is polarization that occurs at the boundary interface between the electrode and tissue. This polarization causes an artificial contribution to measurement impedance referred to as electrode polarization impedance (EPI). EPI is frequency dependent, can mimic the dispersion-like behaviour of tissue and provides an undesired contribution to impedance especially at low frequencies.

Boundary artifacts can be avoided or reduced by stripping away the SC layer, by invasive insertion of electrode contact points in deeper tissue layers and by specific denoising processing techniques (see Section 3.5). Alternatively a tetrapolar electrode arrangement is used whereby current injection and voltage measurements occur with different electrode pairs. This results in an impedance measurement that is characteristic of tissue deeper than the boundary layer.

In a tetrapolar arrangement with separate current carrying (CC) and voltage pick up (PU) electrodes transfer impedance is measured. When voltage amplifiers are ideal no current passes through the PU electrodes and therefore electrode-boundary impedance is not introduced in the measurement. The contribution of each volume element of the tissue to the measured impedance is determined by the sensitivity of the electrode system; this is defined as

$$S = J_{cc} \cdot J_{reci} \tag{2.1}$$

and transfer impedance is

$$Z = \int_V \rho J_{cc} \cdot J_{reci} \tag{2.2}$$

where J_{cc} and J_{reci} are current density vectors representing unity current through the current carrying (CC) and pick up (PU) electrodes respectively and ρ is the specific impedance of each tissue element. Transfer impedance is then affected by the arrangement and spatial configuration of the PU and CC electrodes. In general increasing inter-electrode distance results in a impedance measurement corresponding to deeper tissue layers. Analysis of the sensitivity field is important for proper interpretation of tetrapolar impedance. For instance a negative sensitivity field may result from certain electrode arrangements and provide measurement artifacts in the form of negative impedance values.

Sources of error of *in-vivo* BIS impedance measurements may include interaction of stray capacitance with elements of the measurement device as well as other devices. A study compared the impedance output of two BIS devices (Xitron 4000B and an SFB3) in an ICU setting where the environment and participants were controlled so as to be comparable for both experiments. The instruments used a tetrapolar lead arrangement, a constant current source and measured impedance at logarithmically spaced frequencies (5-500 KHz - Xitron 4000B, 4 -1000 KHz - SFB3). Measurements taken in the ICU room resulted in significant changes to the high frequency end of the spectrum when compared to measurements taken in a room free of electronic devices. These changes were hypothesized to be predominantly

due to the interaction of stray capacitance between the device and surrounding objects. The authors of the study recommend minimizing stray capacitance by active screening of the pick up voltage leads and by careful design of the current drive system.

Research has shown evidence of the role of biological tissue properties in causing deviation in impedance spectra from theoretical model predictions. Whole body and segmental BIS impedance spectrum were simulated using an equivalent electric circuit in the frequency range 5-1000 KHz; stray capacitance pathways considered included cable capacitance, inter-electrode lead capacitance, capacitance between body segments and the earth and capacitance between the device and earth. The latter two pathways were shown to cause significant dispersion artifacts for frequencies over 500 KHz. For segmental measurements, impedance spectra was found to be sensitive to switching electrode leads. Remaining artifacts are thought to be caused by the superposition of different tissue type dispersions.

Other sources of error are related to the intrinsic variability of biological tissue. Change in tissue geometry, blood perfusion and skin conductivity are all factors that affect impedance. Studies have shown that the tissues samples undergo significant conductivity changes a short time after extraction.

3 Modelling

3.1 Model Accuracy

Bioimpedance data acquisition provides a measurement of impedance amplitude and phase, Z_i, at a set of frequencies, f_i. As discussed in the previous section, these measurements are subject to both random and systematic sources of error, which vary as a function of frequency, typically with large errors at higher frequencies. The next step in data analysis is typically to fit the measured values to a model and to extract the best-fitting parameter values.

Models, $M(\cdot)$, are linear or non-linear functions of a small number of parameters, p, which predict the impedance, $\hat{Z}_i$ as a function of frequency. Thus,

$$\hat{Z}_i = M_p(f_i), \tag{3.1}$$

where M is evaluated using parameters, p. Model fitting is the process of finding the parameter values, $\hat{p}$, which best fit the measured data, as expressed by

$$\hat{p} = \underset{p}{\operatorname{argmin}} \left(\sum_i \left|\left| Z_i - M_p(f_i) \right|\right| \right) \tag{3.2}$$

This equation can be interpreted as finding the p values for which the model's predictions are as close as possible to the measured Z values. Several variants of the norm $||Z_i - M_p(f_i)||$ are used to account for differences in the experimental conditions. One important consideration is to use a weighted norm at each frequency proportional to the data reliability at that frequency (i.e. the inverse of the estimated error).

If the model is linear, there is often a one-step algebraic solution for the parameter values. However, for non-linear models, an iterative solution for (3.2) is required. In most cases, such algorithms are not guaranteed to find a correct solution. Instead, they suffer from the problem

of local minima, where the calculated solution is less than nearby values, but is not the overall best parameter.

The most important limitations of models is due to the systematic errors introduced by the instrumentation. For example, cables can introduce a bias due to parasitic capacitance and impedance at high frequency. These instrumentation-related changes will occur in addition to the "real" tissue-related changes in the sample. A model fitting will choose parameters which match both the real and the parasitic effects. However, if the model parameters are designed to explain only frequency-related changes in the tissue properties then the interpretation will be incorrect.

The most important method to manage model accuracy is through careful calibration. The experimental configuration must be measured with (at least one, but preferably multiple) samples of known properties from which the instrumentation biases can be measured. Using such a calibrated system allows measurements to be associated with tissue rather than the instrumentation.

The complexity of biological system necessitates using certain simplifying assumptions to provide meaningful ways of analyzing and interpreting the data. These simplified assumptions form the basis of two general methods of representing bioimpedance data.

Explanatory models consider the anatomical and electrical models of the tissue being measured. Generally tissue electrical properties are considered as lumped, e.g capacitance and conductance may be modelled as a single circuit comprised of a resistor and capacitor in parallel. Model complexity may however range from single electrical circuit models to complex finite element models that aim to represent both tissue anatomy and electrical properties with a high degree of fidelity. Here we consider the most common equivalent circuit models.

