NMR measurements for hazelnuts classification

Domenico Di Caro

Unione Europea

UNIVERSITÀ DEGLI
STUDI DI SALERNO

FONDO SOCIALE EUROPEO

Programma Operativo Nazionale 2000/2006
"Ricerca Scientifica, Sviluppo Tecnologico, Alta Formazione"
Regioni dell'Obiettivo 1 – Misura III.4
"Formazione superiore ed universitaria"

Department of Industrial Engineering

Ph.D. Course in Industrial Engineering
(XVI Cycle-New Series, XXX Cycle)

NMR measurements for hazelnuts classification

Supervisor
Prof. Consolatina Liguori

Ph.D. student
Domenico Di Caro

Ph.D. Course Coordinator
Prof. Ernesto Reverchon

To Adi

Publications resulting from this work

Di Caro, D., Liguori, C., Pietrosanto, A. and Sommella, P., 2016, May. Investigation of the NMR techniques to detect hidden defects in hazelnuts. In *Instrumentation and Measurement Technology Conference Proceedings (I2MTC), 2016 IEEE International* (pp. 1-5). IEEE.

Di Caro, D.; Liguori, C.; Pietrosanto, A.; Sommella, P.; Gibboni, A. (2016). UTILIZZO DI TECNICHE DI RISONANZA MAGNETICA PER IL RILIEVO DI DIFETTI OCCULTI NELLE NOCCIOLE. In: Atti del XXXIII Congresso Nazionale dell'Associazione GRUPPO MISURE ELETTRICHE ED ELETTRONICHE Benevento 19-21 settembre 2016 Vol.1, Pag.31-32

Di Caro, D., Liguori, C., Pietrosanto, A. and Sommella, P., 2017. Hazelnut Oil Classification by NMR Techniques. *IEEE Transactions on Instrumentation and Measurement*, *66*(5), pp.928-934.

Di Caro, D., Liguori, C., Pietrosanto, A. and Sommella, P., 2017, May. Using a SVD-based algorithm for T 2 spectrum calculation in TD-NMR application to detect hidden defects in hazelnuts. In *Instrumentation and Measurement Technology Conference (I2MTC), 2017 IEEE International* (pp. 1-6). IEEE.

Di Caro, D.; Liguori, C.; Pietrosanto, A.; Sommella, P.; Gibboni, A. (2017). CLASSIFICATION OF HAZELNUTS IN SHELL BY MEANS OF TD-NMR TECHNIQUES. In: Atti del I Forum Nazionale delle Misure, Modena 14-16 Settembre 2017, Pag.125-126

Di Caro, D., Liguori, C., Pietrosanto, A. and Sommella, P., 2018, May. Advanced TD-NMR based estimation of the inshell hazelnuts quality parameters. In *Instrumentation and Measurement Technology Conference (I2MTC), 2018 IEEE International* (pp. 1-6). IEEE.

Contents

List of figures

List of tables

Abstract

In this work, a method for the quality detection of the in-shell hazelnuts, based on the low field NMR, has been proposed. The aim of the work is to develop an in-line classification system able to detect the hidden defects of the hazelnuts. After an analysis of the hazelnut oil, carried out in order to verify the applicability of the NMR techniques and to determine some configuration parameters, the influence factors that affect these measurements in presence of solid sample instead of liquids have been analyzed. Then, the measurement algorithms were defined.

The proposed classification procedure is based on the CPMG sequence and the analysis of the transverse relaxation decay. The procedure includes three different steps in which different features are detected: moisture content, kernel development and mold development. These quality parameters have been evaluated .analyzing the maximum amplitude and the second echo peak of the CPMG signal, and the T_2 distribution of the relaxation decay. In order to assure high repeatability and low execution time, special attention has been put in the definition of the data processing. Finally, the realized measurement system has been characterized in terms of classification performance. In this phase, because of the reduced size of the test sample (especially for the hazelnuts with defects) a resampling method, the bootstrap, was used.

Introduction

In recent years, quality control in agricultural and food products is becoming a key factor for several factors, like increasing safety and customer satisfaction, and reducing financial losses due to the storage and selling of low quality products. Several standards, related to the processes and products, have been adopted to define the minimum requirements that must be satisfied by companies. In order to check if the quality requirements are met, the firms use laboratory equipment to detect the main properties of their products. This kind of instrumentation guarantees reliability and precision, but it is expensive and requires high skilled personnel. For these reasons, farmers and food industries also require non-destructive techniques for in-line evaluation of the quality of their products. Several techniques, like X-rays, infrared spectroscopy, ultraviolet-visible (UV-VIS) spectroscopy, have been investigated to perform quality analysis on agricultural and food products. Among them, Nuclear Magnetic Resonance (NMR) is increasing its usability, showing to be an effective detection method in the food quality control. In this work, Time Domain – NMR (TD-NMR) has been employed for the development of a system for the in-line classification of the hazelnuts in shell. Several techniques have been adopted to check the quality requirements related to the hazelnuts. The dimension of the shell and the blank nuts are easily detected using machines based on mechanical techniques (e.g. compressed air). They are not enough accurate but characterized by low cost and high throughput. A different method to detect empty hazelnuts, based on the analysis of the acoustic signal generated from the impact of the nut on a steel plate, has also been proposed. Using the same principle, a sorting method of cracked, hollow and regular shell has been developed for hazelnuts and pistachio nuts, exploiting Frequency Domain (FD) signal processing techniques and Artificial Neural Networks (ANNs) for the

classification. Moreover, RF impedance measurement has been employed to determine the moisture in a batch of in-shell hazelnuts. These techniques provide classification methods based on single features or they are not suitable for in-line detection.

For many years, NMR techniques were used in chemical analysis and medical diagnosis, but recently, thanks to the development of permanent magnets and the improvements in the signal processing, several industrial applications have been investigated. In particular, the NMR methods based on time domain analysis proved to be suitable for food quality detection related to fat, oil and water components. The quality of the hazelnuts is mainly determined by the moisture content and the quantity and quality of their main component, which is the oleic acid. For these reasons, the use of the TD-NMR appeared to be appropriate for the analysis of this kind of product. In a first phase of the work, the tests have been conducted on oil samples extracted from healthy and unhealthy hazelnuts. This allowed focusing the analysis on the composition of the material without considering the effect of external influence factors like moisture or the shape of the sample. Thanks to these tests, the optimal parameters for the system setup and the pre-processing algorithms were defined. Moreover, an algorithm for the classification of the hazelnut oils, based on Singular Value Decomposition (SVD), has been proposed. Later, an analysis of the influence factors has been carried out and then a two-stage classification method for the in-shell hazelnuts has been developed. It is based on the analysis of the transverse relaxation decay obtained from the CPMG signal. In the first stage, the empty hazelnuts, with a not well developed kernel and with a high moisture content are detected exploiting the differences among the amplitude of the CPMG signal, while in the second stage, molded hazelnuts are detected analyzing the differences among the components of the multi-exponential decays of the transverse relaxation signal. The thesis is organized as follows: Chapter I describes the quality control parameters involved in the hazelnuts processing and the main techniques used in the food quality control, focusing on the NMR applications. Chapter II shows an overview of the NMR techniques and a description of the system. Chapter III contains a description of the influence parameters and the data processing and test oil samples. Chapter IV describes the signal processing techniques employed in the analysis of the hazelnuts and Chapter V shows the classification algorithm for the in-shell hazelnuts.

Chapter I
Quality control in hazelnuts processing

I.1 Main properties and hidden defects of the hazelnuts

Hazelnuts are widely used in the confectionary industry for their flavor and taste. They are characterized by a high nutritional value due to the presence of several components, mainly lipids (about 60% w/w), carbohydrates, proteins, sugar and dietary fibers (Memoli *et al.*, 2017). Turkey is the most important producer of hazelnuts in the world, with more than 70% of the overall production, followed by Italy and United States (FAOSTAT, 2017). Both the yearly world production and the hazelnuts quality depend on the climate condition. In order to guarantee a suitable level of quality, several standards have been defined by the producing countries and regions. They define the minimum quality requirements of the fruits in terms of dimension, aspect, hidden defects. Other requirements, related to the quality of the hazelnuts and the place of production, can be found, for example, in the European Union certification like the Protected Geographical Indication (PGI). The most recent international standard has been set by the Organization for Economic Co-operation and Development (OECD, 2011). It defines the quality requirements for the in-shell hazelnuts and hazelnut kernels intended for direct consumption or for food. The minimum requirements are described in the following. The shell must be:

- Intact; only slight superficial damages are allowed.
- Clean; free of any foreign matter.
- Free from blemishes and stains affecting more than 25% of the surface of the shell.
- Well formed.

The kernel must be:

- Free from rancidity.

- Sufficiently developed; kernels should fill at least 50% of the shell cavity.
- Not desiccated.
- Free from blemishes and stains affecting more than 25% of the surface of the shell.
- Well formed.

The whole nut must be:

- Free from mold filaments.
- Free from living pests.
- Free from damage caused by pests.
- Free of any foreign smell or taste.
- Dried; moisture content not greater than 12% for the whole nut or 7% for the kernel.

The standard also defines the maximum allowed defects related to each requirements providing three quality classes: extra class, class I and class II. Figure I-1 shows the aspect of healthy and unhealthy kernels.

Figure I-1 *Hazelnut kernels: (1) mold and pest damage; (2) desiccated and not well developed; (3) healthy*

There are many varieties of hazelnuts cultivated in several areas. Each cultivar has specific characteristics and can produce fruits with different properties. Among the cultivars, “Tonda di Giffoni”, that is produced in the south of Italy by the “Consorzio di Tutela Nocciola di Giffoni I.G.P., is appreciated for its quality which is particularly suitable for the processing industry. It is certified by the European Union with the Protected Geographical

Indication (PGI) that defines the following features: dimension of the shell not less than 18 mm; dimension of the kernel of the unshelled hazelnuts not less than 13 mm; moisture content after drying not more than 6% (Commission Regulation (EC) No 1257/2006 of 21 August 2006).

The price of the hazelnuts it is strictly related to the quality. In particular, it is determined on the basis of the percentage of the defects found in a batch. For this reason, farmers need nondestructive techniques to detect the hidden defects of the hazelnuts in order to avoid a price reduction selling the product.

I.2 Techniques to detect hazelnuts quality

The main technique used to detect the hidden defects of the hazelnuts is the visual inspection. This is a sample based method in which a batch of at least 100 nuts are examined and classified. A procedure to carry out this kind of test has been defined by the U.S. Food and Drug Administration in the Macroanalytical Procedures Manual (MPM) (FDA, 1984). The document defines the test procedures for several food products, like fruits, vegetables, grain, dairy, cheese and seafood. Nut products method is described in the chapter V-10. It defines:

- Sample preparation.
- Sequential sampling plans.
- Visual and organoleptic examination.
- Classification of reject nuts.
- Report.

This method requires time and skilled personnel for a correct nut classification. Other techniques have been investigated in order to achieve automatic systems to check the quality parameters of the hazelnuts.

The dimension of the shell and the blank nuts are detected by means of machines based on mechanical techniques. In particular, calibrator machines able to sort the whole hazelnuts on the basis of their diameter are available making use of sieves with different diameters. Air compressed-based machines can detect and separate some kinds of foreign body, like leaves or pieces of wood, from the whole hazelnuts. They can also detect blank nuts, even if they are not accurate. These kinds of equipment are useful for a pretreatment of the product but they does not represent quality control system. A different method to detect empty hazelnuts, based on the analysis of the acoustic signal generated by the impact of the nut on a steel plate, has been proposed (Onaran *et al.*, 2005). It exploits both time-domain and frequency-domain analysis of the signal and can be used for a real-time detection. Using the same principle, a sorting method of cracked, hollow and regular shell has been developed for hazelnuts (Kalkan and Çetisli, 2011) and pistachio nuts (Mahdavi-Jafari *et al.*, 2008), exploiting Frequency Domain (FD) signal processing techniques and Artificial Neural Network (ANNs). Several studies have been conducted related to the moisture content determination. This is an

important quality parameter because a high level of moisture causes fungal diseases and mold formation, and then postharvest losses. Based on the weather condition around the harvest period, the moisture content of the hazelnuts can be more than 30%. They are dried using industrial dryer or, often, by means of sun drying techniques. After this process, farmers need instruments to verify that the moisture content is lower than the limits imposed by the quality standards. The most common method to evaluate the moisture content in food is based on oven drying in which the sample is heated under specified conditions and the amount of moisture is determined calculating the loss of weight:

$$\%Moisture = \frac{weight\ wet\ sample - weight\ dry\ sample}{weight\ wet\ sample} \times 100 \qquad (1)$$

Other methods, related in particular to seeds and nuts, have been proposed. RF impedance measurement has been employed to determine the moisture in single grain and peanut kernels placed in a small parallel plate capacitor (Nelson *et al.*, 1992). The samples are placed between two electrodes and a measurement of the impedance is made at two different frequencies. The difference of the measured impedances is related to the moisture content. A similar method has been applied to determine the amount of moisture in in-shell hazelnuts placed in a probe with two vertical parallel-plate electrodes able to host a sample of 250 g of nuts (Solar and Solar, 2016). More complex analytical techniques have been used to evaluate some quality characteristics of the hazelnuts. For example, a study for the identification of chemical markers of hazelnut roasting, based on Headspace Solid Phase Microextraction (HS-SPME) coupled with Gas Chromatography – Mass Spectrometric (GC-MS) detection has been proposed (Nicolotti *et al.*, 2013). Near Infrared (NIR) Spectroscopy has also been used to detect flawed kernels and to estimate lipid oxidation (Pannico *et al.*, 2015). These kinds of techniques allow extracting complex features in a non-destructive way, but are suitable for laboratory testing and cannot used for industrial applications.

I.3 Nondestructive techniques in food quality control

In recent years, the quality control in both processes and products is becoming an important issue in the food supply chain. In order to guarantee the compliance to the international standards and regulations, farmers and food industries need nondestructive techniques to check the quality of their products. Nondestructive techniques can be defined as the qualitative and quantitative measurements that has been surveyed without any chemical, physical, mechanical and thermal damage to the product (Aboonajmi and Faridi, 2016). Several studies have been conducted using different techniques to carry out quality control on agricultural and food products at different stages, from the plantation, to the post-harvesting and then the processing.

