Notes

A great deal more is known than has been proved
Richard Feynman

Section 2.2
Determination of the embedding dimension by the method of false nearest neighbour is described by Kennel et al. (1992). Embedding for non-uniformly sampled time series is covered by Huke and Broomhead (2007).

Section 2.4
Alternative classification by coefficients of global dynamical models is considered in (Kadtke, 1995).

Section 2.5.4
The performance of various synchronization measures is compared by Ansari-Asl et al. (2006) for a number of numerical models of brain activity.

Section 2.7
Recurrence plots are a general tool to analyze dynamical systems, with manifold applications (Webber, Jr. and Zbilut, 1994; Marwan et al., 2007). Joint recurrence plots even allow to compare the dynamics of two dynamical systems defined on different phase spaces, overcoming the main theoretical problem when comparing systems measured with distinct modalities (Marwan et al., 2007). However, quantitative measures defined for recurrence plots do not fulfill metric properties and cannot be used in a multivariate context.
Section 3.1

A good introduction to the physiology of breathing is (Guyton and Hall, 2006). The book by Kulish (2006) contains some advanced topics and corresponding mathematical models.

The branching nature of the bronchial tree is described by Weibel (1963) and by the model of Horsfield et al. (1971). A three-dimensional generalization has been obtained by Kitaoka et al. (1999). Optimality principles that explain the geometry of the lung were considered by Weibel (1991) and criticized by Imre (1999).

The control of breathing is described in (Whipp, 1987; Bronzino, 2006). Mathematical models for this regulatory process include the basic Mackey-Glass model (Keener and Sneyd, 1998) and the famous Grodins model (Grodins et al., 1967a), both incorporating time delays.

General cardiovascular models are described in (Ottesen et al., 2004; Batzel et al., 2006).

Section 3.2

The forced oscillation method is reviewed in (MacLeod and Birch, 2001; Oostveen et al., 2003); the article by Nucci and Cobelli (2001) considers mathematical details. Normal values are described by Landser et al. (1982). Details of the frequency-domain approach and parameter estimation are given by Michaelson et al. (1975). The frequency-dependence of FOT was observed and validated in Jackson et al. (1987).

A few facts about lung mechanics are: The tidal lung volume is about 1L, with a dynamic driving pressure of about $-1\text{mmHg}$. Airway resistance is highest at segmental bronchi and lower at higher airway generations. Similarly, resistance decreases nonlinearly with lung volume from 4 (2 L) to about 0.5cmH2O/L (6 L); conductance increases linearly from about 0.25 (2 L) to about 2L/cmH2O (6 L) (Herman, 2007)[pg. 539f]. Flow is at Reynolds number of about 1600, so mostly laminar. However, turbulence occurs because the walls are not smooth. Interestingly, the work to breathe can take up to 20% of total body energy consumption.

Partitioning of FOT signals has been pioneered by DuBois et al. (1956a). In this model, transfer and input impedance are partitioned as follows,

$$Z_{tr} = Z_{aw} + Z_t + \frac{Z_{aw}Z_t}{Z_g},$$

$$Z_{in} = Z_{aw} + \frac{Z_gZ_t}{Z_g + Z_t}$$

where (t = tissue): $Z_t = R_t + i2\pi f L_t - iE_t/(2\pi f)$ and $Z_g = -iE_g/(2\pi f)$ (g = gas, compressibility).

The frequency-dependence of FOT parameters is modelled in the constant-phase model (Hantos et al., 1992; Peták et al., 1997). Thereby, $Z_{in}$ is separated into airway and tissue components, since $R_t = G/(2\pi f)^\alpha$ with a frequency dependence parameter $\alpha$. Suki et al. (1997) considered tissue nonlinearities in the context of this model.

At frequencies below 2 Hz mainly rheologic properties of the tissues are dominant, as well as mechanical heterogeneities. At frequencies above 100 Hz FOT obtains information about acoustic properties.

Ionescu and Keyser (2008) review other commonly used partitioning models. Resistance, compliance and inertance can also be considered in terms of electrical analogs and the math-
The mathematical theory of the resulting equivalent electrical circuits has been considered quite generally by Smale (1972).

The recent book by Bates (2009) is a general introduction to the modeling of lung mechanics. Viscoelasticity can be incorporated into the simple model (3.7) by an additional quadratic term,

\[ P_A(t) - P_0 = R_{rs} \dot{V}(t) + E_1 V(t) + E_2 V^2(t). \]

Including an inertance term, this becomes a second-order nonlinear model,

\[ P_A(t) - P_0 = R_{rs} \dot{V}(t) + E_1 V(t) + E_2 V^2(t) + L_{rs} \ddot{V}(t). \]

The inertial pressure is in counterphase with respect to elastic recoil pressure under harmonic forcing and thereby compensates the stiffness of the respiratory system. It becomes dominant over elastance at frequencies greater than 5 to 10 Hz (Peslin and Fredberg, 1986).