Descriptive models consider biological systems as black boxes; this phenomenological approach emphasizes analysis of parameters such as impedance, phase, time course and frequency without necessarily any attempt of correlation with known or assumed electrical properties of the measured tissues. Included in this model family are multivariate models that provide a method of describing features based on intrinsic organization and structure of the data.

3.2 Explanatory Models

Biological tissue can be modeled as an aggregation of cells with high impedance membranes containing intra-cellular fluid surrounded by extra-cellular fluid. At low frequencies, current will tend to flow through extra-cellular fluid, whereas high frequency currents will shunt through cellular membranes and intra-cellular fluid. In general the electrical properties of biological materials are dependent on internal tissue structure and chemical composition. At specific frequency bands rapid drops in impedance, referred to as dispersion zones, occur. Schwan identified three major dispersion zones in living tissue referred to as the α, β, and γ dispersions where α, β, and γ correspond to frequencies 100 Hz, 1 MHz and 10 GHz respectively. The β dispersion is thought to be affected by macroscopic tissue structures such as cell membranes, oedema and membrane polarization. This dispersion range is thus considered to provide most of the clinically relevant information and is therefore of primary importance in the design of BIS experimental and instrument design. Equivalent circuit models are therefore

(a) Debye (b) Cole

Figure 3.1: Idealized circuit diagrams for (a) Debye and (b) Cole models

designed to mimic the dispersion of biological tissue through an arrangement of resistors and capacitors.

3.2.1 DEBYE SINGLE DISPERSION MODEL

Simple systems that are characterized by a single relaxation time constant have a single dispersion, referred to as the Debye single dispersion, separating two permitivity zones at low and high frequencies respectively. The permittivity levels are separated by a transition zone around a characteristic relaxation frequency f_c. This system can be modelled by a $1R-2C$ electrical circuit. The circuit has no parallel resistor to allow DC flow since this represents a purely dielectric material with no free flow of ions. In terms of admittance Y where $Y = \frac{1}{Z}$, the complex admittance and capacitance of this circuit is written as:

$$Y = j\omega C_\infty + \frac{j\omega\Delta C}{1 + j\omega\tau} \tag{3.3}$$

where the characteristic time constant $\tau = R\Delta C$. The characteristic relaxation frequency of the circuit is $f_c = \frac{2\pi}{\tau}$. The relationship (3.3) is illustrated as a circuit in Figure 3.1a. At low frequencies the impedance of the circuit is largely a function of the parallel combination of C_∞ and ΔC, whereas at high frequencies admittance of the circuit is dominated by C_∞.

3.2.2 COLE MODEL

When tissue reactance is plotted in the complex plane against real resistance, an arc of a semi-circle results, with frequency intercepts on the real resistance axis at R_0 and R_∞. In order to account for an observed depression of the center of the semi-circle below the real resistance axis, Cole proposed modifying the capacitor of the 2R-1C electrical circuit to an ideal conductance element G in parallel with a non-realizable electric element referred to as a Fricke constant phase element (CPE). This proposed circuit for modelling impedance data is referred to as the Cole$_z$ system. The CPE is characterized by phase value Φ that can be anywhere between 0° and 90°. The CPE is modelled as a parallel resistor and capacitor; both these elements are frequency-dependent so that the phase can be set to a constant which is independent of frequency. The tissue impedance spectra can then be modelled as

$$Z(f) = R_\infty + \frac{R_0 - R_\infty}{1 + (j\omega\tau)^\alpha} \tag{3.4}$$

Changing the resistance terms R in the fraction to conductance terms G, so that $\Delta R = R_0 - R\infty = 1/\Delta G$, (3.4) can be rewritten as

$$Z(f) = R_\infty + \frac{1}{\Delta G + \Delta G(j\omega\tau)^\alpha} \tag{3.5}$$

where, $Z(f)$ is the frequency dependent impedance, ω is the angular frequency in Hertz, α is a dimensionless phase quantity between 0 and 1 and τ is RC - the characteristic time constant of the tissue. The relationship (3.5) is illustrated as a circuit in Figure 3.1b.

The Cole equation may be rewritten in admittance form, as

$$Y = G_o + \frac{\Delta G}{1 + (j\omega\tau_\gamma)^{-\alpha}} \tag{3.6}$$

where G_0 is the conductance at low frequency.

3.2.3 Feature Extraction of Explanatory Models

Cole model parameters are the most common explanatory features extracted from bioimpedance measurements. Cole modelling results in the reduction of impedance spectra to characteristics of a lumped electronic circuit and is the primary method used for feature extraction.

The essence of Cole parameter extraction is the fitting of a semi-circular arc in the complex plane where the horizontal axis is the resistance or the real part of the impedance and the vertical axis is the reactance or the imaginary impedance component (Figure 3.2). The right and left intersection of the arc with the horizontal axis are R_0 and R_∞ respectively. The semi-circle thus has an approximate radius of $\frac{R_0 - R_\infty}{2}$ when α approaches one. α is a dimensionless quantity that assumes values between 0 and 1 and provides a measure of the position of the Cole semi-circle with respect to the horizontal (resistance) axis , and τ is the characteristic time constant which corresponds to characteristic frequency f_c (at which reactance is at a maximum). The resulting extracted parameters can then be defined by the vector $m = [R_0, R_\infty, f_c, \alpha]$.

Cole parameters can be fitted without direct impedance measurements (through knowledge of the measurement filter transfer function and gain characteristics), by impedance magnitude measurements alone, or through use of both real and imaginary impedance components. Advantages of the first two methods are a reduction in complexity of device design since phase measurements are not required; in practice however Cole fitting algorithms usually consider real and imaginary impedance since most BIS devices are designed to measure phase information.