Among these techniques, visual spectroscopy exploits the light energy absorption at specific wavelengths, in the visible region (380-750 nm), of the chemical components of the food materials. In this way, some compositional information can be determined from the spectra of the samples under test. On the same principle, but in a different range of wavelength (780-2500 nm), is based the NIR spectroscopy. This kind of techniques are not reliable in case of heterogeneous materials, and they usually require a considerable calibration effort (Dalitz *et al.*, 2012). Microwave sensing is an emerging technique for internal quality detection in water-containing materials taking advantage of the dipolar nature of water and dispersion associated with its dielectric properties at microwave frequencies (Trabelsi and Nelson, 2016; Meng *et al.*, 2017). Several applications, based on both acoustics and ultrasound techniques, have been proposed (Aboonajmi and Faridi, 2016). They are based on the analysis of the waves, generated by a source, that pass through the product tissue, and they allow evaluating the freshness and the ripeness of some kinds of food. X-ray imaging is another method that has been applied in food inspection. Some applications to fresh agricultural products, using medical grade computed tomography, are described in (Donis-Gonzalez *et al.*, 2014). Moreover, a study for the quality assessment of the onions, using x-ray computed tomography, can be found in (Speir and Haidekker, 2017).

Among the nondestructive techniques, NMR is increasing its usability in food quality control. In the next paragraph an overview of the food applications of this technique will be presented.

I.4 Using NMR in food quality control

NMR has been mainly used, in the past, in chemical analysis and medical diagnosis. The applications in food quality control have been developed progressively, thanks to the technical improvement in the instrumentation, the signal processing and especially the development of permanent magnets that allowed a significant cost reduction of the NMR equipment.

The NMR systems can be classified, on the basis of the technique used to carry out the analysis, in three main categories:

- NMR Spectrometry.
- NMR Relaxometry.
- MRI (Magnetic Resonance Imaging).

NMR Spectrometry, also named High Field NMR or Fourier Transform NMR (FT-NMR), makes use of superconducting magnets with a static magnetic field usually higher than 1 T. It is based on the frequency domain analysis of the NMR signals and it allows obtaining quantitative information related to the chemical composition of the samples under test.

Relaxometry mainly differs from the spectrometry for the use of permanent magnets that makes it the cheapest NMR technique and, for this reason, the most interesting to investigate industrial applications. It is also named Low

Field NMR or Time Domain NMR (TD-NMR) because it employs magnets with a static magnetic field that is usually lower than 1 T and the signal analysis, differently from the spectrometry, is carried out in the time domain.

MRI is a technique highly used in medical diagnosis because it allows achieving images of the organs of the body. It is more complex than the previous techniques because it makes use of gradient magnetic fields in addition to the main static magnetic field to produce the images.

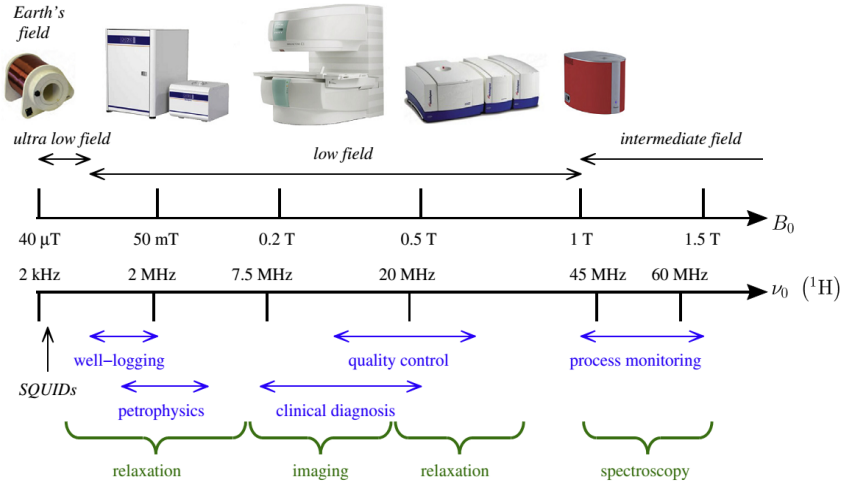

Figure I-2 *Summary of the NMR applications on the basis of the static magnetic field strength for permanent magnets (Source: Mitchell et al., 2013)*

All of the above techniques have been investigated in food analysis. An overview of the food applications of the liquid state FT-NMR can be found in (Mannina *et al.*, 2012). It presents a review of the NMR methodologies in food quality control, analyzing several aspects as the sample preparation, which is an important step to obtain reproducible spectra, the spectral analysis and the statistical analysis. It also describes some applications to several products like alcoholic beverage, fruits and vegetables, milk and dairy products. An example of food authentication is described in (Parker *et al.*, 2014), in which a bench-top spectrometer has been used for the analysis of olive oil with adulteration of hazelnut oil. A study on the compositional analysis of wine in a full intact bottle has been proposed in (Weekley *et al.*, 2003). In the paper, the capability to detect acetic acid spoilage has been investigated using NMR spectrometry system. An overview of the potential applications of NMR and MRI for compositional and structural analysis, inspection of microbiological, physical and chemical quality, food authentication, on-line monitoring of food processing is presented in (Marcone *et al.*, 2013). A review focused on the

development of TD-NMR and MRI based sensors for different food applications, can be found in (Kirtil *et al.*, 2017).

Several studies have been carried out on methods and applications of Low Field NMR. A method for Solid Fat Content (SFC) and simultaneous oil and moisture determination has been described in (Todt *et al.*, 2006). Applications on quality control of different food products have been proposed: the evaluation of quality of oranges during storage (Zhang *et al.*, 2012); Analysis of water dynamic states and age-related changes in mozzarella cheese (Gianferri *et al.*, 2007). The interest towards the industrial applications of the TD-NMR is constantly increasing. A study of the perspective of the use of the TD-NMR as in-line industrial sensor can be found in (Colnago *et al.*, 2014). Among the proposed applications in in-line monitoring, Low Field NMR has been employed for the evaluation of the internal browning of whole apples (Chayaprasert and Stroshine, 2005). Moreover, an automated system for oil and water content determination in corn seeds has been proposed in (Wang *et al.*, 2016).

The main applications of the TD-NMR in food quality control are related to the evaluation of the presence of water, moisture, oil, fat, that are well detected by this technique. The quality of the hazelnuts is strictly related to the presence of moisture and fatty acids and then, in this work, Low Field NMR has been taken into account as a suitable technique for the analysis of the hazelnuts quality. In the next chapter, an overview of the NMR methods and the equipment used to carry out the analysis will be depicted.

Chapter I

Chapter II
NMR methods and equipment

II.1 NMR Theory

NMR is based on the magnetic properties of the nuclei of the atoms. The nuclei are characterized by a nuclear spin quantum number (I) that can be equal to or greater than zero, with values multiples of ½. Only the nuclei with a value of I different by zero are NMR-sensitive, while the nuclei with I=0, which are those with atomic mass and atomic number both even, are NMR-silent. The main characteristic of the spinning nuclei is that they possess an angular momentum (P) and charge, and the motion of the charge causes the presence of a magnetic moment (μ). The magnetic moment is proportional to the angular momentum:

$$\mu = \gamma \cdot P \qquad (2)$$

Where γ is the gyromagnetic ratio, which is constant for each nucleus.

When the nuclei are placed in an external static magnetic fields (B_0), the microscopic magnetic moments align themselves relative to the field in a discrete number of orientation, depending on the energy states involved (they depend on the possible spin states). The effect of the static magnetic field on the magnetic moment is to impose a torque that causes a motion of the vector representing the magnetic moment on the surface of a cone around the direction of the field (Fig. II-1).

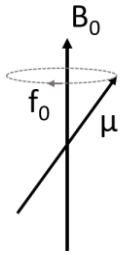

Figure II-1 *Motion of a magnetic moment (μ) of a nucleus in a static magnetic field (B_0)*

This motion is referred to as Larmor precession and its rate depends on the strength of the applied field and the gyromagnetic ratio of the nucleus.

The rate can be expressed in terms of angular velocity ω_0 (rad s^{-1}):

$$\omega_0 = -\gamma \cdot B_0 \quad (3)$$

Or frequency f_0 (Hz):

$$f_0 = -\frac{\gamma}{2\pi} B_0 \quad (4)$$

This is named Larmor frequency of the nucleus. For the proton (^{1}H) the gyromagnetic ratio (γ) is $26.75 \cdot 10^7$ rad $T^{-1}s^{-1}$. With a 1 T static magnetic field the larmor frequency is 42.25 MHz. In the table II.1, the gyromagnetic ratio, natural abundance and relative sensitivity (to the proton) for the main NMR-sensitive nuclei is shown.

Table II.1 *Natural abundance, gyromagnetic ratio and relative sensitivity (to ^{1}H) of some NMR-sensitive nuclei*

Nucleus	Natural abundance [%]	γ [rad $T^{-1}s^{-1}$]	Relative sensitivity [%]
^{1}H	99.98	$26.75 \cdot 10^7$	100.0
^{13}C	1.11	$6.73 \cdot 10^7$	1.6
^{19}F	100.00	$25.18 \cdot 10^7$	83.3
^{31}P	100.00	$10.84 \cdot 10^7$	6.6

The ^{1}H nucleus has the highest NMR sensitivity and can be found in most materials. For this reason it is the most used in the NMR experiments.

At a macroscopic level, the effect of the static magnetic field is a net magnetization vector (M_0) on the same direction of the applied field, which is conventionally considered the z-axis of a Cartesian co-ordinate system (Claridge, 2016). When the nuclei are placed in the field, the magnetization is not instantaneous, but it increases exponentially (eq. 5):

$$M_z(t) = M_0 \left(1 - e^{-\frac{t}{T_1}}\right) \quad (5)$$

The time constant T_1 is the longitudinal relaxation time. It is an important parameter because it can be used to get information about the sample under test during the NMR experiments. Moreover, when a sample is placed in the magnetic field, or after an experiment, it is necessary to wait a period of at least $5T_1$ to be sure that the magnetization has recovered. This period, on the basis of the material under test, can be several seconds.

When a material is placed in the static magnetic field and the magnetization process is complete, to obtain a NMR signal it is necessary a radiofrequency (RF) pulse at the Larmor frequency (which is the resonance frequency of the system) that generates an oscillating magnetic field B_1 in a plane orthogonal

to the direction of B_0. This can be made transmitting the RF pulse to a coil placed in the static field (Fig. II.2).

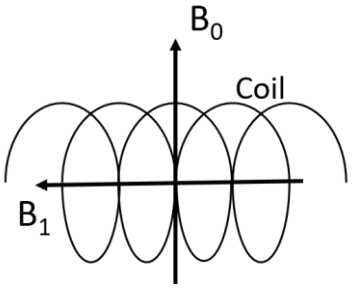

Figure II.2 *Oscillating magnetic field B_1 generated by a RF pulse transmitted to a coil placed orthogonally in the static magnetic field B_0*

The effect of B_1 is to rotate the magnetization vector, which is along the z-axis, by an angle θ, called flip angle that can be expressed as:

$$\theta = \gamma \cdot t_p \cdot B_1 \qquad (6)$$

Where t_p is the duration of the RF pulse. At the end of the RF pulse the magnetization vector returns to the equilibrium and, during this process, it induces a RF voltage in the coil. The maximum amplitude of the voltage is obtained when $\theta = \pi/2$. When the magnetization vector returns to the equilibrium, the component M_z increases according to eq. (5). This process is named longitudinal relaxation. At the same time, another independent process occurs, the transverse relaxation, in which the component of the magnetization in the x-y plane decreases to zero. This process is always faster than the longitudinal relaxation. It can be expressed by the eq. (7):

$$M_{xy}(t) = M_{xy} e^{-\frac{t}{T_2}} \qquad (7)$$

Where the time constant T_2 is the transverse relaxation time. This is, together with the longitudinal relaxation time (T_1), one of the main parameters used to get quality information about the material under test, in particular in the NMR relaxometry. Conversely, the NMR spectroscopy is based on the frequency domain analysis of the signal induced in the coil at the end of the RF pulse. In the next paragraph a description of the NMR instrument will be depicted.

II.2 NMR equipment

A NMR instrument is composed, regardless of the use as spectrometer or relaxometer, by four main components:

- Magnet;
- Probe;

- Transmitter;
- Receiver.

II.2.1 Magnet

The magnet is used to generate the static magnetic field B_0. The field must be extremely homogeneous in the volume that contains the sample under test, in order to have the same resonance frequency in the different sections of the sample. The most homogeneous field can be obtained with the superconducting magnets, which are composed by a coil with a wire of superconducting material immersed in liquid helium to maintain the coil at the low temperature needed to guarantee the superconducting property of the wire. These kind of magnets are mainly used in the applications that need high field with the highest homogeneity, like the NMR spectroscopy for chemical analysis. In the low field NMR, permanent magnets are used instead of the more expensive superconductive magnets. They are not able to produce a perfect homogeneous field as required in the NMR spectroscopy, but can be used in the Relaxometry, where several pulse sequences techniques have been developed to limit the effect of the field inhomogeneity and to obtain quality information about the samples under test. Using the permanent magnets allowed a great reduction of the overall cost of the systems and the emergence of the NMR industrial applications.

The main parameter of the magnet is the static magnetic field (B_0), which also determines the resonance frequency of the NMR system (see eq. 4). Other parameters are:

- Field homogeneity ($\Delta B_0/B_0$);
- Dimension of the field homogeneity area

II.2.2 Probe

The probe is the component that allows transmitting the excitation signals to the samples and receiving the NMR signal in response to the excitation. It is connected both to the transmitter and the receiver and its purpose is to deliver the rotating magnetic field B_1 to the sample and to detect the signal at the end of the excitation RF pulse, during the relaxation time. The main component of the probe is the coil (with inductance L) which is placed inside the magnet in order to obtain a B_1 field orthogonal to B_0 when the transmitter sends the RF pulses. The sample under test is arranged in the coil of the probe in order to be subjected both to the static and the rotating fields. The probe also includes two variable capacitors for the frequency tuning (C_{tune}) and the impedance matching (C_{match}) (Fig. II.3). The probe circuit is characterized by three quantities (Teng, 2012):

- Resonance frequency;
- Impedance at the resonance;

- Q factor.