Similarly can the linear resistance term be replaced by a flow-dependent resistance (Rohrer, 1915; Wagers et al., 2002). In an actively breathing patient an estimate of pleural pressure \( P_{pl} \) is needed to partition between lung and chest wall characteristics. This can be measured approximately by oesophageal pressure (minimally invasive measurement by means of an oesophageal balloon). Another possibility to measure FOT signals is to use the pressure forcing generated by the heart, leading to so-called output impedance (Bijaoui et al., 2001). The problematic upper airways shunt has been studied by (Cauberghs and de Woestijne, 1989) and Bijaoui et al. (1999) discuss how to detect it and estimate impedance in its presence.

In recent years modern multifrequency FOT measurements, e.g., by impulse oscillometry (IOS) have become possible, but necessitate the use of complex, short pressure perturbations (Kuhnle et al., 2000; Klein et al., 80).

**Section 3.3**

Both asthma and COPD are obstructive lung diseases; under this heading fall also emphysema and chronic bronchitis (excessive mucus production). Narrowing of airways occurs in asthma, due to edema (thickening of airway walls or muscle hypertrophy), which reduces wall springiness and increases compliance.

Prevalence of COPD has been modeled by (Hoogendoorn et al., 2005). Influential projections of disease burden were published in (Murray and Lopez, 1997) and extended in (Mathers and Loncar, 2006).

Clinical application of FOT measurements is reviewed by (Goldman, 2001). LaPrad and Lutchen (2008) is a recent review of FOT with a focus on applications in asthma. (Lutchen et al., 1996) consider disease-related changes in FOT signals and criticize (the use of time averages of) single-frequency FOT in severely diseased lungs. Increased variability of \( Z_{rs} \) in asthma was reported by Que et al. (2001), but could not be confirmed later (Diba et al., 2007).

The book by Hamid et al. (2005) covers many further aspects of the physiology of healthy and diseased lungs.

Variability and fluctuation analysis is reviewed in Seely and Macklem (2004). A critical assessment of detrended fluctuation analysis was given by Maraun et al. (2004).
Section 3.5

The idea of “dynamical disease” was popularized by Glass and Mackey (1988).

Section 3.7

A further method that can be considered to highlight differences in FOT time series is bispectrum analysis (Mendel, 1991), which was pioneered in the analysis of EEG data (Barnett et al., 1971). Global dynamical models are reviewed in the recent article of Aguirre and Letellier (2009).

Approaches to detect artifacts in FOT measurements are described by Marchal et al. (2004), who also published considerations specific to measurements in young children (Marchal et al., 2005). Filtering to improve FOT estimates was discussed by Schweitzer et al. (2003).

Fluctuations in the respiratory system are discussed by Suki (2002). The model of Venegas et al. (2007) shows that exacerbations in asthma might be the result of a self-organizing cascade of ventilatory breakdowns.

The work of Bailie et al. (2009); Sassi et al. (2009) shows that heartbeat interinterval series are nonchaotic and multifractal. Wessel et al. (2009) argue that this might be caused by respiratory coupling.

Section 4.1

Since the publication of the first magnetoresonance (MR) image (Lauterbur, 1973), MR imaging has become one of the most important medical imaging methods. The mathematics behind image reconstruction in MRI is described by Twieg (1983); Ljunggren (1983).

Section 4.2

An important alternative to histogram estimates is kernel density estimation (Silverman, 1986), which results in reduced bias away from the interval mid-points. Silverman (1981) described one approaches to bump-hunting, i.e., estimation of the location of the peak of a density. Classification by likelihood ratios in this context has been considered by Silverman (1978). The choice of the bandwidth is still an issue, however. Adaptive bandwidth selection overcomes many of the problems with a single bandwidth (Sain and Scott, 1996; Sain, 2002). Kernel density estimation also offers the possibility to estimate total variation distances. In a different context this has been discussed by Schmid and Schmidt (2006). Since densities are infinite-dimensional objects, multivariate analysis of densities (discriminant analysis, PCA) needs to be based on a finite approximation. The distance-based approach is a natural way to avoid the bias and instability of histograms. A similar, popular approach is offered by kernel discriminant analysis (Shawe-Taylor and Cristianini, 2004).