Biophysical interpretation of the Cole parameters is challenging due to the simplifying model assumptions as well as noise factors that affect the measurements. In general R_0 and R_∞ are considered to be indicative of tissue conductive and capacitive properties as well as functions of body segment geometry and electrode configuration. f_c and α are properties of measured tissue; α is considered a measure of tissue heterogeneity scaled between 0 - highly

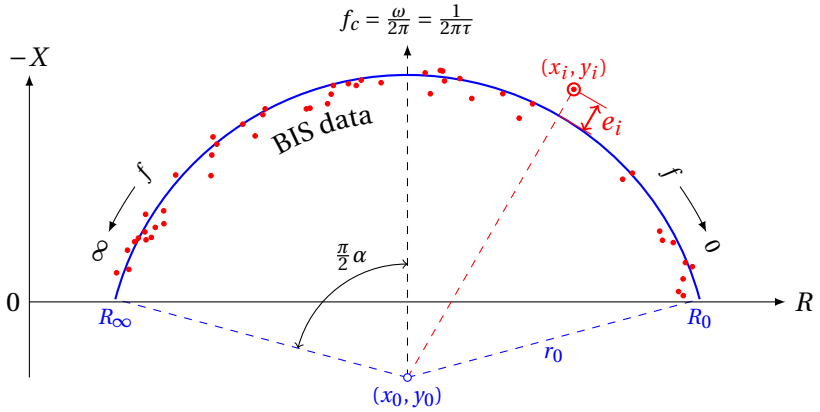

Figure 3.2: BIS data errors; data (red) collected at a variety of frequencies f are fitted to a Cole model $\alpha, F_c, R_0, R_\infty$ (blue) with error e_i for each data point (x_i, y_i)

homogeneous and 1 - highly heterogeneous, whereas f_c is the characteristic frequency at maximum tissue reactance. α has been shown to be sensitive to measurement variability and is relatively insensitive to variations of tissue properties whereas f_c has a higher sensitivity to different tissue properties but is relatively stable to measurement configuration changes. f_c is thus of singular importance for discrimination between various tissue types.

3.2.4 Gradient Minimization Fitting Methods

Cole model fitting methods F are usually based on the least squares (LS) gradient minimizing method. The radial error e_i is the distance between the data point (x_i, y_i) and the fitted semi-circle (see Figure 3.2). The total LS error function for N measured frequencies is then:

$$F_{\mathrm{LS}}(x_0, y_0, r_0) = \sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} \left(r_0 - \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \right)^2 \tag{3.7}$$

where x_0, y_0, r_0 and x_i, y_i are fitted and experimental points respectively. Optimally fitted points are obtained by iteratively minimizing the error function through the gradient method. In general BIS data fitting is a non-linear process, thus the reduction of the error function is referred to as a non-linear least squares (NLLS) problem. Limitations of this method are that standard NLLS algorithms are sensitive to outliers and non-normally distributed errors. This type of noise is common in BIS measurements since data is usually not normally distributed due to the limited number of frequency points.

An alternative fitting method is the least absolute deviation (LAD) algorithm. The LAD error function is the sum of the absolute values of e_i and is written as:

$$F_{\mathrm{LAD}}(x_0, y_0, r_0) = \sum_{i=1}^{N} |e_i| = \sum_{i=1}^{N} \left| r_0 - \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \right| \tag{3.8}$$

LAD performance and accuracy were shown to be superior to typical LS methods for simulated data. When Cole parameters were extracted from simulated data with outliers (30% deviation

at 2 frequency points) and with random noise (±10% at 16 frequency points), LAD derived parameters had significantly lower error than parameters fitted by the LS method. The LAD method is however computationally more demanding than the LS algorithm.

3.3 Evolutionary Computational Fitting

Despite the superior performance of LAD relative to LS, both methods are still affected by noise and may converge to a local minimum. Stochastic methods from the family of evolutionary computation have been evaluated as alternate fitting techniques. The principal steps of evolutionary algorithms can be summarized as: generation of a random group of individuals, evaluation of the fitness of the individuals based on predetermined criteria (i.e. convergence), selection of the best fit individuals for reproduction, breeding new individuals, evaluation of the fitness of the new individuals and repetition of these steps until the problem has converged to a desired level.

An example of an evolutionary computation method is the Bacterial Foraging Optimization (BFO) algorithm which is designed to mimic the swarm foraging and cooperative behaviour of *E. Coli*. The BFO algorithm initializes bacteria population position, evaluates the *fitness* of each bacteria based on predefined criteria, simulates random and directed bacteria movement, reproduces healthy bacteria based on fitness criteria and reinitilizes bacteria position until convergence is reached.

In the context of BIS measurements, bacteria position can be defined based on a fitness function J as:

$$J = \sum_{i=1}^{N} |Z_i - F_i| \tag{3.9}$$

where Z_i and F_i are experimental and measured BIS data points respectively for N number of frequencies. Bacteria position is represented by

$$\theta^i(j,k,l) \tag{3.10}$$

which represents the i^{th} bacterium at the j^{th} chemotactic, k^{th} reproduction and l^{th} elimination-dispersal step. The fitness function of the i^{th} bacterium can then be written as

$$J(i,j,k,l) = J(i,j,k,l) + J_{cc}\left(\theta^i(j,k,l), P(j,k,l)\right) \tag{3.11}$$

where $P(j,k,l)$ are position coordinates of the i^{th} bacterium and J_{cc} is a function that represents added fitness by simulating bacteria swarming (including cell-to-cell attraction and repelling effects).

BFO algorithm performance was evaluated by fitting simulated BIS data to the Cole model with added radial random noise (±10% and ±30%). Average Cole parameters were estimated after running the algorithm 50 times. When compared to a standard LS algorithm, BFO fitting had a lower relative error and lower standard deviation. The LS fitting performance deteriorated as noise levels were increased whereas the BFO algorithm was much more robust with a slight decrease in accuracy resulting from the noise increase.

A significant limitation of evolutionary algorithms is higher computational complexity. The LS method is much more efficient in terms of execution time. This suggests that for

applications where speed and efficiency is paramount, such as real-time monitoring, the LS method may be more suitable. In most cases however the superior accuracy of the evolutionary algorithm fitting methods outweigh the limitations caused by slow processing speeds.

3.4 Descriptive Models

Principal Component Analysis (PCA) is a non-parametric *descriptive* method of modelling and analyzing data where no assumptions are made concerning the physiology or physics of the measured phenomenon. PCA is used to reveal internal structure of the data in a way that best explains the variance in the data.