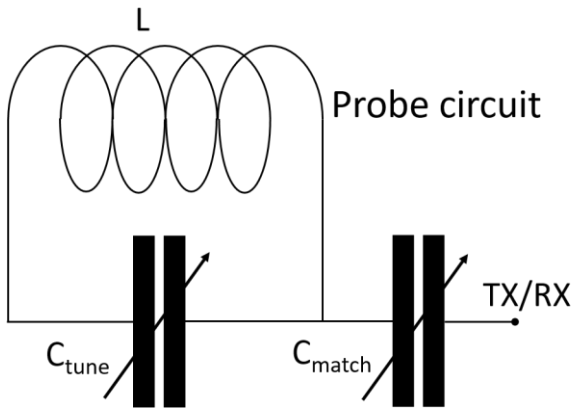

Figure II-3 *Example of probe circuit*

The two capacitors are used to adjust the value of the resonance frequency of the circuit to the resonance frequency of the NMR system, and to obtain, at that frequency, an impedance of 50 Ω. The Q factor depends on the resonance frequency and it can be defined by:

$$Q = \frac{f_0}{\Delta f} \qquad (8)$$

Where f_0 is the resonance frequency and Δf is the bandwidth of the resonance circuit. In order to increase the probe sensitivity, the probe design is made to obtain the highest value of Q.

II.2.3 Transmitter

The transmitter provides the RF pulses to send to the probe for the sample excitation. It has to be able to generate the RF signals with the requested frequency, phase, amplitude and pulse width. The signals have to be amplified in a power range that depends on the NMR application, and it can be from 10 W to hundreds watt.

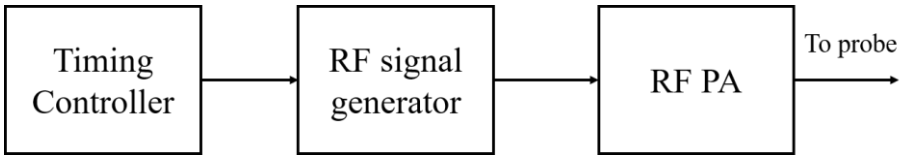

Figure II-4 *Block diagram of the transmitter*

Figure II-4 shows a block diagram of the NMR transmitter. It is composed by three diffeent blocks:

- A timing controller that manages the pulse width and the timing of the pulse sequences;
- A RF signal generator;
- A RF power amplifier.

The transmitted RF pulse is a sinusoidal signal at the resonance frequency of the system (f_0) with duration t_p. It can be expressed in the time domain as:

$$x(t) = A \prod \left(\frac{t}{t_p}\right) cos(2\pi f_0) \quad (9)$$

And in the frequency domain by its fourier transform:

$$X(f) = \frac{At_p}{2}\left[sinc(f - f_0)t_p + sinc(f + f_0)t_p\right] \quad (10)$$

The RF pulses excites the frequencies around the resonance frequency in a bandwidth that depends on the pulse duration t_p.

II.2.3 Receiver

The receiver has to be able to detect the weak RF signal from the probe. As previously described (Fig. II-3), the coil of the probe is used both for exciting and for detecting the signal. The excitation pulse is a power signal, so the receiver must be protected when the TX pulses are applied. To do this, the receiver is connected to the probe by means of a device called duplexer and a λ/4 cable (Fig. II.5). The duplexer is made of fast switching diodes. It allows, together with the λ/4 cable, protecting the receiver:

- When the pulse is on, the high power RF is routed to the probe and the receiver is protected by disconnecting it or shorting it to ground;
- When the pulse is off, the receiver is connected to the probe and the transmitter is disconnected.

The received signal is amplified and then a quadrature detection is carried out to obtain a baseband NMR signal from the RF signal.

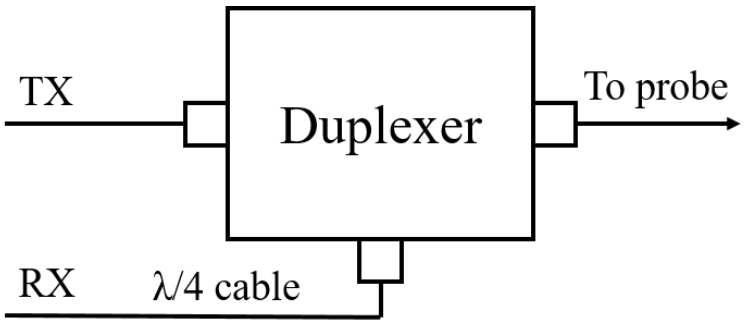

Figure II-5 *Transmitter-receiver connection to the duplexer*

The following figure shows a block diagram of the receiver:

Figure II-6 *Block diagram of the receiver*

The received signal has a low-level amplitude that has to be amplified using a low-noise preamplifier that is design to add the lowest possible noise level. In addition, a bandpass filter (BPF) is used in order to filter the noise outside the band of interest before the quadrature detector.

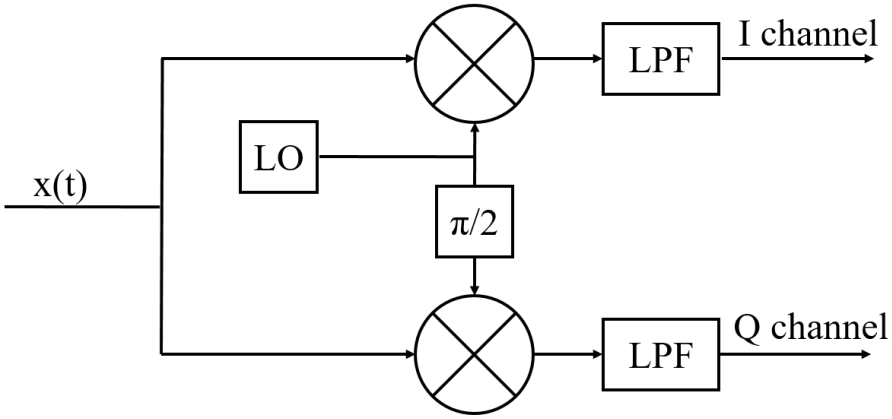

Figure II-7 *Block diagram of the quadrature detector*

The quadrature detector is composed by:

- Local oscillator (LO): it provides a sinusoidal signal with frequency f_{LO};
- $\pi/2$ phase shifter: it provides a sinusoidal signal with a $\pi/2$ phase shift respect to the signal generated by the LO;
- Mixer: it multiplies the input signal x(t) and the signal of the LO in the I channel, and the input signal and the phase shifted signal in the Q channel;
- Lowpass filter (LPF).

The signals obtained in the I and Q channel have the frequency: $f_0 - f_{LO}$, where f_0 is the frequency of the input signal x(t).

The input signal, received from the probe, can be written as:

$$x(t) = Acos(2\pi f_0 t) \quad (11)$$

Where f_0 is the Larmor frequency; while the signal generated by the LO and the $\pi/2$ phase shifter can be written, respectively:

$$x_{LO}(t) = cos(2\pi f_{LO} t) \quad (12)$$

$$x_{SH}(t) = -sin(2\pi f_{LO} t) \quad (13)$$

The output of the mixer for the I channel is:

$$x(t) \cdot x_{LO}(t) = Acos(2\pi f_0 t) \cdot cos(2\pi f_{LO} t) = \frac{A}{2}[cos2\pi(f_0 + f_{LO})t + cos2\pi(f_0 - f_{LO})t] \quad (14)$$

While the output for the Q channel is:

$$x(t) \cdot x_{SH}(t) = Acos(2\pi f_0 t) \cdot (-sin(2\pi f_{LO} t)) = \frac{A}{2}[-sin2\pi(f_0 + f_{LO})t + sin2\pi(f_0 - f_{LO})t] \quad (15)$$

As can be seen from eq. (14) and eq. (15), at the output of the mixer the signals can be expressed as the sum of two sinusoidal components with frequency f_0+f_{LO} and f_0-f_{LO}. The component with frequency f_0+f_{LO} is filtered using the LPF, and then only the component with frequency f_0-f_{LO} is present at the output of the quadrature detector.

The frequency of the local oscillator must be equal to the resonance frequency of the system ($f_{LO} = f_0$) in order to have a baseband NMR signal at the output of the receiver. This is useful because the NMR signals have a narrowband frequency spectrum around the resonance frequency, so the baseband conversion allows simplifying the signal processing.

II.3 NMR methods

Several techniques have been developed to obtain NMR signals in order to extract information like the frequency spectrum, in the case of FT-NMR, or the longitudinal (T_1) and transverse (T_2) relaxation time, the diffusion coefficient (D), in the case of TD-NMR. In this work, ^{1}H TD-NMR has been employed and two pulse sequences have been taken into account:

- Single pulse;
- CPMG.

These techniques does not require complex hardware and are relatively fast, so they are suitable for the industrial applications.

II.3.1 Single pulse sequence

Single pulse is the simplest sequence to obtain an NMR signal response. A short and strong RF pulse at the resonance frequency of the system is produced by the transmitter to excite the frequencies in the NMR spectrum; during the RF pulse, the magnetization of the nuclei inside the sample rotates, and the duration of the pulse is chosen to obtain a rotation of the magnetization vector of $\pi/2$ rad. At the end of the RF pulse, the nuclei of the sample return to the initial position generating an NMR signal (FID signal).

In order to acquire the FID signal correctly, it is necessary to synchronize the transmitter and the receiver by means of a control signal (Blanking signal).

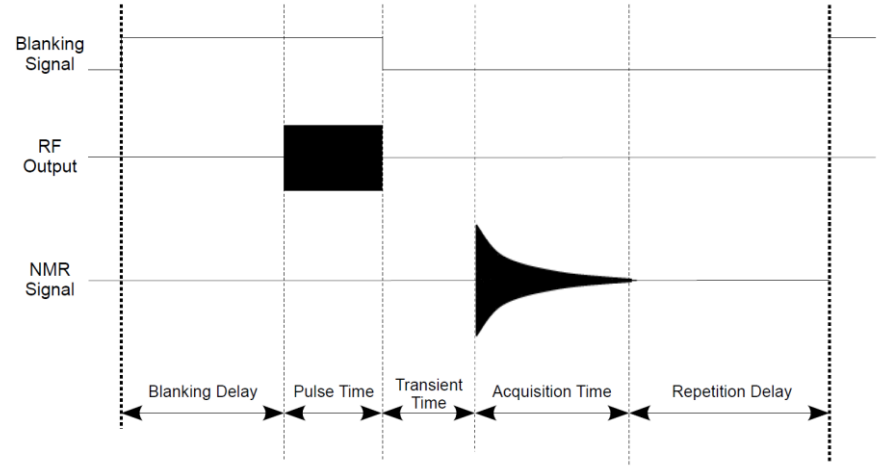

Figure II-8 *Time diagram of the single pulse sequence (Source: Spincore Technologies Inc., 2014)*

Figure II-8 shows a time diagram of the single pulse sequence. When the blanking signal becomes active, the RF power amplifier (RFPA) is switched on, and the RF pulse is transmitted to the probe after the blanking delay, that is the time needed to warm-up the RFPA before the RF pulse can start (Fig. II-9). After the pulse time (t_p) the blanking signal becomes de-active and the RFPA is switched off, but the receiver can start to acquire the NMR signal only after the transient time, also named dead time, which is the time needed to switch off the RFPA. After the acquisition time, in which the receiver acquires the FID signal, a new pulse sequence can be started on the same sample only after the repetition delay, which is the time needed to completely recover the magnetization of the sample (eq. 5).

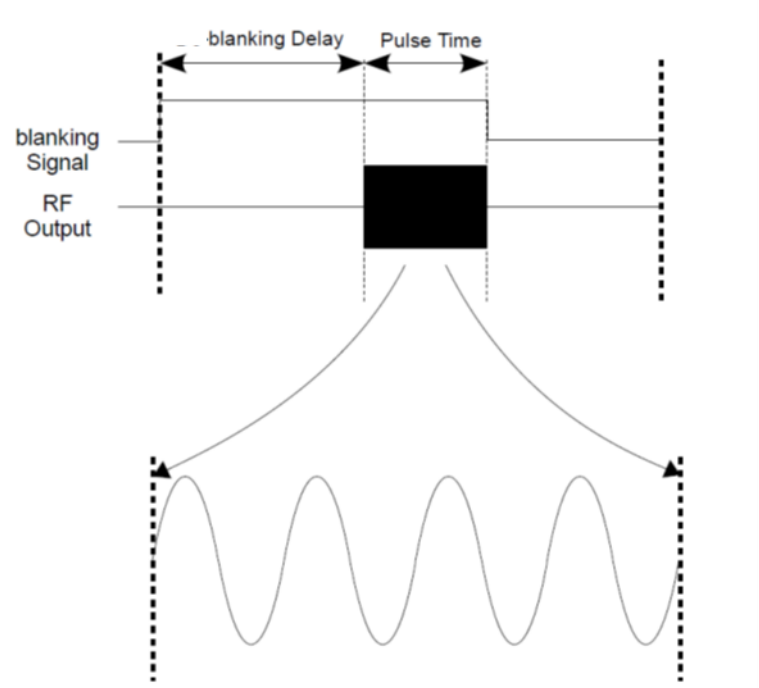

Figure II-9 *Blanking signal and RF pulse at the resonance frequency (Source: Spincore Technologies Inc., 2014)*

Both the blanking delay and the transient time are related to the RF power amplifier and have to be the shortest possible. In particular, the transient time does not allow acquiring all the FID signal (Fig. II-10).

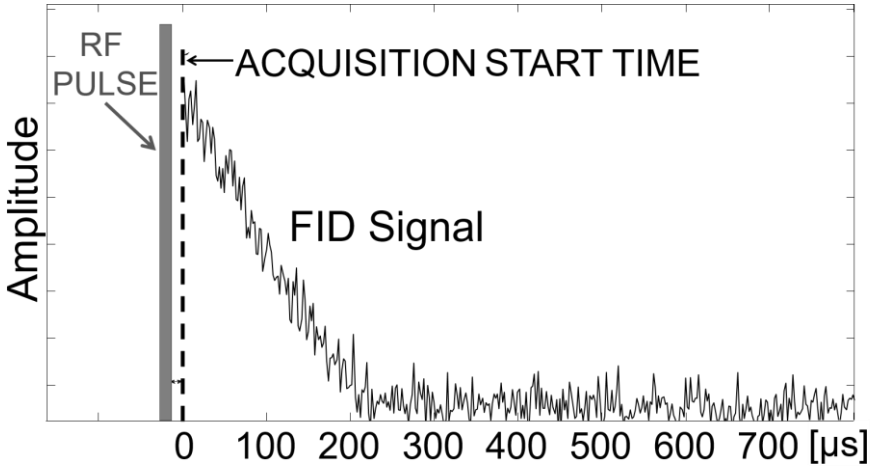

Figure II-10 *FID signal acquired by means of single pulse*

In a perfectly homogeneous field ($\Delta B/B_0<10^{-6}$), the FID signal corresponds to the transverse relaxation decay (eq.7). In this case, it can be used to calculate the transverse relaxation time (T_2) or, alternatively, a Fourier transform can be carried out to obtain the spectral components of the signal and then the chemical composition of the sample (FT-NMR). Conversely, in presence of a non-homogeneous field, the FID signal decreases faster than the transverse relaxation decay. In this case, the decay is mainly due to the field inhomogeneity, and not to the magnetization of the sample. The exponential decay of the FID signal can be expressed in terms of the time constant T_2^* instead of T_2, where:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \gamma \frac{\Delta B_0}{2} \qquad (16)$$

As can be seen from eq. (16), the time constant T_2^* is inversely proportional to the static magnetic field variation (ΔB_0).
The field inhomogeneity causes a resonance frequency variation in the different sections of the sample. The result is that it is not possible to perform the spectral analysis to determine the molecular structure of the sample. Figure II-11 and Figure II-12 show the effect of the field inhomogeneity in the time-domain and in the frequency-domain.