Section 4.3

Yet another approach to quantitative MRI is offered by finite mixture models of parameter distributions (McLachlan and Peel, 2000; Wehrens et al., 2002), that can be fitted in the Bayesian
Figure D.1: Example: Fitting a two-component Gaussian mixture to an empirical magnetic transfer ratio distribution.

framework (Fraley and Raftery, 2002). Although computationally involved, this approach allows to work with substantial hypotheses and is a promising direction for future research. Figure D.1 shows an example obtained with the MCLUST software (Fraley and Raftery, 2007). For a true application, the fitting procedure needs to be extended to three-way analysis: One component should describe the background (brain) MTR distribution common to all subjects, whereas the remaining components should describe individual variations in the MTR parameter. This necessitates constrained and hierarchical fitting procedures which are not available at the moment.

The recent work of Oh and Raftery (2007) considers Bayesian clustering of Euclidean representations, and is similar to the distance-based analysis in its philosophy.

Section 4.4

The literature on Alzheimer’s Disease is too numerous to review here. Good starting points are Goedert and Spillantini (2006); Roberson and Mucke (2006). Our publication (Musculus, Scheenstra, Braakman, Dijkstra, Verduyn-Lunel, Alia, de Groot and Reiber, 2009) offers a comprehensive review of small animal models of Alzheimer’s disease and mathematical approaches to its genesis and disease severity quantification. Voxel-based relaxometry (Pell et al., 2004) has gained much interest in recent years and is based on estimating significance probabilities for single voxels in a group-wise comparison. Note that the $T_2$ parameter should be considered multi-exponential (Whittall et al., 1999), but a large (time-consuming) number of echo times is needed to resolve this sensibly and uniquely. Further details can be found in (Whittall et al., 1991).

Section 6.1

A comprehensive introduction into neuronal science is (Kandel et al., 2000). Magnetoencephalography is reviewed by Hämäläinen et al. (1993) and the resourceful book of Nunez...

Section 6.3

Functional connectivity of the brain and its relation to anatomical connectivity has been studied by Honey et al. (2007) by functional MRI. Partial coherence is a recent extension of coherence that has been applied in functional MRI to quantify directionality (Sun et al., 2004).

Synchronization approaches to brain dynamics model brain activity as self-sustained coupled non-linear oscillators. A general introduction to all aspects of synchronization is (Pikovsky et al., 2003). The Kuramoto model (Acebrón et al., 2005; Strogatz, 2000) has been very influential in this area (Cumin and Unsworth, 2007). In the time domain, Carmeli et al. (2005) introduced a bivariate synchronization measure that is based on geometrical changes (contractions) in delay embeddings.

A very different approach is considered in Albrecht and Palus (1991), where distance measures between power spectral densities are studied.

Section 6.5

Resting state connectivity in general has been studied by Greicius et al. (2003) and by Beckmann et al. (2005) in functional MRI. Effective connectivity in the context of auditory information processing is discussed by Gonçalves et al. (2001), again for functional MRI.

Section 6.6

The standard reference for the electrophysiological inverse problem is Sarvas (1987). Apart from beamforming, an important alternative approach to source location has been pioneered by Pascal-Marqui (2002). The book by Kaipio and Somersalo (2004) discusses the statistical approach to inverse problems and features a section on MEG.

Section A.1

Alternatively, and in a certain sense dual to the set-theoretic foundation of mathematics, it is possible to base all of mathematics on the notion of functional relationships, i.e., to build mathematics, and thereby also distance geometry, from the notion of a category (MacLane, 1985). This approach is not considered here.
Distance geometry started with the works of Menger (Menger, 1928, 1930), and the account of Blumenthal (1953) is still the canonical reference.

**Section A.2**

Euclidean point representations derived from multidimensional scaling can often be interpreted in terms of substantial properties. Beals et al. (1968) was influential with regard to this interpretative approach, and a more recent discussion is given by Gati and Tversky (1982).

The embedding criterion was already established by Young and Householder (1938). The double centering operation is studied in a more abstract setting by Critchley (1988). Gower (1966) put multidimensional scaling on a solid theoretical foundation, stressing the difference between R-matrices (coefficients-of-association between pairs of characteristics) and Q-matrices (coefficients-of-association between pairs of samples).

An introduction to nonmetric scaling procedures is given by (Borg and Groenen, 2005). Recent local embedding algorithms allow to accommodate geodesic distances, confer (Roweis and Saul, 2000; Tenenbaum et al., 2000).

For reconstruction of noisy distances, the method of Singer (2008) is available. More generally, the bound smoothing approach of Havel et al. (1983) allows to find optimal representations if lower and upper bounds of all pairwise distances are given. Thereby, the bound smoothing algorithm of Dress and Havel (1988) allows to respect the triangle inequality constraint optimally. In three dimensions, errors and missing entries were considered by Berger et al. (1999), and general distance matrix completion was studied by Trosset (2000).

Large scaling problems can potentially be solved by a divide and join strategy as described by Tzeng et al. (2008).

The interesting article of Laub et al. (2006) considers whether there is a use in classificatory task for negative eigendirections, arising from non-Euclidean pair-wise data.