In the context of BIS measurements a data matrix X, can be generated composed of $M \times N$ rows and columns where M is the number of *observations* or measurements and N is the number of frequencies at which the measurements are taken. The data is then mean centered $X - \bar{X}$. The covariance of the mean adjusted data is calculated as

$$C_x = [c_{ij}] = \frac{1}{M-1} \sum_{k=1}^{M} x_{ki} x_{kj}, \quad i, j = 1, 2, ..., N \tag{3.12}$$

where x_{ij} is the element corresponding to the i^{th} impedance measurement at the j^{th} frequency. C_x can then be expanded according to

$$C_x = \frac{1}{N-1} \sum_{i=1}^{N} \sigma_i^2 V_i^T V_i \tag{3.13}$$

where $\sigma_i = \sigma_1 \geq \sigma_2 \geq ... \geq \sigma_N \geq 0$ are singular values and $v_i = [v_{1i}, v_{2i}, ..., v_{Mi}]$ where $i = 1, 2, ..N$ are singular or eigenvectors of C_x respectively. The expansion of the covariance matrix provides a linear orthogonal basis that represents the directions of variance in the data.

The singular value decomposition (SVD) may be used to efficiently calculate the PCA or the singular values and vectors may be used directly. The SVD technique involves direct expansion of the data matrix X of M measurements (rows) and N frequencies (columns) according to

$$X_{MxN} = \sum_{i=1}^{N} U_i \sigma_i V_i^T \tag{3.14}$$

where $\sigma_i = \sigma_1 \geq \sigma_2 \geq ... \geq \sigma_N \geq 0$ are singular values of X and U_i and V_i are the left and right singular vectors of X.

Generally singular values are truncated beyond the rank r of the data matrix, where r is assumed to be smaller than the number of frequencies; non-zero singular values beyond the r-th singular value are assumed to be caused by noise.

3.4.1 Feature Extraction of Descriptive Models

In PCA, principal components indicate the direction and magnitude of data variability; axis orientation is provided by the covariance eigenvectors and magnitude by eigenvalues. Principal components are ranked in order with the largest principal component corresponding with the axis oriented along the largest direction of data variability. Usually most of data

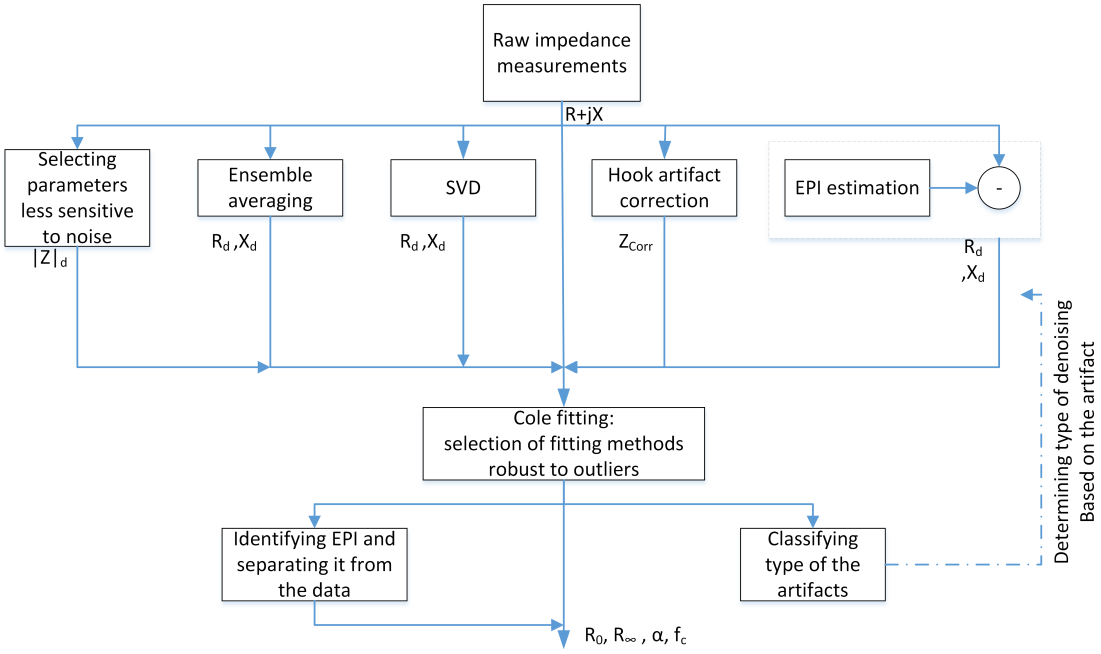

Figure 3.3: Denoising methods presented in BIS literature. R and X in the figure represent resistance and reactance over a defined frequency range. R_d and X_d represent the resistance and reactance over the same frequency range after denoising.

variance and meaningful information is contained within the first few principal components, thus PCA provides a useful method of reducing dimensionality of multivariate data, like BIS measurements, without loss of essential information.

Feature extraction is generally comprised of PCA scores which are obtained by mapping (by the dot product operation) impedance spectra on the new set of uncorrelated orthogonal vectors V_i. The scores, essentially projections of impedance measurements on the new orthogonal basis, are considered as features of the spectrum and can be fed to a classifier for further analysis.

Although most of the information content of the data or covariance matrix is represented in the first few eigenvectors, the precise demarcation between essential information and noise can be subject to trial and error depending on the application. For example a classification task of categorizing arm position achieved highest accuracy when the second principal component was selected as an essential feature of the impedance spectra.

3.5 Denoising

Reduction of noise and pre-processing are crucial elements of analysis and interpretation of inherently variable and noisy data like bioimpedance measurements. Sources of noise as well as ways to reduce noise and artifacts by following specific experimental procedures or by using particular measurement configurations were detailed in Section 2. In this section, we will focus on algorithmic solutions for reducing the noise. The denoising block shown

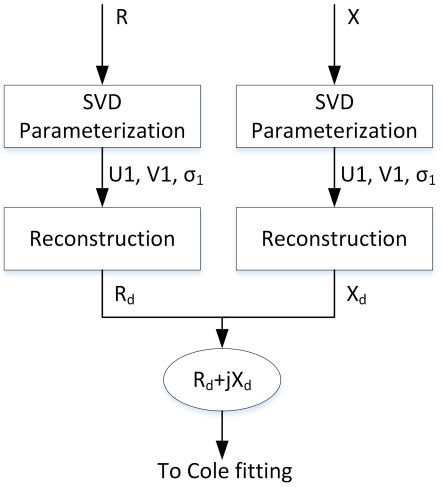

Figure 3.4: SVD denoising sequence. Real and imaginary impedance components are parametrized using SVD to the singular vector V_1. Remaining vectors are discarded, BIS complex data is reconstructed and Cole parameters are then extracted.

in Figure 1.1 has raw bioimpedance data as an input and produces bioimpedance data with improved signal-to-noise ratio at the output that is connected to modelling blocks. More detailed representation of denoising steps are shown in Figure 3.3. As can be seen in the figure, the majority of denoising methods are implemented on raw bioimpedance data and therefore are part of the pre-processing steps. A number of methods however are described as processing and classification of noise after data modelling; these techniques perform denoising as part of the data post-processing.