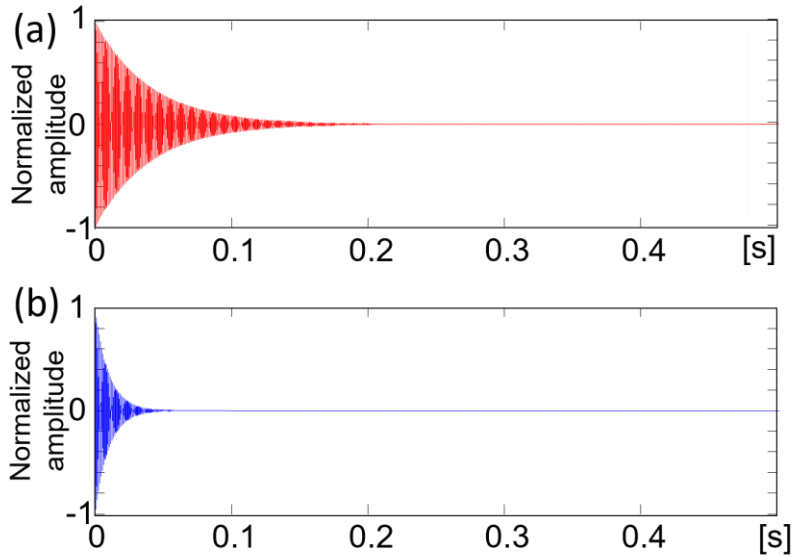

Figure II-11 *Effect of inhomogeneity in time domain. (a) Homogeneous field; (b) inhomogeneous field*

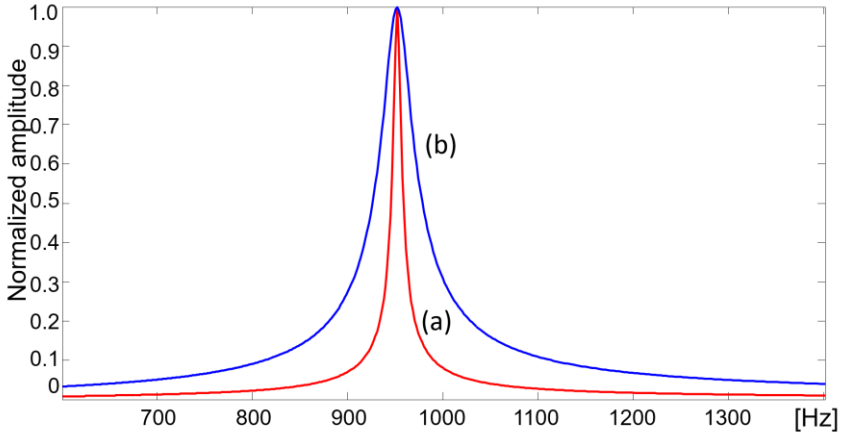

Figure II-12 *Effect of inhomogeneity in frequency domain. (a) Homogeneous field; (b) inhomogeneous field*

In the frequency-domain, the faster decay in time corresponds to a broader spectral line, and then a higher frequency resolution that does not allow discriminating the several spectral components related to the molecular composition of the sample. In this case is only possible to perform a quantitative analysis of the liquid part of the sample (that are the molecules containing ^{1}H atoms excited at the resonance frequency).

II.3.1 CPMG sequence

This technique is largely adopted in Low Field NMR because it allows obtaining qualitative information about the sample in presence of field inhomogeneity. In more details, it is used to determine the transversal relaxation time (T_2) which is the time constant of the exponential decay of the NMR signal in presence of a perfect homogeneous field.

The CPMG sequence (Fig. II-13) consists of a $\pi/2$ RF pulse (same as the single pulse) followed by a train of π RF pulse (it takes twice the time of the $\pi/2$ RF pulse). The distance between the $\pi/2$ RF pulse and the first π RF pulse is denoted as τ; the time interval between two π RF pulses is 2τ. The effect of this kind of sequence is to obtain a FID signal after the $\pi/2$ RF pulse which has a fast decay due to the field inhomogeneity, and a train of echo signals in the middle of the distance between the π RF pulses; the amplitude of the echo signals gradually decreases after each π RF pulse. The maximum of the amplitude of each echo gives the relaxation curve that can be used to estimate the decay time T_2.

The acquisition of the FID signal starts at the time 0. It is generated by the $\pi/2$ RF pulse (that is not shown in the figure); after a time τ a π RF pulse is

generated and after a time τ there is the peak of the first echo; an echo is visible each 2τ

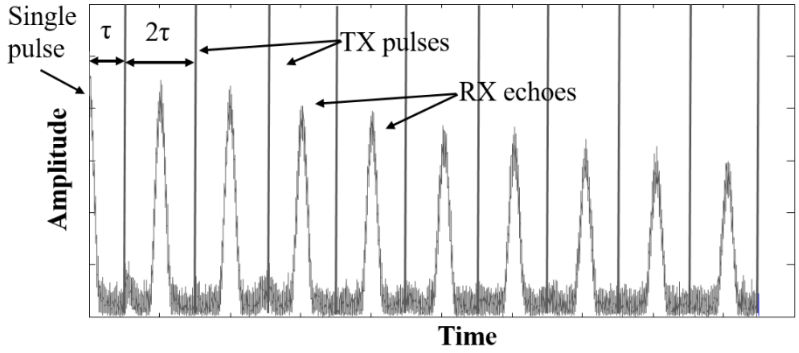

Figure II-13 *CPMG sequence (first nine echoes)*

In the evaluation of the transverse relaxation curve, the odd echoes are usually discarded because the amplitude can be affected by an error due to an incorrect length of the π R pulse, therefore only the even echoes are taken into account.

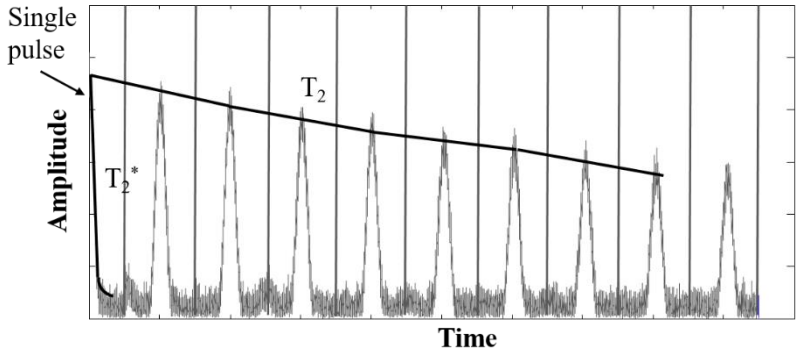

Figure II-14 *Comparison between the transverse relaxation curve obtained with the CPMG (decay with time constant T_2) and the decay curve of the FID signal affected by field inhomogeneity (decay with time constant T_2^*)*

Figure II-14 shows an example of the evaluation of transverse relaxation curve with the CPMG and the comparison between the decay of the curve obtained using the CPMG and the decay of the FID signal that is affected by the field inhomogeneity (see eq.16). As can be seen from the figure:

$$T_2^* \ll T_2 \qquad (17)$$

II.4 The adopted TD-NMR measurement system

The TD-NMR instrument used for the experimental work is composed by three main components:

- NMR console;
- Magnet and probe;
- Personal computer.

II.4.1 NMR console iSpinNMR™

The NMR console is manufactured by Spincore Technologies Inc. It includes all the electronics to perform the generation and transmission of the excitation pulse sequences as well as the acquisition of the RF signals from the probe. It allows employing a resonance frequency up to 37.5 MHz. The main parameters are:

- Internal RF PA 20 W PEP;
- Internal Arbitrary Waveform Generation;
- DDS for RF output pulse with 14-bit resolution @300 Ms/s;
- ADC with 14-bit resolution @75 Ms/s;
- Internal Digital Down Conversion;
- Internal Signal Averaging of Baseband data in multiple acquisitions;
- USB 2.0 Interface to transfer data to PC.

A block diagram of the NMR console is depicted in figure II-15. It is composed by three block: (1) The RadioProcessor USB is a digital board that manages the data acquisition, the excitation pulse generation and the timing. In the receiving section, the pre-amplified RF signal from the probe is acquired by a 14-bit resolution ADC and then is demodulated in a digital down converter. In the digital transmitter the RF pulses are generated and converted to analog using a 14-bit resolution DAC. (2) The RF Power Amplification is composed by a RFPA and a low-pass filter. (3) The Small-Signal Pre-amplification is composed by a low-noise 60 dB pre-amplifier and a low-pass filter.

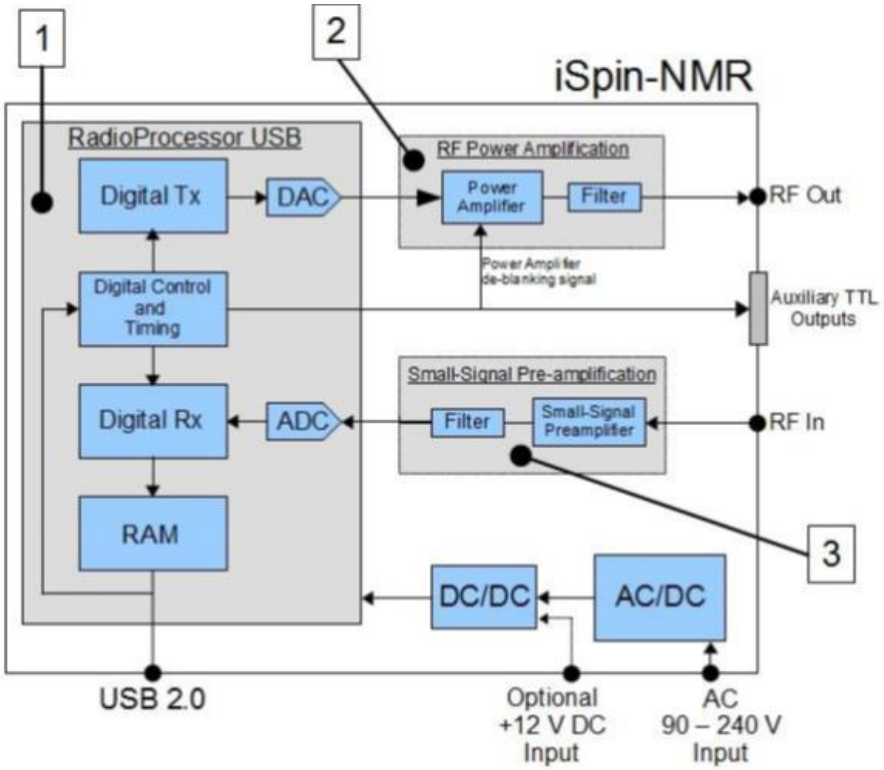

Figure II-15 *iSpinNMR™ console block diagram (Source: Spincore Technologies Inc., 2014)*

II.4.2 Magnet and probe

These components have been designed for the specific application considering that the probe, placed in the magnet, has to be able to host a whole hazelnut. To this aim, the probe has a cylindrical coil characterized by height and diameter both equal to 25 mm. The permanent magnet has the following characteristics (Figure II-16):

- Air gap: 40 mm;
- Pole face diameter: 150 mm
- Dimension of the uniform magnetic field: 25 mm diameter, 25 mm high;
- Field uniformity: $10^{-4}\ \Delta B_0/B_0$;
- Static magnetic field: 0.513 T at 23.0 °C.

Considering the value of the static magnetic field, the resonance frequency of the system is 21.8 MHz

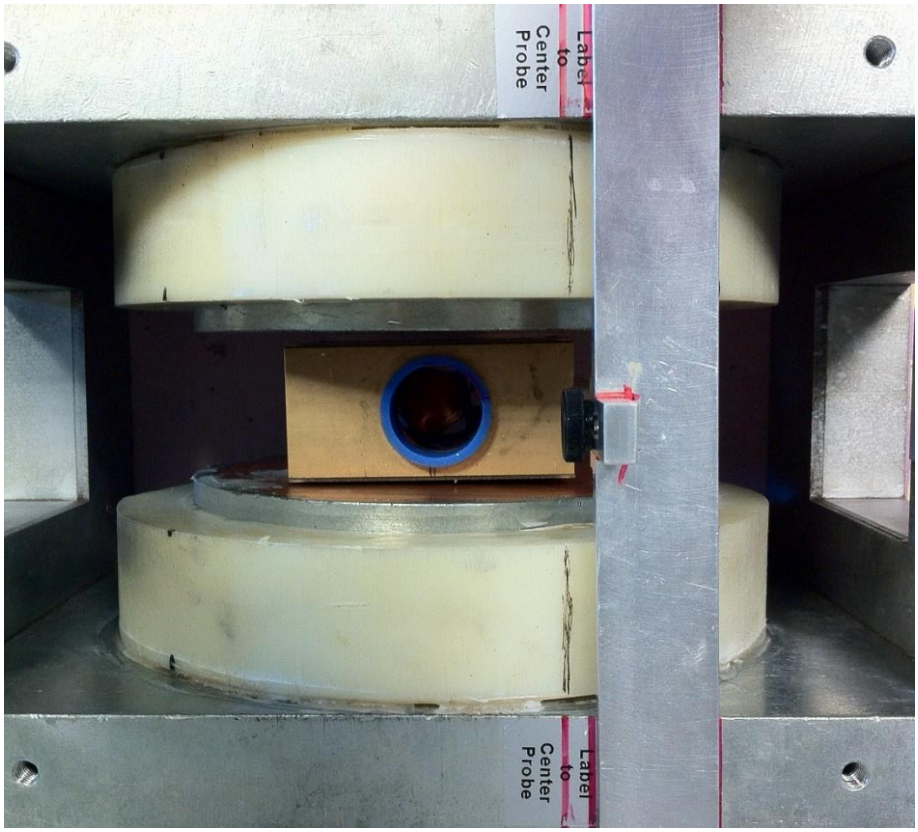

Figure II-16 *Permanent magnet and probe*

The field homogeneity of the magnet is 10^{-4} $\Delta B_0/B_0$.This is a typical value for a commercial permanent magnet. The figure II-17 shows the field variation in function of the position in the x-y plane. Considering a ^1H probe, a variation of 1mT of the static magnetic field causes a variation of the Larmor frequency greater than 40 kHz.