Multidimensional scaling is essentially based on a Gram matrix (obtained by double centering), so it can be considered in the framework of kernel methods (Shawe-Taylor and Cristianini, 2004). This allows to apply the methods of kernel discriminant analysis, in particular, the recently developed support-vector machine classification strategies (Schölkopf and Smola, 2001; Hastie et al., 2008). However, these methods usually need much larger sample sizes to outperform linear discriminant analysis.

**Section A.3**

The multiple response permutation test can be applied in a much more general setting; these developments are described by Mielke and Berry (2007).

Although our distance-based approach was developed independently, generalized discriminant analysis on distances was first considered by Anderson and Robinson (2003). The solution of the out-of-sample problem, necessary for leave-one-out cross-validation, was obtained by Trosset and Priebe (2008). Confer (de Leeuw and Meulman, 1986) for a more general solution.

A recent issue of the Journal of the ICRU has been devoted entirely to receiver-operator characteristics (Receiver operating characteristic analysis in medical imaging, 2008). An advanced
measure of diagnostic accuracy in a probabilistic setting is the Brier score (Spiegelhalter, 1986; Redelmeier et al., 1991).

A seminal article with regard to the combination of distinct classifiers was (Xu et al., 1992); it considers averaging, voting, Bayesian methods, and Dempster-Shafer theory.

Section B.2

The optimal transportation problem was considered by Kantorovich (Kantorovich, 1942b, 1948). The discrete version is variously called the Hitchcock transportation problem (Hitchcock, 1941). In 1D it is efficiently solved by monotone rearrangement (Villani, 2003; Brandt et al. (1991) contains a derivation in the discrete case, where this simple fact is proved by duality (!). For finite points distributed on the circle, see (Cabrelli and Molter, 1995).

The “dictionary of distances” (Deza and Deza, 2006) discusses many other distance measures.

Section B.3

The Wasserstein distance is also called the Hutchinson metric (Hutchinson, 1981). In the image analysis literature it is often referred to as the “Earth mover’s distance” (EMD). Its metric properties and the lower bound by the difference in means is given by Rubner et al. (2000) in the discrete case. Note that the EMD is often defined more generally, namely, for two positive measures. Its solution is then given by the optimal transport that matches the smaller measure optimally to the larger. This construction does not result in a distance, however. The behavior of the Earth mover’s distance under transformations is considered by Klein and Velkamp (2005).

The Wasserstein distance has applications as a goodness of fit test in statistics (del Barrio et al., 1999).

An elementary proof of the triangle inequality was recently given by Clement and Desch (2008).

Hoffman (1963) showed that the optimal transportation problem can be solved in linear time $O(m + n)$ if the cost coefficients fulfill the Monge property,

$$c_{ij} + c_{rs} \leq c_{is} + c_{rj} \quad \text{for} \quad 1 \leq i < r \leq m, 1 \leq j < s \leq n.$$ 

This holds, for example, if $c_{ij} = u_i + v_j$. See the review by Burkard et al. (1996). This property can be generalized to Monge sequences (Alon et al., 1989), and recognition of permuted Monge matrices seems possible in $(O(mn + m \log m)$ time (confer Burkard et al. (1996) for references). It is related to the so-called quadrangle inequality that allows significant computational speedups in many algorithms (Yao, 1980).

A feasible solution is always available by the greedy northwest corner rule (Burkard, 2007). The fastest algorithm for solving minimum cost flows still seems to be Orlin’s algorithm, with a running time of

$$O(n' \log n' (n' \log n' + m')),$$

where $m' = n + m$ and $m'$ is the number of (finite) entries in the cost matrix (Orlin, 1988). The network simplex algorithm is described in detail by Kelly and O’Neill (1991).
Dell’Amico and Toth (2000) compare computational codes for the assignment code.

The Hungarian algorithm (Kuhn, 1955) runs with complexity $O(n^3)$ (Jonker and Volgenant, 1986). In the plane, (Vaidya, 1989; Atkinson and Vaidya, 1995) have obtained improvements to $O(n^{2.5} \log n)$ for the Euclidean bipartite matching problem and $O(n^{2.5} \log n \log N)$ for the transportation problem, where $N$ is the magnitude of the largest cost. Agarwal et al. (1999) consider an $\epsilon$-approximation algorithm with complexity $O(n^{2+\epsilon})$. A probabilistic algorithm that results in an $1+\epsilon$ approximation with probability at least $1/2$ has been given by Varadarajan and Agarwal (1999); it has complexity $O((n/\epsilon^3) \log^6 n)$. Kaijser (1998) considers another improvement in the plane.

If many related transportation problems need to be solved, relaxation algorithms should be considered, e.g., the Auction algorithm of Bertsekas and Castanon (1989).