Reducing the complexity of the BIS instrumentation and fitting algorithms is one strategy being pursued to eliminate some sources of error; for instance, methods of fitting Cole parameters from resistance or impedance magnitude only measurements have been developed. These type of methods are described as “selecting parameters less sensitive to noise” in Figure 3.3. Parameters extracted from these measurements have been demonstrated to have an accuracy exceeding or comparable to parameters extracted from complex impedance (real and imaginary components). The implications of reduced hardware complexity is that removal of phase sensitive measurements eliminate the cause of many errors such as noise generated from capacitive leakage.

Ensemble averaging can reduce zero mean, unbiased noise from impedance measurements. Averaging can be calculated immediately after data acquisition or following extraction of Cole parameters. However, averaging of model parameters is known to introduce new errors so averaging raw data is the preferred approach.

In some cases averaging repeated measurements is not effective because noise may have systematic biases even for single subject measurements. In such cases the SVD method has been used successfully as a preprocessing, denoising method (sequence steps are shown in Figure 3.4). The raw impedance data is first decomposed by the SVD method to a new

orthogonal basis. The first singular vectors of the real and imaginary components are assumed to represent the primary information content and they are retained, while the remaining vectors are discarded. The real and imaginary impedance are then reconstructed and Cole parameters are extracted. In order to compare the efficacy of SVD-based noise removal, BIS data was simulated using an equivalent circuit mimicking the electrical properties of the various arm tissues (muscle, fat and blood) and additive Gaussian white noise (AGWN) was added to simulated voltage measurements in the time domain. Cole parameters were estimated with simulated data with and without pre-processing. The mean and the standard deviation of relative error of Cole parameters extracted from SVD-preprocessed data was shown to be much less than for parameters derived from uncorrected data.

The Hook effect appears as a hook-like deviation in Cole semi-circle - most noticeably at high frequencies. This artifact is caused by capacitive leakage which generates errors in all impedance components (phase,reactance, resistance and modulus). Time delay (T_d) compensation, an established method of removing the hook deviation, consists of multiplying the measured impedance spectrum, $Z_{meas(\omega)}$ by an exponential in the form of $e^{-j\omega T_d}$ where T_d is a scalar. Deficiencies of the time delay method include the inability to correct phase and magnitude distortions at all frequencies with a single scalor factor. A newer compensation approach involves multiplying the impedance spectra with a complex valued, frequency dependent value of Td that is a function of both the measured impedance $Z_{meas(\omega)}$ and parasitic capacitance C_{par}. Specifically Td is given by:

$$T_d(\omega) = \frac{Log[1 - j\omega Z_{meas(\omega)} C_{PAR}]}{j\omega} \tag{3.15}$$

the corrected impedance is then:

$$Z_{Corr(\omega)} = Z_{meas(\omega)} \frac{1}{1 - j\omega Z_{meas(\omega)} C_{PAR}} \tag{3.16}$$

As frequency is increased, tissue susceptance, the inverse of reactance, decreases towards 0. Increase of susceptance at high frequencies is due to capacitive coupling which occurs in parallel to the measurement circuit; the susceptance slope can therefore be used to estimate C_{par}.

Polarization of the electrodes, the EPI artifact, can be eliminated by modifications to the measurement setup presented in Section 2. Other methods include estimation of EPI by calibration, or by varying measurement electrode distances, followed by subtraction from the measurement data as shown in Figure 3.4. Post-process modelling techniques can be used for removal of EPI: tissue and EPI are generally modelled as a series of Cole impedances which can be converted to a parallel admittance circuit. The advantage of post-processing is that for low frequencies where EPI is dominant, admittance converges to a low value; whereas low frequency impedance will become very large or diverge. Assuming the characteristic time constants of EPI and measured tissue are sufficiently different, the influence of EPI can be identified and separated from the total measurement data.

The use of supervised machine learning has been proposed as a post-processing method for identification and classification of different types of artifacts in BIS impedance spectra.

First, five frequency bands were defined as a function of characteristic frequency f_c: VLF ($f < f_c/5$), LF ($f_c/5 < f < f_c/2$), MF ($f_c/2 < f < 2f_c$), HF ($2f_c < f < 5f_c$) and VHF ($f > 5f_c$). Next, analysis of a database of 1502 real BIS spectra was used to generalize impedance artifacts in 6 different categories distributed through the frequency bands. These errors included various degrees of Hook effect due to capacitive leakage and abnormal decrements in reactance. The next step was feature extraction; for each impedance measurement the magnitudes of resistance, reactance, conductance, susceptance, impedance modulus, and angle were calculated after fitting to the Cole model. The relative error between the fit and measurement were then calculated over the defined frequency bands providing a total of 30 features. The features were labelled and then fed to a least squares linear discriminant analysis (LDA) classifier. The classifier was then trained by adjusting the weights using an evolutionary algorithm (EA) to reduce the mean squared error (MSE) (defined as the difference between the LDA output and true values). The potential efficacy of this approach was demonstrated by classifier performance; impedance spectra were labelled with high accuracy giving error rates as low as 0.16%.

With recent advances in machine learning, we expect that classification and machine learning methods for identifying and removing noise and artifacts will be the focus of increased research efforts over the next several years. This potential step of using classifiers for intelligent noise removal is illustrated with the dotted line in Figure 3.4.

The best approach for noise removal is not limited to a single technique and may be based on a combination of the methods described here. In all cases sources of noise and impedance artifacts should be carefully studied before defining an optimal strategy of pre- and post-process denoising techniques.