In order to work in the highest homogeneity area, the centre of the coil of the probe must be placed in correspondence of the (0,0) x-y coordinates (Fig. II-17). This point can be found considering the maximum amplitude of the FID signal for a sample at the resonance frequency, achieved moving the probe inside the magnet.

The probe needs to be tuned to the resonance frequency of the system (f_0=21.8 MHz) and, in addition, the impedance has to be matched to be matched to 50 Ω. This has been made by means of a Vector Network Analyzer (Fig. II.18), adjusting the two variable capacitors C_{tune} and C_{match} (Fig.II-3)

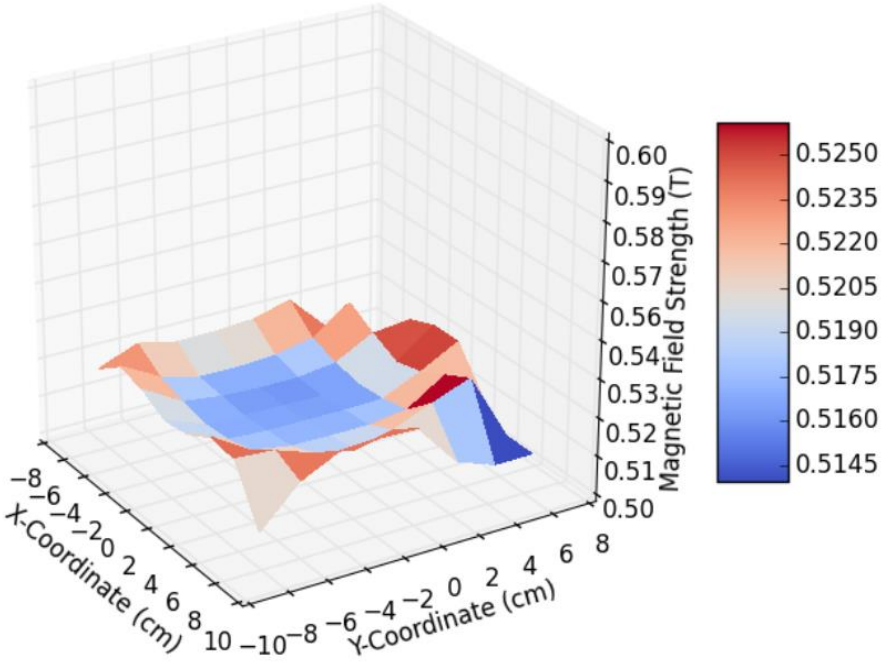

Figure II-17 *Field strength as a function of the position in the magnet*

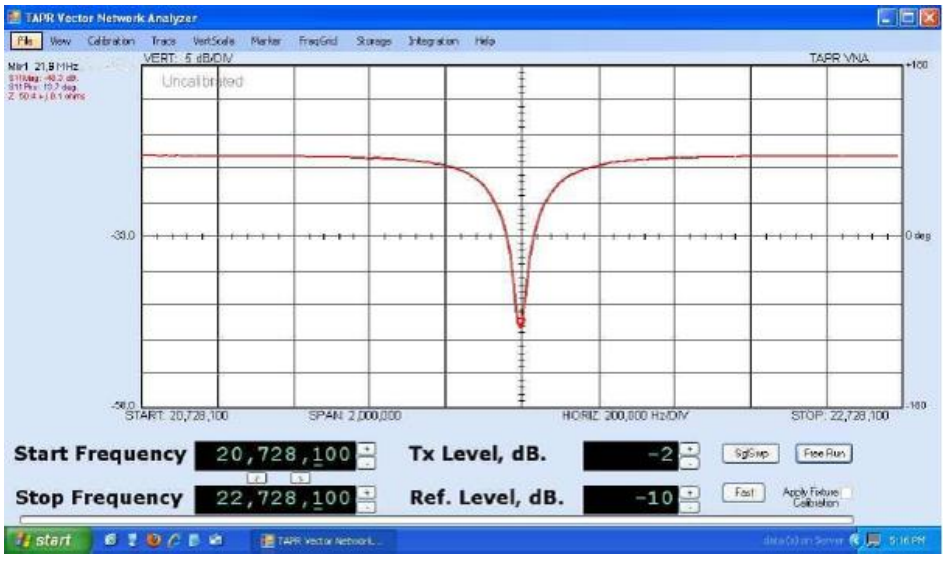

Figure II-18 *Probe tuning and matching*

II.4.3 Personal computer

The NMR console is connected to a personal computer through an USB 2.0 interface. The complex data of the acquired NMR signal, after a baseband digital down conversion, are stored in an internal RAM memory and then can be transferred to the PC for the automatic processing. A Labview user interface has been implemented to send the setting commands to the NMR console and to carry out the signal processing and the classification algorithm on the received signals.

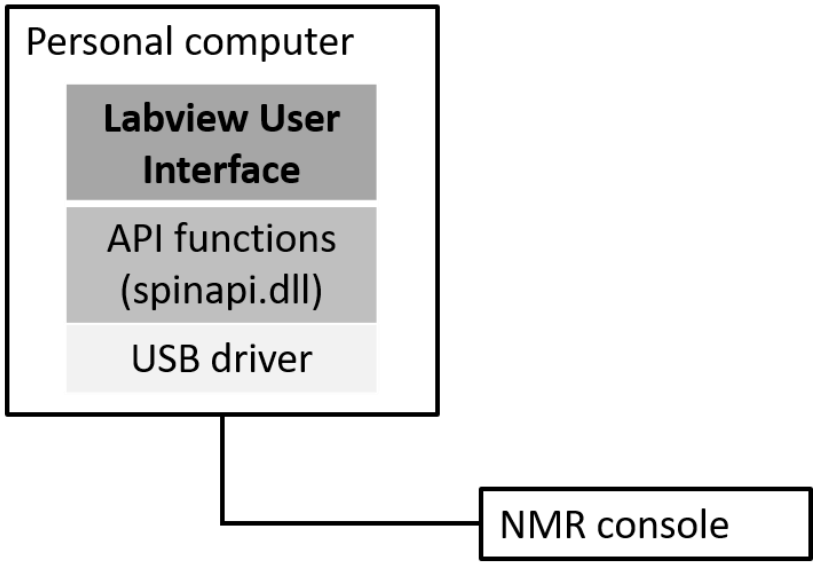

Figure II-19 *Software component diagram of the Labview user interface*

The Labview user interface makes use of the API functions, provided in a spinapi.dll library by the Spincore Technologies Inc., to send the setting commands to the programmable RadioProcessor board (Fig. II.19).

II.5 Influence parameters in liquid and solid samples

The development of industrial applications with TD-NMR has been possible thanks to the decreasing costs of the electronic components and the use of permanent magnets instead of the more expensive superconducting magnets. However, there are two main disadvantages of using TD-NMR with permanent magnet:

- Field inhomogeneity;
- Dependence of the field from the temperature

The effects of the field inhomogeneity have been already illustrated in the previous paragraphs.

Permanent magnets have a coefficient of temperature in the order of -1000 ppm/°C, so the static magnetic field, and then the resonance frequency, changes when the temperature changes. In order to limit the resonance frequency variation due to the temperature variation, the magnet has been equipped with a temperature controller and insulated by means of sheets of extruded polystyrene (Fig. II-20).

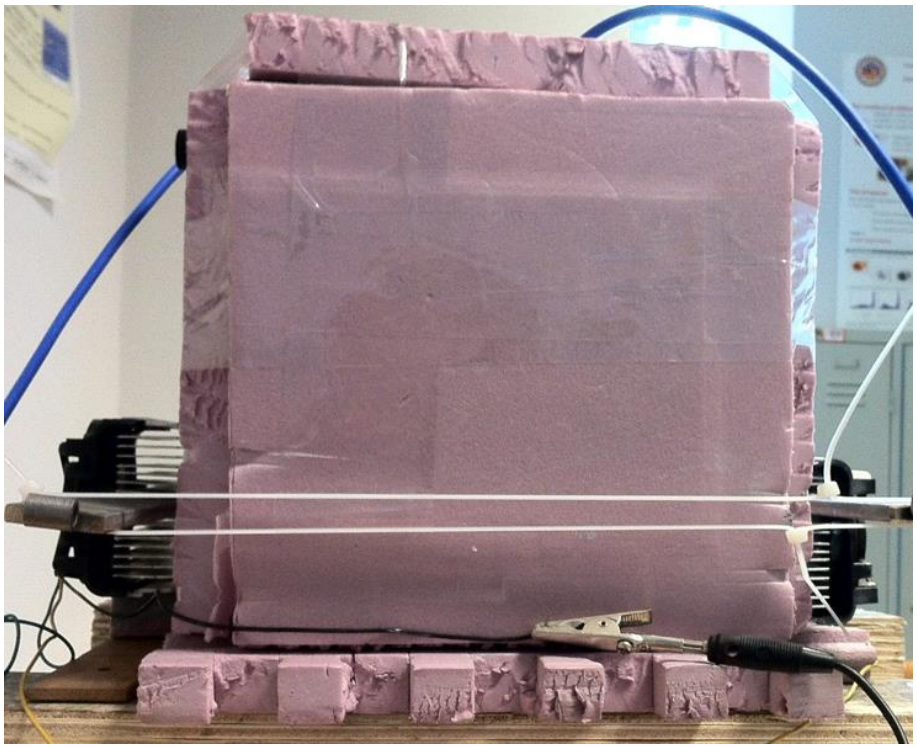

Figure II-20 *Magnet insulated by means of extruded polystyrene sheets*

II.5.1 Temperature controller

The temperature controller is composed by:

- 2 RTDs PT100 A class placed inside the magnet, near each pole;
- 2 Peltier Modules, for heating and cooling, placed outside the magnet, behind each pole;
- 1 microcontroller Microchip PIC® 18F4550, implementing the PID control algorithm

The temperature controller allows reducing the temperature variation of the magnet at ± 0.1 °C.

In this way, when the operating temperature is kept at 23.0 ± 0.1 °C the resonance frequency variation is lower than 5 kHz around the nominal value.

The following figure shows a block diagram of the temperature controller system:

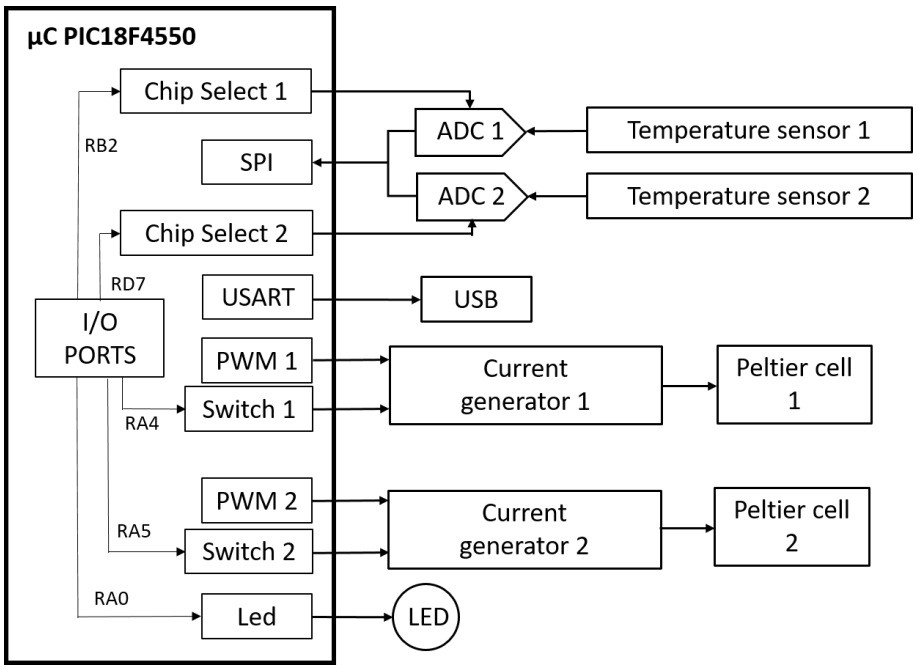

Figure II-21 *Block diagram of the temperature controller*

The system is composed by two independent channels, one for each pole of the magnet. The microcontroller acquires the temperature from the two channels on the SPI port. Each temperature sensor is connected to a 22 bit resolution ADC that provides the temperature data on a SPI port connected to the microcontroller (Fig.II-21), which can select the channel by means of a chip select signal. The PID control algorithm, on the basis of the read temperatures, set the value and the direction of the current in the current generators of the two channels, by means, respectively, of a PWM signal and a switch signal. In this way, it is possible to control the heating and cooling level of the peltier cells connected to the current generators.

II.5.2 Other influence parameters

In offline applications, NMR experiments are frequently conducted on liquid samples; solid samples are pre-treated to extract liquid parts, or cut in

pieces for being inserted into the test tube, which is arranged into the probe. In this way, it is possible to use commercial benchtop NMR systems. These kind of systems are used in a laboratory environment, where a strict control of the parameters related to the equipment and the samples is allowed. In particular, the temperature of the sample must be controlled because the NMR signals, e.g. the CPMG signal, are affected by the temperature variations (Carosio *et al.*, 2016). Moreover, some other issues have to be highlighted working with solid samples. This is often the case of inline applications, in which there is no pre-treatment of the material under test. Solid samples are, in many cases, heterogeneous materials with different weight and shape; this causes several effects that do not occur in liquid samples:

- Amplitude variation of the NMR signal and then S/N variation, due to the weight variation of the samples;
- resonance frequency variation due to the different positions of the samples inside the magnet;
- presence of several time constants into the transverse relaxation decay due to heterogeneous samples.

Finally, the physical properties of hazelnuts influence the NMR signals, because of the presence of the shell. Indeed, the hazelnuts are mainly composed by lipids, which are well detected by TD-NMR, while the shell presents some other components, making this product a heterogeneous material. In details, the presence of the shell causes a faster decay of the transverse relaxation signal compared to the unshelled nut (Fig. II-22). Moreover, even considering nuts with the same size, the shape and the weight of the kernel may be very different, affecting the S/N ratio of the NMR signal related to each nut. In particular, being the weight of the kernel closely related to the maximum signal amplitude, the S/N ratio is smaller for lighter nuts (Di Caro *et al.*, 2017).