4 Data Classification

4.1 General Classification Sequence

Classification is defined as the problem of identifying to which set of categories or classes a new observation belongs. Generally the classification sequence (see Figure 4.1) begins by extraction of features from a set of measurements; in the case of multivariate data such as BIS measurements, data dimensionality is usually reduced through fitting to a simplifying model or by other parametrization techniques. Parameters can be derived from explanatory models which are based on tissue equivalent circuits such as the Cole model. Alternatively features can be extracted from descriptive multivariate based methods such as PCA or SVD which do not consider the biological origin or structure of the data source. The data features are then provided to the classifier for training whereby the classifier, using a predefined rule based system, separates the measurements into categories. The trained classifier is then used to categorize a new unlabeled measurement test set based on features derived from the same parameter extraction method.

Appropriate classifier selection is comprised of a number of steps including:

- determining essential features that are relevant to the type of diagnostic decision required (for example, conductivity of tumors is known to be different than healthy tissue)

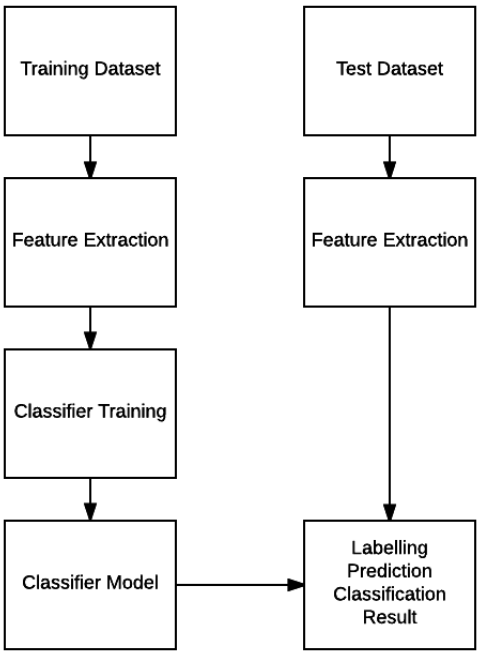

Figure 4.1: Schematic of general classification sequence. Features extracted from a training dataset are used for training a classifier. Parameters extracted from a test dataset are fed to the trained classifier which then provides a prediction of the class label of the new data.

- determination of the optimal classifier (best decision boundaries) in the feature space for discrimination between classes
- setting the required accuracy of the classifier

Classifier performance is usually quantified by comparison to a "gold standard." In the context of BIS, measurements may be used to assist in formulating a diagnostic decision about the malignancy of a tissue. Classification performance can be compared to direct clinical measurements such as tissue biopsy and lab analysis.

A classification confusion matrix is shown in Table 4.1. The percentage of correct determinations of tissue pathology (true positive (TP)) is referred to as the classifier sensitivity $Se = TP/(TP+FN)$, whereas the percentage of correct determinations of tissue health (true negatives(TN)) is the specificity of the classifier $Sp = TN/(TN+FP)$. In our example a false positive error (FP) is the incorrect classification of tissue as malignant, whereas a false negative error (FN) is the incorrect classification of tissue as healthy. Classifier accuracy can then be quantified as the ratio of correct decisions made over total decisions (correct and incorrect); stated formally this is:

$$ACC = \frac{(TP+TN)}{TP+FP+FN+TN} \tag{4.1}$$

Although 100 % accuracy is desirable, in practice as sensitivity increases, specificity is usually reduced; this tradeoff is demonstrated in a receiver operating characteristic (ROC) curve where sensitivity is plotted vs the FP rate or 1 - specificity. Performance is commonly defined as the area under the curve (AUC) which is a measure of the classifier discrimination. An AUC of 1 demonstrates (a technically unachievable) perfect discrimination whereas an AUC of 0.5 is indicative of a random predictor.

4.2 Classifier types

Both statistical and hierarchical supervised classifiers are used for BIS classification tasks. Classification trees are hierarchical classifiers for categorical variables which provide a high degree of flexibility and exploratory power while not assuming any underlying probability distribution. Each feature with the potential for classifying a data set occupies a node in the decision tree which partitions the data set into two subgroups.

Statistical classifiers include different types of supervised learning methods such as linear discriminate analysis (LDA), k nearest neighbour (k-NN) and artificial neural networks (ANN). The LDA classifier assumes Gaussian class-conditional densities where classes have equal covariance. The boundary between classes is linear which represents a limitation for typically non-linear BIS measurements. LDA is computationally efficient and can be computed directly from the data without using any search algorithms. LDA can however result in over-fitting; a proposed solution has been to use PCA for data dimensionality reduction prior to LDA classification.

k-NN returns k points closest to a new data point and assigns a label based on the status of the majority of the k returned points. A smaller value of k leads to smaller training errors while larger values lead to more stable predictions due to a voting effect. In BIS applications, accuracy of k-NN classifier was improved after data dimensionality had been reduced using PCA.

Soft independent modelling of class analogy (SIMCA) consists of a collection of PCA models, each one defined for a different class. Limits are defined around the class of each PCA model. A new data point is classified based on the distance between the point and the nearest class. A limitation of the SIMCA method is the difficulty of interpreting and relating inter-class distance to data categories.

ANNs require relatively large training sets; this is one of the reasons why they are not commonly used in BIS applications where the number of subjects are relatively small. ANNs

Table 4.1: Confusion matrix

-	Predicted condition - positive	Predicted condition - negative
True condition - positive	**True positive**	**False Negative (Type II error)**
True condition - negative	**False Positive (Type I error)**	**True negative**

perform well when dealing with multidimensional features and when there is a non-linear relationship between the input and output features.

4.3 Applications

Table 4.2 shows some examples of BIS data used in classification tasks for both assisting in a diagnostic decision and non-diagnostic applications.

Classification applications for assisting in diagnosis include discrimination of normal versus cancerous tissue mainly for breast and skin cancer, detection of diabetes related change in the skin and distinguishing between stroke patients and healthy volunteers.

Body composition is known to change with diabetes due to skin and connective tissue related changes - beginning first in the arms and legs. Classification based on principal component regression was used on a sample of 16 subjects with diabetes and 12 control subjects. First the measurements were classified according to gender (male and female) based on PCA scores of capacitance. After separation of the data by gender, data was further classified into patient and control groups. Use of leg and feet measurements resulted in the highest classification accuracy.

For skin cancer application, two types of electrodes were considered: non-invasive and microinvasive electrodes. Features were extracted using PCA and then fed to an LDA classifier. A separation of 96% sensitivity and 86% specificity between benign nevi and basal cell carcinoma was achieved using the regular non-invasive probe while a similar high sensitivity/specificity separation between benign nevi and malignant melanoma was achieved using the microinvasive electrodes.

Cole parameters R_0 and R_∞ were used for distinguishing between normal and stroke patients. Classification was based on a set of pre-defined criteria characterizing healthy and pathological states. The study involved 6 volunteers and 3 patients who suffered a unilateral stroke. In healthy controls, 100% of measurements were successfully classified, while a lower accuracy was achieved in detecting patients with stroke.

Non-diagnostic application include monitoring tasks such as determining the onset of hypoxia or measuring body parameters such as relative levels of fat and fluid. Other applications include measuring muscle status (contracted or relaxed), categorizing tissue type and classifying body orientation (arm position).

Sixteen measurements were obtained from a single subject with relaxed and contracted biceps brachii muscle where the arm was oriented in different positions. SIMCA was applied to complex impedance values consisting of magnitude and phase; contracted and relaxed muscles were classified with an accuracy of 80%.

A Bayesian classifier was used to determine different types of tissues at the tip of a needle containing a bioimpedance probe. An animal study demonstrated that the classifier could distinguish between fat, muscle, tendon and blood tissues with an accuracy higher than 85.5%.

The accuracy of classification algorithms was compared for a bench mark test described as properly labeling arm orientation (up, down and horizontal). Data was collected from eight subjects over three different sessions. PCA and the Cole model were used for feature extraction; features were then fed into various classifiers for evaluation of accuracy. Classifiers tested included SIMCA, LDA, quadratic discriminant analysis (QDA), decision tree and k-NN. It was

Classification Type	Description
Assisting in Diagnostics	
Lindholm et al. 1998	**Application**: Detection of diabetes-related changes in the skin **Validation**: Skin biopsy **Number of subjects** 28: **Classification method** Principal component regression:
Atefi et al. 2012	**Application**: Detection of stroke **Validation**: Previous diagnosis of stroke **Number of subjects**: 6 healthy controls and 3 stroke patients **Classification method**:Rule-based using R_0 and R_∞ Cole parameters
Aberg et al. 2005	**Application**: Skin cancer **Validation**: Skin biopsy **Number of subjects**: 99 subjects with pigmented benign nevi, 28 with basal cell carcinoma, and 13 with malignant melanoma **Classification method**: PCA and LDA
Non-Diagnostic	
Zagar et al. 2008	**Application**: Muscle contraction **Validation**: **Number of subjects**: 1 volunteer 32 sets of data **Classification method**: PCA + SIMCA
Nejadgholi et al. 2015	**Application**: Posture (arm orientation) detection **Validation**: **Number of subjects**: 8 volunteers **Classification method**: PCA + k-NN had the best accuracy, specificity and sensitivity
Kari et al. 2015	**Application**: **Validation**: In vivo tissue classification **Number of subjects**: Animal study **Classification method**: Bayesian classification

Table 4.2: BIS classification applications

shown that PCA features (as opposed to Cole parameters) fed to a k-NN classifier resulted in the highest accuracy (90%) for determining arm orientation.

Classifier accuracy is also determined by choice of fitting method to the Cole model. As described earlier, a classification task was defined as properly categorizing arm position (up,

down and horizontal). Cole parameters were extracted from the raw impedance data using the LS and BFO algorithm, including genetic algorithms (GA) and particle swarm optimization (PSO), and fed to various classifiers (k-NN, LDA, QDA and decision trees). In all cases, parameters extracted using the BFO method, followed by GA and PSO, resulted in higher classifier accuracy than parameters provided by the LS method.

5 Conclusion

The objective of this chapter was to present an overview of overview of the major processing processing steps of BIS data including modelling, denoising and classification. In order to provide context to the challenges inherent in denoising and classifying methods, we first described different types of systemic and random measurement errors as well as common artifacts.

As shown in Figure 1.1, sources of noise originate from non-ideal instrumentation and experimental conditions. Measurement setup also significantly affects error levels; for instance noise can be substantially reduced by appropriate measurement configuration such as using a tetrapolar setup, modification of electrode properties to reduce the EPI effect, enhancing electrode contact and so on. Denoising strategy (e.g. choise of denoising algorithm) is dependent on type of data artifact; characterizing noise type is thus an integral part of data processing.

Denoising is typically implemented after raw data acquisition - before modelling and feature extraction. Different types of denoising methods described in this chapter include averaging, SVD decomposition, as well as removal of known artifacts (e.g. Hook artifact). Denoising can also be applied as a post processing step after feature extraction and model fitting; for instance classifiers have been used to distinguish between spectral features that are characteristic of noise and those of clean data. The classifier approach represents a novel "smart method" of noise removal algorithms based on learned parameters; we believe this will be an increasingly important focus of future research.

Data reduction from the impedance spectra to several representative parameters is accomplished using explanatory or descriptive models. The most popular explanatory model is the Cole model where impedance data is fitted to a semi-circular arc in the complex plane. Both gradient based and stochastic optimization methods are used for fitting. The gradient method is more appropriate for applications that require fast processing such as on-line monitoring and the stochastic approach is better suited for applications requiring a high degree of accuracy. Typically PCA is used as descriptive method of modelling whereby data dimensionality is reduced to a set of core eigenvectors/values; this is a compact way of representing the complex multivariate data without losing essential information. This method also allows the removal of noise and other sources of variability.

The use of classifiers in labeling features extracted from both explanatory and descriptive models has been demonstrated in a number of studies. No consensus however exists yet concerning a universally acceptable classification method for the BIS applications being explored. Larger studies will allow for applying more advanced learning techniques due to the increase of data available for analysis and classifier training. We expect to see more of ANN, deep learning approaches and novel classifier combinations in the future to deal

with highly non-linear BIS data. Other important research directions include integration of BIS classification into larger diagnostic models, for example based on Bayesian networks; better characterization of non-linear tissue properties and use of quantification of uncertainty and sensitivity analysis for analyzing the sensitivity of various BIS models to different model parameters and inputs.