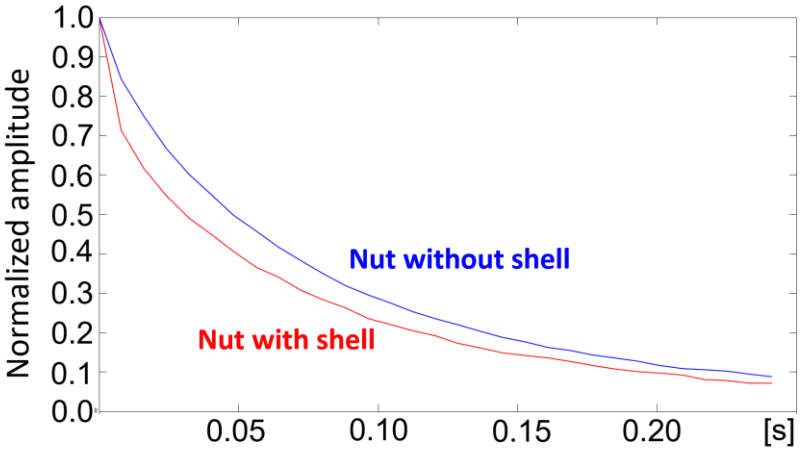

Figure II-22 *Example of CPMG signal for nut with shell and nut without shell*

Chapter III
Hazelnut oil characterization

In order to set-up the instrument for hazelnuts diagnosis, preliminary tests were conducted on oil samples extracted from the hazelnuts. In this way, it was possible to concentrate the analysis on the composition of the material without considering some external parameters like humidity or the shape of the sample that can affect the results. This allowed determining the parameters related to the CPMG sequence, and the pre-processing and processing algorithms to carry out the analysis.

III.1 Instrumental parameters

Before starting the analysis on the samples, a test of the NMR instrument has been carried out in order to verify the operation of the equipment (Fig. III-1).

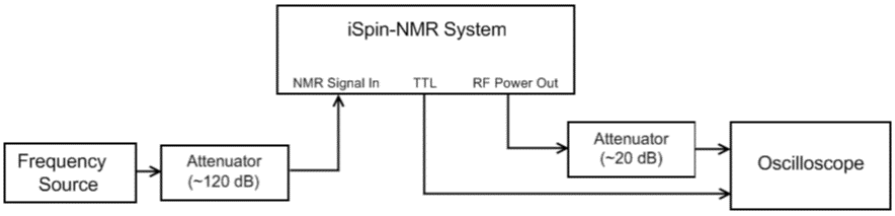

Figure III-1 *NMR system test setup*

The NMR console provides a TTL output connector where the main internal digital signals are available, this allows checking the timing synchronization of the transmit and the receiving section of the instrument.

In particular, two parameters have been set:

- Blanking delay;
- Transient Time (Dead time).

These parameters depend on the system hardware, in particular the RF power amplifier. The former is the time needed to warm-up the amplifier; it

affects the total time of a single NMR experiment because the pulse transmission can start at the end of this delay. The latter is the time needed to ring down the high voltage induced in the coil at the end of the RF pulse. It affects the receiver section because the sample acquisition can start after this time.

The values set are:

- Blanking delay = 3 ms;
- Transient time = 16 µs.

The main parameter related to the pulse sequences (both single pulse and CPMG) is the width of the π/2 RF pulse. This is the value that allows rotating the magnetization vector by π/2 respect to the direction of the static magnetic field (eq. 6), and it can be determined using a pulse width finder procedure consisting in calculating the amplitude of the FID signal for several values of the pulse width (t_p). The width corresponding to the maximum amplitude of the FID signal is the width of the π/2 RF pulse (Fig. III-2).

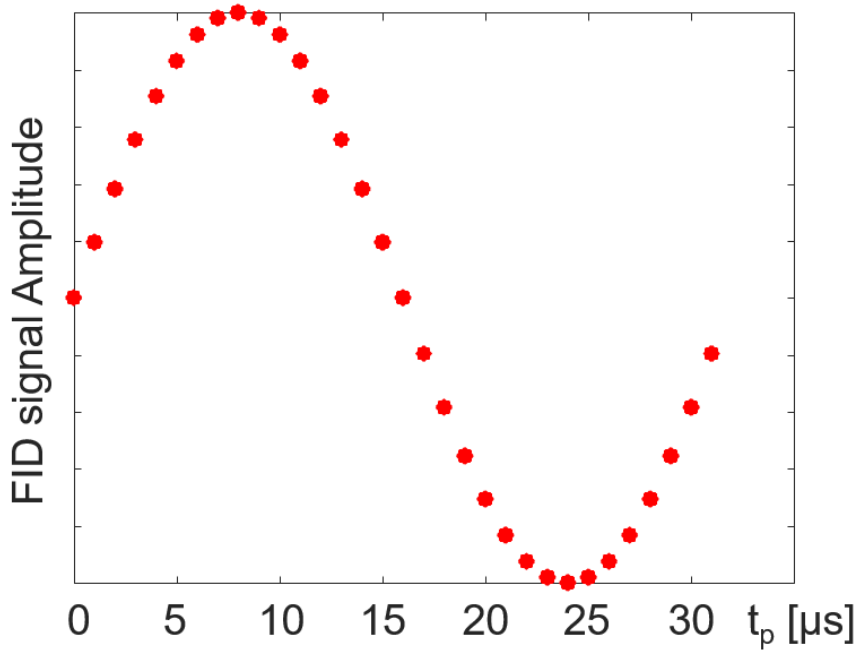

Figure III-2 *Pulse width finder procedure*

This procedure allows determining the π RF pulse used in the CPMG sequence, which is the pulse width with a null FID signal amplitude. The π RF pulse with is approximately twice as much the π/2 RF pulse width.

Hazelnut oil characterization

Applying the pulse width finder procedure in the case of interest the following values have been found:

- π/2 RF pulse width = 8 μs;
- π RF pulse width = 16 μs.

III.2 CPMG setup

As described in the chapter II, in order to limit the effect of the field inhomogeneity due to the use of permanent magnets, it is suitable to make use of the CPMG sequence instead of the single pulse

In addition to the width of the π/2 RF pulse and the π RF pulse, the CPMG sequence requires the definition of other parameters (Fig. II-13):

- τ, which determines the distance among the π RF pulse and then the echoes (2τ);
- number of echo to acquire, which determines the acquisition time.
- sampling frequency of the received signal;

The time interval between two π RF pulses has been set to 4 ms (τ = 2 ms). This is also the time between the echoes that represent the points of the transverse relaxation decay, so it has been chosen in order to have enough point of the relaxation curve.

For each CPMG sequence, 50 echoes were acquired; the number of echoes is limited by the S/N ratio: when the amplitude of the echo signal is below the noise level, it is impossible to detect the peak.

A sampling frequency of 500 kHz was adopted to have enough points of the CPMG signal for each echo.).

The acquisition time can be calculated from the previous parameters:

$$CPMG\ acquisition\ time = N_{echo}(\pi\ pulse\ width + 2\tau) + \tau \quad (18)$$

In this case, the acquisition time is about 200 ms.

III.3 Echo response pre-processing

NMR signals have low amplitude (about 10 μV) and are affected by a high level of noise, due, both to RF interferences and A/D conversion. In the CPMG sequence the amplitude of the echoes decrease at each π RF pulse while the noise level does not change, therefore it is difficult to estimate the amplitude of the last echoes (Fig. III.3).

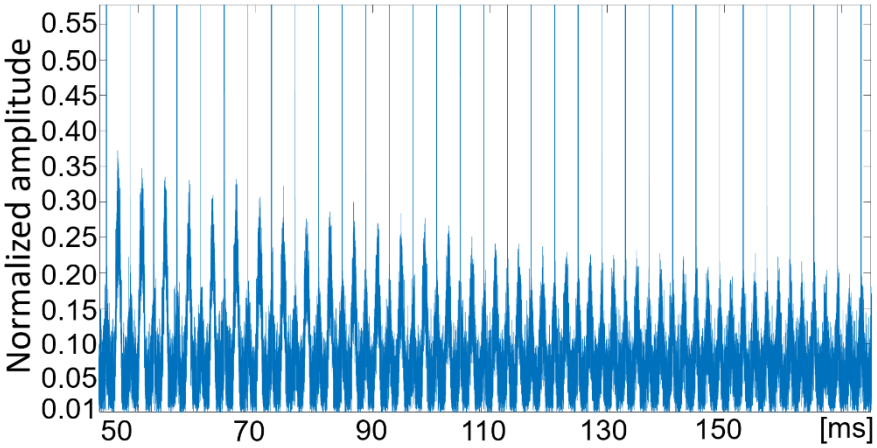

Figure III-3 *Example of acquisition of CPMG signal after 50 ms*

Usually, the NMR instruments are able to perform an internal signal averaging over several acquisitions. This allows increasing the S/N ratio of the received signal. In particular:

$$\frac{S}{N} \propto \sqrt{N} \qquad (19)$$

Where N is the number of acquisitions of the signal (Fig. III-4).

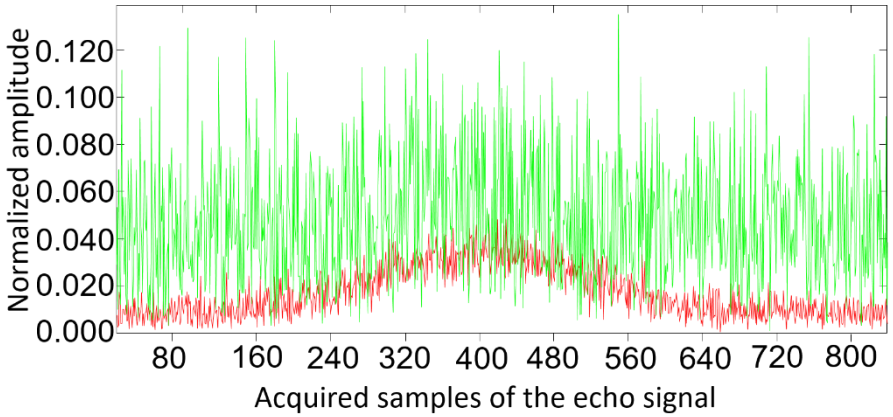

Figure III-4 *Acquired echo signal (45th echo) with a single sequence (in green) and with an average over 32 sequence (in red)*

Considering the acquisition time, and the time interval between two acquisitions on the same sample, which has to be greater than the repetition delay (the time to recover the magnetization at the end of the pulse sequence),

the applicability of this procedure is limited to applications that do not require a fast execution time (e.g. laboratory applications).

In this work, in order to better evaluate the amplitude of the echoes that have a lower S/N ratio, avoiding the use of the time consuming multiple acquisitions procedure, a pre-processing of the signal has been performed.

As previously described, each echo peak corresponds to a point of the transverse relaxation decay, and so the relevant information related to the CPMG signal are contained in the acquired samples of the echoes.

Three methods have been evaluated and compared to detect the echo peak amplitudes (Di Caro *et al.*, 2016):

- Mean value of k samples around the peak position;
- Sum of the spectral components of the magnitude of the FFT for the echo signal;
- Maximum of the curve obtained with a polynomial regression on the echo signal.

The echo peaks in the CPMG sequence are located in the middle of the time between two π RF pulses, therefore they are in a known position in the acquired signal.

III.3.1 Mean value

The Mean value method exploits this property to determine the peak position and estimate its value applying a smoothing factor to reduce the effect of noise.

The position (n) of the peak of the i-th echo in the acquired samples is:

$$n = i \cdot (2\tau + \pi\ pulse\ width) \cdot f_s \qquad (20)$$

Where f_s is the sampling frequency of the acquired signal.

The amplitude of the i-th echo peak is calculated as the mean value of the samples in the interval: $\left[n - \frac{k}{2}, n + \frac{k}{2}\right]$.

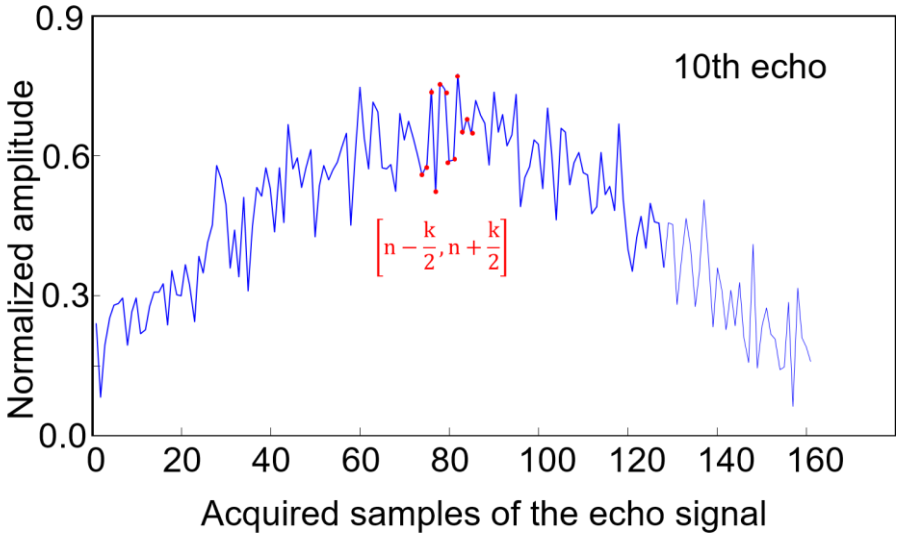

Figure III-5 *Mean value method on echo signal with k=12*

III.3.2 FFT magnitude

The second method relies on the property that the number of nuclei excited at the resonance frequency is proportional to the area of the FFT magnitude. This is because, for the magnet inhomogeneity, the nuclei in the different sections of the sample are subjected to a different resonance frequency, so their contribution to the FFT magnitude is located in a different frequency respect to the resonance frequency of the system (Fig.II-12). The noise level is calculated and subtracted considering the FFT magnitude outside the frequency range of interest.

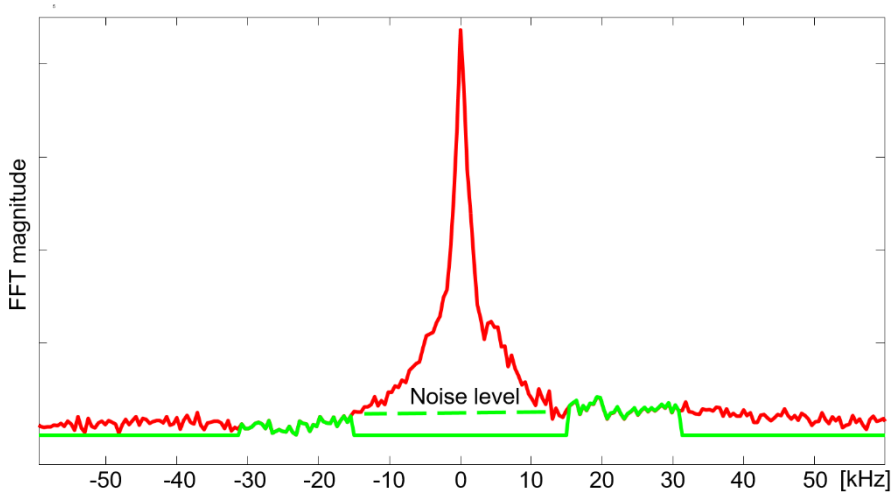

Figure III-6 *FFT magnitude of the echo signal (in red); noise level (in green)*

The noise level (NL) has been calculated on the basis of the mean value and the standard deviation of the FFT magnitude corresponding to the noise (the green regions in fig. III-6):

$$NL = \overline{noise} + s(noise) \quad (21)$$

III.3.3 Polynomial regression

About the third method, a second order polynomial regression has been applied to the echo signal (Fig. III-7), considering, for each echo, 160 samples around the peak, located at the position according to eq. (20); the echo peak is the maximum of the regression curve (Eq. 22) - (Eq. 24):

$$y_{reg} = P_2 x^2 + P_1 x + P_0 \quad (22)$$

$$\Delta = P_1^2 - 4P_2P_0 \quad (23)$$

$$max = -\frac{\Delta}{4P_2} \quad (24)$$

Where max is the echo peak.

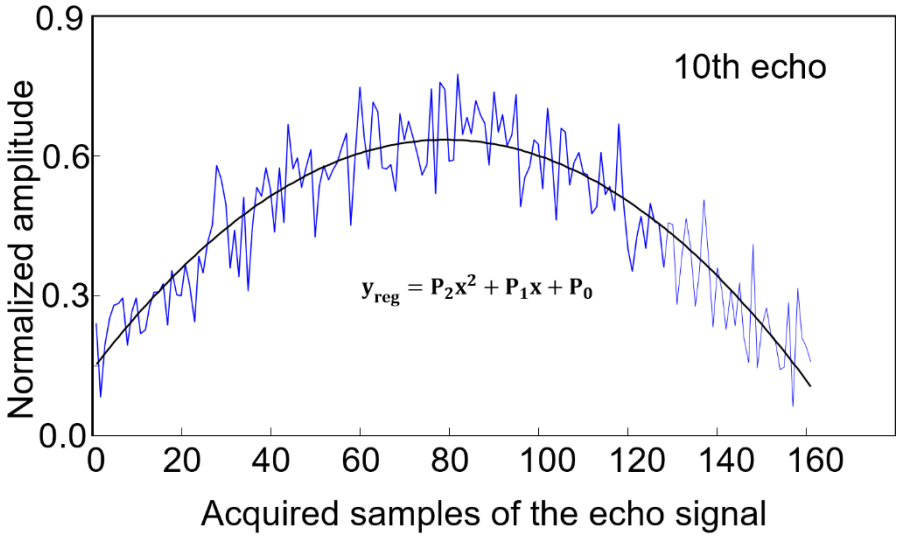

Figure III-7 *Second order polynomial regression on echo signal*

III.3.4 Comparison among the three methods

In order to choose the most robust method, tests were made on signals suitably corrupted by noise and the repeatability in the peak estimation was valuated.

In particular, two signal (one from healthy hazelnuts oil (HO), one from unhealthy hazelnuts oil (UO)) have been acquired and Gaussian noise was added to the signals generating 100 noise signals for each one of these. The three pre-processing algorithms were run on the two set of signals and the standard deviation in the peak estimation was calculated. In Figure III-8, for both HO and UO oil, the measured standard deviation is reported. As can be seen the polynomial regression algorithm is characterized by the best repeatability over all on the smallest peaks that characterize the oil response, in detail the relative uncertainty is quite constant for all the peaks.

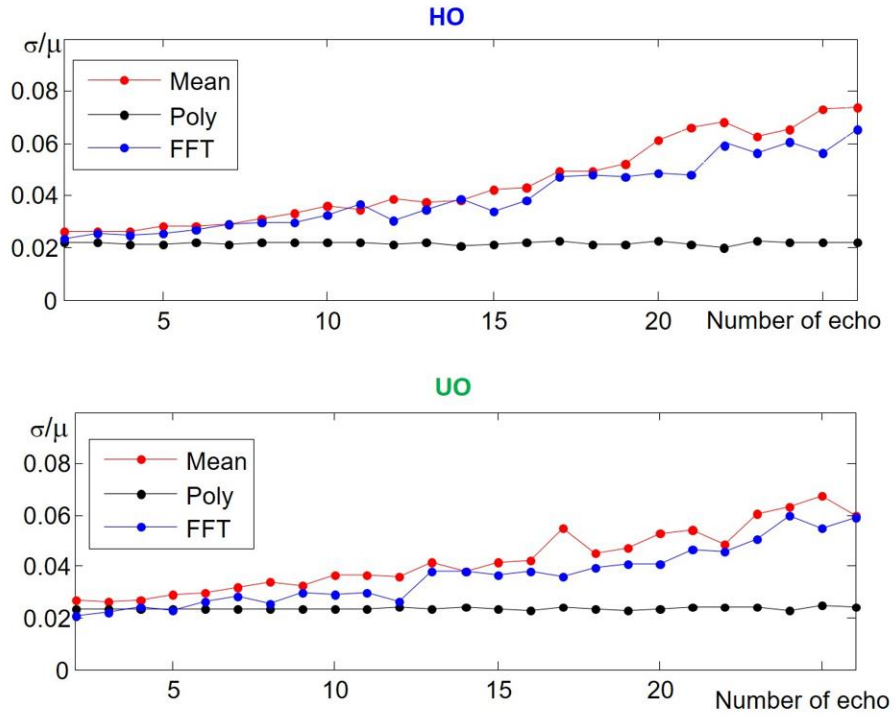

Figure III-8 *Comparison among the three pre-processing methods*

III.4 Study of the echo decay envelope

As described in the chapter II, the echo decay envelope represents the transverse relaxation decay, which provides qualitative information about the sample. Over the acquired CPMG signal, composed by the FID signal and 50 echoes, the odd echoes have been discarded because the amplitude can be affected by an error due to an incorrect length of the π RF pulse. This error doesn't affect even echoes, therefore only 25 echoes and the initial FID signal have been used in the analysis; moreover, normalized amplitudes have been used (Fig. III-9).

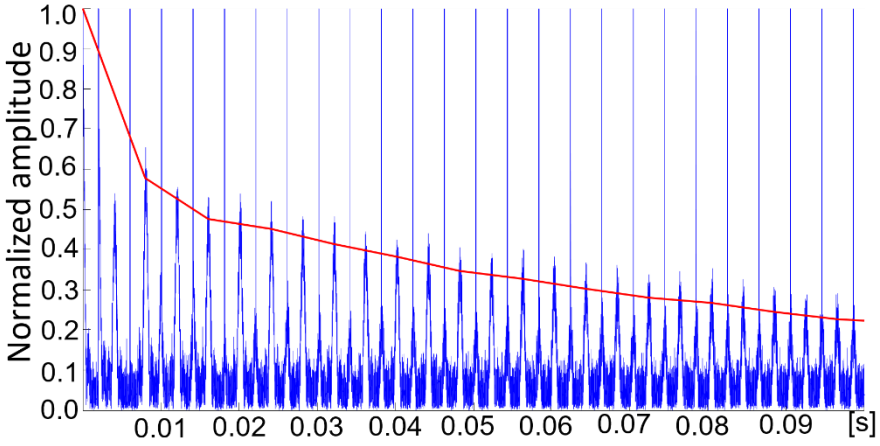

Figure III-9 *Acquired CPMG signal and echo envelope*

Two different approaches were set up. Using these quantities it is possible to discriminate between oil extracted from healthy hazelnuts and oil extracted from unhealthy hazelnuts.

The first method defines a parameter, ES, that synthetizes the normalized amplitudes of the final echoes, the second one approximate the echo decay envelope by means of a multi-exponential function.

III.4.1 ES parameter

The following parameter has been introduced as a synthetic information about the final trend of the response:

$$ES = \sum_{i=p}^{26} \frac{A(i)}{A(1)} \qquad (25)$$

Where p was fixed on the basis of experimental results, A(i) is the amplitude of the i-th echo, A(1) is the amplitude of the FID signal.

The ES parameter takes account of the slowest relaxation components of the decay signal.

III.4.2 Multi-exponential approximation

The transverse relaxation decay obtained from the CPMG signal can be expressed as a multi-exponential function:

$$y(t_i) = \sum_{j=1}^{m} x_j e^{-\frac{t_i}{T_{2j}}} \quad 1 \leq i \leq n \qquad (26)$$

Where $y(t_i)$ is the amplitude of the i-th echo; x_j and T_{2j} are the amplitude and the time constant of the j-th relaxation component.

Hazelnut oil characterization

A multi-exponential inversion algorithm should be used to get the relaxation spectrum (the distribution of the time constants of the relaxation decay). There are many algorithms implemented to solve this kind of problem (see Appendix). In this work, an algorithm presented in (Chen *et al.*, 2009), based on singular value decomposition (SVD), has been used. This method makes use of the matrix theory to solve the eq. (26).

The first step is to rewrite eq. (26) in matrix form:

$$Y = A \cdot X \qquad (27)$$

Where Y is the array with the n samples of the measured signal, X is the array containing the amplitude of the m relaxation components, A is a $n \times m$ matrix with coefficients:

$$A_{n\times m} = [a_{ij}] = \left[e^{\frac{-t_i}{T_{2j}}}\right] \qquad (28)$$

T_{2j} are m pre-assigned, logarithmically spaced, relaxation time constants in the T_2 distribution range.

t_i are the sampling times of the measured signal.

The solution of eq.(27) determines the T_2 distribution. This is a typical highly ill problem and requires an iterative solution in which a non-negatively condition for the components of the X array is applied.

First, the singular value decomposition of the matrix A is calculated:

$$A_{n\times m} = U_{m\times m} \cdot S \cdot V^t_{n\times n} \qquad (29)$$

Then, the solution of the equation (27) with linear least square algorithm can be found:

$$X = V \cdot S^{-1} \cdot U^t \cdot Y \qquad (30)$$

If some components of the X array are negatives, the non-negativity condition is applied and the inversion algorithm is computed again, until all elements of X become non-negatives (Fig. III-10).

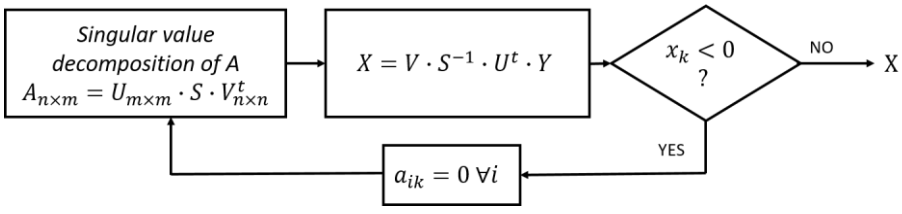

Figure III-10 *SVD algorithm*

The following parameters have been used: the number of samples of the measured signal n = 26; the number of pre-assigned time constant can't exceed

n (otherwise the equation (27) is undetermined), in this case is 25, in a range from 0.01 s to 0.25 s.

III.5 Experimental results on oil samples

III.5.1 Sample preparation

All the samples of oil used for experimental verification were extracted from hazelnuts of cultivar "Tonda di Giffoni" and prepared using the same procedure, as described in (Amaral *et al.*, 2006) and (Memoli *et al.*, 2017): the hazelnuts were manually deshelled, cracked, and sorted by a visual inspection of the kernels in two different samples [apparently healthy (no damage), and unhealthy (mouldy and cimiciate) hazelnuts]. Thereafter, 100 g of finely chopped nuts were suspended in diethyl ether (1:3 w/w) and kept under magnetic stirring at room temperature in a stoppered dark flask. After 2 h, the suspension was filtered through filter paper and the residues were treated once again in the same fresh solvent, applying the described procedure twice. The residues were washed twice with 10 ml of diethyl ether; thereafter, the solvent was rotary evaporated under reduced pressure at 40 °C. The oil extracted was stored at 4 C in tubes protected from light with aluminium foil and flushed with nitrogen. Finally, the chemical analysis on the two kinds of oil samples was performed by means of the gas chromatography (Memoli *et al.*, 2017).

III.5.2 Experimental results

Eighteen oil samples (9 from healthy hazelnuts, 9 from unhealthy hazelnuts) were prepared for the training phase in order to set up the processing algorithms whereas, different fifty oil samples (25 from healthy hazelnuts, 25 from unhealthy hazelnuts) were prepared for testing and validation (Di Caro *et al.*, 2017b).

Figure III-11 shows the curves of the maximum values of the echoes for the training samples. As can be seen, as expected, the curves of the two kind of oils are well separated when the number of echo increases.