6 Further Reading

O. G. Martinsen and S. Grimnes, Bioimpedance and Bioelectricity Basics. Chapter 3 pp. 58-91, Chapter 4 pp. 93 - 124, Chapter 8 pp. 283-331, Academic Press, 2011

C. Gabriel, S. Gabriel and E. Corthout, "The dielectric properties of biological tissues: I. literature survey", Physics in medicine and Biology, vol.41,no 11, p. 2231, 1996.

References

[1] A. Smirnov, D. Nikolaev, and V. Kolesnikov, "On measurement errors of the impedance spectrum of human body in vivo," in *Journal of Physics: Conference Series*, vol. 224, no. 1. IOP Publishing, 2010.

[2] B. Lindholm-Sethson, S. Han, S. Ollman, I. Nicander, G. Jonsson, F. Lithner, U. Bertheim, and P. Geladi, "Multivariate analysis of skin impedance data in long-term type 1 diabetic patients," *Chemometrics and Intelligent Laboratory Systems*, vol. 44, pp. 381–394, 1998.

[3] D. Ayllón, F. Seoane, and R. Gil-Pita, "Cole equation and parameter estimation from electrical bioimpedance spectroscopy measurements-a comparative study," in *Engineering in Medicine and Biology Society. Annual International Conference of the IEEE*, 2009, pp. 3779–3782.

[4] D. Ayllón, R. Gil-Pita, and F. Seoane, "Detection and classification of measurement errors in bioimpedance spectroscopy," *PloS one*, vol. 11, no. 6, pp. 1–19, 2016.

[5] D. Tao and A. Adler, "In vivo blood characterization from bioimpedance spectroscopy of blood pooling," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, pp. 3831–3838, 2009.

[6] F. Seoane, S. R. Atefi, J. Tomner, K. Kostulas, and K. Lindecrantz, "Electrical bioimpedance spectroscopy on acute unilateral stroke patients: Initial observations regarding differences between sides," *BioMed research international*, 2015.

[7] H. Kalvøy, G. K. Johnsen, Ø. G. Martinsen, and S. Grimnes, "New method for separation of electrode polarization impedance from measured tissue impedance," *The open biomedical engineering journal*, vol. 5, pp. 8–13, 2011.

[8] H. Scharfetter, P. Hartinger, H. Hinghofer-Szalkay, and H. Hutten, "A model of artefacts produced by stray capacitance during whole body or segmental bioimpedance spectroscopy," *Physiological measurement*, vol. 19, no. 2, pp. 247–261, 1998.

[9] I. Nejadgholi, H. Caytak, M. Bolic, I. Batkin, and S. Shirmohammadi, "Preprocessing and parameterizing bioimpedance spectroscopy measurements by singular value decomposition," *Physiological measurement*, vol. 36, no. 5, pp. 983–999, 2015.

[10] I. Nejadgholi and M. Bolic, "A comparative study of pca, simca and cole model for classification of bioimpedance spectroscopy measurements," *Computers in biology and medicine*, vol. 63, pp. 42–51, 2015.

[11] J. Kari, K. Annala, P. Annus, V. Seppä, and K. Kronström, "A thin needle with bioimpedance measuring probe: tissue recognition performance assessed in in vivo animal study," Injeq Oy Ltd., Tech. Rep., 2015.

[12] M. Bolton, L. Ward, A. Khan, I. Campbell, P. Nightingale, O. Dewit, and M. Elia, "Sources of error in bioimpedance spectroscopy," *Physiological measurement*, vol. 19, no. 2, pp. 235–245, 1998.

[13] O. G. Martinsen and S. Grimnes, *Bioimpedance and Bioelectricity Basics*. Academic press, 2011.

[14] P. Aberg, I. Nicander, J. Hansson, P. Geladi, U. Holmgren, and S. Ollmar, "Skin cancer identification using multifrequency electrical impedance-a potential screening tool," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 12, pp. 2097–2102, 2004.

[15] P. Aberg, P. Geladi, I. Nicander, J. Hansson, U. Holmgren, and S. Ollmar, "Non-invasive and microinvasive electrical impedance spectra of skin cancer, a comparison between two techniques," *Skin Research and Technology*, vol. 11, pp. 281–286, 2005.

[16] R. Buendia, F. Seoane, and R. Gil-Pita, "A novel approach for removing the hook effect artefact from electrical bioimpedance spectroscopy measurements," in *International Conference in Electrical Bioimpedance*, vol. 224, no. 1. IOP publishing, 2010, pp. 1–5.

[17] R. Buendia, R. Gil-Pita, and F. Seoane, "Cole parameter estimation from the modulus of the electrical bioimpeadance for assessment of body composition. a full spectroscopy approach." *Journal of Electrical Bioimpedance*, vol. 2, no. 1, pp. 72–78, 2011.

[18] R. Buendía, P. Bogónez-Franco, L. Nescolarde, and F. Seoane, "Influence of electrode mismatch on cole parameter estimation from total right side electrical bioimpedance spectroscopy measurements," *Medical engineering & physics*, vol. 34, no. 7, pp. 1024–1028, 2012.

[19] S. Kotsiantis, "Supervised machine learning: A review of classification techniques," *Informatica*, vol. 31, pp. 249–268, 2007.

[20] S. Gholami-Boroujeny and M. Bolic, "Extraction of cole parameters from the electrical bioimpedance spectrum using stochastic optimization algorithms," *Medical & biological engineering & computing*, vol. 54, no. 4, pp. 643–651, 2016.

[21] S. R. Atefi, F. Seoane, T. Thorlin, and K. Lindecrantz, "Stroke damage detection using classification trees on electrical bioimpedance cerebral spectroscopy measurements," *Sensors*, vol. 13, no. 8, pp. 10 074–10 086, 2013.

[22] T. Zagar and D. Krizaj, "Multivariate analysis of electrical impedance spectra for relaxed and contracted skeletal muscle," *Physiological measurement*, vol. 29, no. 6, pp. 365–372, 2008.

[23] Y. Yang, W. Ni, Q. Sun, H. Wen, and Z. Teng, "Improved cole parameter extraction based on the least absolute deviation method," *Physiological measurement*, vol. 34, no. 10, pp. 1239–1252, 2013.