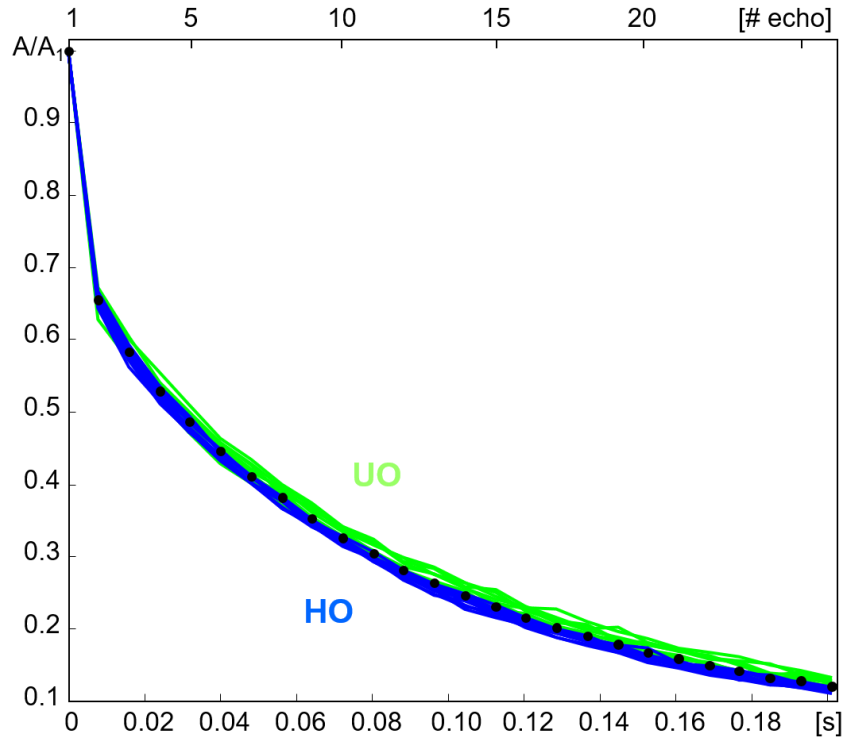

Figure III-11 *Normalized echo amplitude for healthy (blue) and unhealthy hazelnut oil (green)*

ES Parameter

In order to fix the value of p of eq. (25) the value of the MARGIN is evaluated on the training samples. The MARGIN is defined as the difference between the minimum value of ES parameter for oils extracted from unhealthy hazelnuts, and the maximum observed for the other kind of oil:

$$MARGIN = min(ES_{UO}) - max(ES_{HO}) \qquad (31)$$

In figure III-12 the trend of the MARGIN with p is reported. The greater values of the MARGIN are obtained when the ES parameter is calculated taking into account only and all the echoes where the curves of the two kinds of oil are separated, so the value of p can't be too small (p=5 corresponds to 0.024 s in Fig.III-7, where the echo responses are high and very similar). Furthermore, when p is too large (p=20 corresponds to 0.164 s), the ES parameter is calculated on few and smaller echo signals that are affected by a higher level of noise, and the MARGIN decreases. The maximum MARGIN is obtained with: p = 17;
Hereafter this value is used.

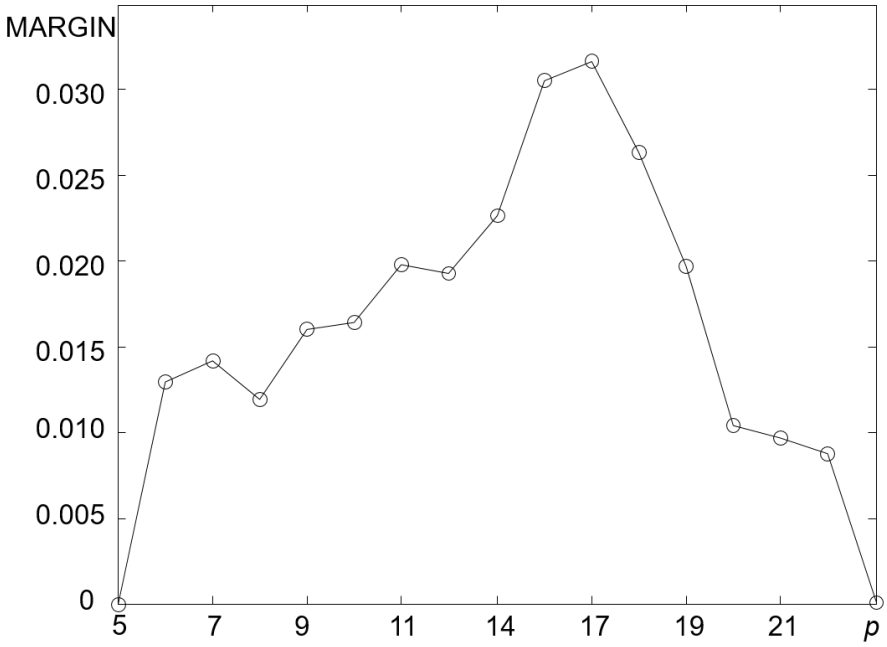

Figure III-12 *MARGIN related to different values of p (p=15 correspond to 0.12 s in Fig. III-7)*

In Table III-1 the results related to 20 measures on each of the 18 training samples (9 healthy hazelnut oils and 9 unhealthy hazelnut oils) are summarized; the mean, the standard deviation and maximum of the ES parameter are reported for healthy hazelnut oil (HO) in the first three columns, and for the unhealthy hazelnut oils (UO) in the next three columns, but the minimum value is indicated.

Table III-1 *The measured parameters on the training healthy (HO) and unhealthy (UO) hazelnut oils on repeated tests*

	ES_{HO} mean	ES_{HO} σ	ES_{HO} max	ES_{UO} mean	ES_{UO} σ	ES_{UO} min
Sample # 1	1.53	0.0027	1.54	1.74	0.0034	1.73
Sample # 2	1.48	0.0032	1.49	1.66	0.0039	1.65
Sample # 3	1.48	0.0036	1.48	1.65	0.0033	1.64
Sample # 4	1.50	0.0025	1.50	1.58	0.0038	1.57

Sample # 5	1.53	0.0033	1.54	1.66	0.0036	1.65
Sample # 6	1.50	0.0030	1.51	1.57	0.0033	1.56
Sample # 7	1.45	0.0031	1.46	1.70	0.0046	1.69
Sample # 8	1.52	0.0044	1.53	1.58	0.0042	1.58
Sample # 9	1.52	0.0034	1.53	1.57	0.0039	1.56

For the classification of the two kinds of oils, all the observed mean values, have been taken into account, the observed statistics on the whole sets are reported in Table III-2. To this aim two thresholds, Upper Threshold (UTHR), and Lower Threshold (LTHR), are defined and consequently three regions are identified: (1) the values of ES greater than UTHR identify the unhealthy hazelnut oil; (2) the values lower than the LTHR identify the healthy hazelnut oil; (3) the values between the thresholds identify the ambiguous region.

Table III-2 *Parameters on the whole population of healthy (HO) and unhealthy (UO) hazelnut oils: Mean Value ($\overline{ES}$), Standard Deviation (s)*

$\overline{ES}_{HO}$	$s(ES_{HO})$	ES_{HO} maximum	$\overline{ES}_{HO}$	$s(ES_{UO})$	ES_{UO} minimum	MARGIN
1.50	0.03	1.53	1.63	0.06	1.57	0.04

The two thresholds (UTHR and LTHR) have been fixed on the basis of the measured ES values, and in particular on the basis of the mean and standard deviation observed on the two kind of oils (Table III-2):

$$\mathrm{UTHR} = \overline{\mathrm{ES}_{\mathrm{UO}}} - \mathrm{s}(\mathrm{ES}_{\mathrm{UO}}) = 1.57 \qquad (32)$$

$$\mathrm{LTHR} = \overline{\mathrm{ES}_{\mathrm{HO}}} + \mathrm{s}(\mathrm{ES}_{\mathrm{HO}}) = 1.53 \qquad (33)$$

The margin value between the regions (1) and (2) is given by:

$$\mathrm{UTHR} - \mathrm{LTHR} = 0.04 \qquad (34)$$

In order to obtain a further verification of the correctness of the data processing, the results in terms of MARGIN were also calculated for the mean and FFT magnitude approaches for the echo peak evaluation

In the following table (III-3), the results on the same samples, applying the previously described Mean value and FFT magnitude algorithms to determine the echo peak, are shown:

Table III-3 *The measured parameters on healthy (HO) and unhealthy (UO) hazelnut oil using the Mean and FFT methods for the echo peak determination*

Method	*Mean*		*FFT*	
	ES_{HO}	ES_{UO}	ES_{HO}	ES_{UO}
Sample # 1	1.62	1.82	1.56	1.66
Sample # 2	1.59	1.72	1.49	1.69
Sample # 3	1.57	1.74	1.52	1.78
Sample # 4	1.55	1.69	1.49	1.64
Sample # 5	1.63	1.73	1.54	1.67
Sample # 6	1.57	1.64	1.59	1.48
Sample # 7	1.52	1.80	1.48	1.67
Sample # 8	1.64	1.76	1.47	1.71
Sample # 9	1.60	1.65	1.54	1.56
Mean Value ($\overline{ES}$)	1.59	1.73	1.52	1.65
Std Deviation (s)	0.04	0.06	0.04	0.09
Min	1.52	1.64	1.47	1.48
Max	1.64	1.82	1.59	1.78
MARGIN		0.00		-0.11

As can be seen from table III-3, the two discarded algorithms show a worse performance, in terms of MARGIN than the polynomial regression. Moreover using these approaches the ES value is not able to discriminate the two kinds of oil.

Using the values obtained in the training phase, in order to verify the classification capability of the proposed method, other tests were made on different oil samples, in particular 25 samples of oil extracted from healthy

nuts and 25 samples extracted from unhealthy nuts, for each sample one measurement was done.

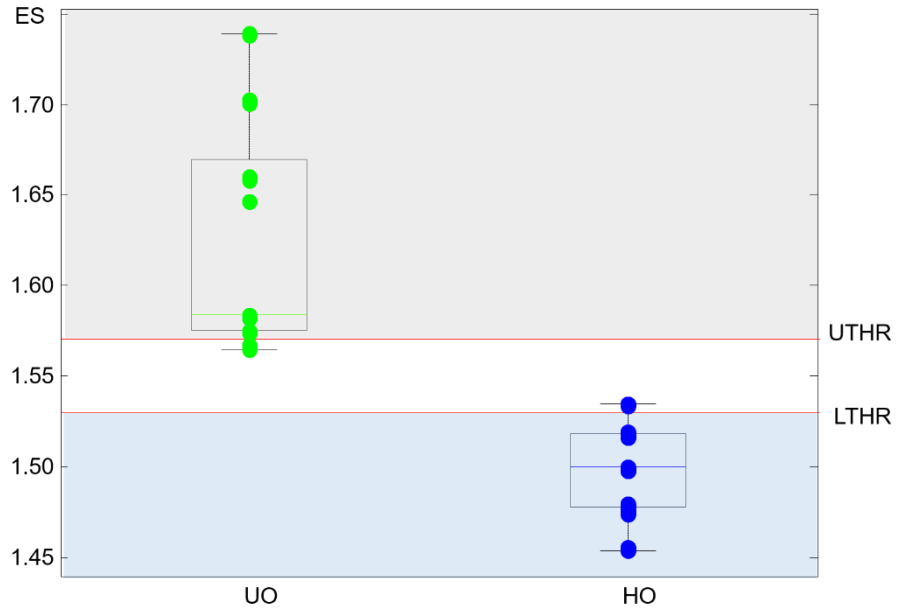

Figure III-13 *Test results on 50 samples (25 healthy hazelnut oil blue points; 25 unhealthy hazelnut oil green points)*

The results of these tests are summarized in the Figure III-13: there has not been any classification error in the fifty samples; seven samples are in the ambiguous region (three unhealthy hazelnut oil and four healthy hazelnut oil), forty-three samples have been correctly detected (twenty-two unhealthy hazelnut oil and twenty-one healthy hazelnut oil).

Multi-exponential approximation

In order to calculate the T_2 distribution on the transversal decay signal, the SVD algorithm has been applied on the same samples used in the former method, using 18 training samples. Applying the algorithm as described in the paragraph III-4 on the 18 training samples, the 9 healthy hazelnut oils show a T_2 distribution different from the 9 unhealthy hazelnut oils. In particular, all samples have a short time constant $T_{2,1}$ = 0.01 s, with same amplitude (maximum percentage of variation lower than 10%). Moreover, all samples also have a second, longer time constant that is different between the healthy and unhealthy hazelnut oils: healthy hazelnut oils have a time constant $T_{2,2}$ = 0.14 s, whereas the unhealthy hazelnut oils have $T_{2,2}$ = 0.15 s.

As for the ES parameter, comparing the results obtained using the polynomial regression algorithm for the echo peak determination with the other two methods, only the first allows discriminating between the two kinds of oil without classification errors (Fig.III-14).

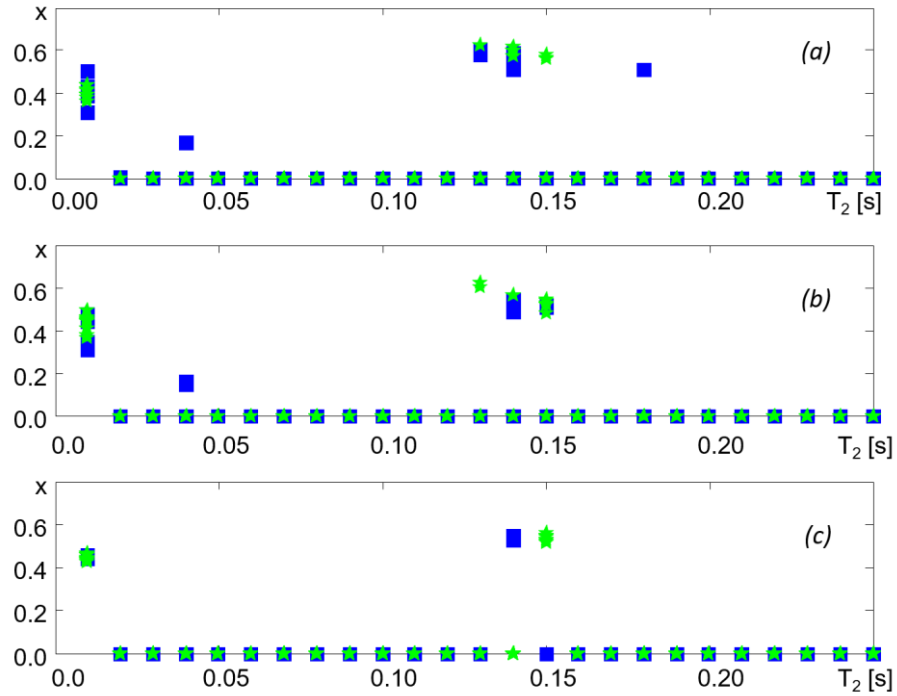

Figure III-14 *T_2 distribution on 9 samples of healthy oil (blue points) and unhealthy oil (green points) using the mean algorithm (a), FFT algorithm (b) and polynomial regression algorithm (c) for the echo peak determination*

Analyzing the previous results, the $T_{2,2}$ time constant can be used as classification parameter for the two kind of oils.

Furthermore, the result obtained about the T_2 distribution is coherent with the result concerning with the ES parameter: a higher level of the ES parameter for the unhealthy hazelnut oil suggests the presence of a higher time constant in the T_2 distribution, and this is confirmed by the new method in which unhealthy hazelnut oils have a slower time constant ($T_{2,2}$ = 0.15 s) greater than the slower time constant of the healthy hazelnut oils ($T_{2,2}$ = 0.14 s).

As for the previous algorithm (ES parameter), to validate the results obtained on the training samples, experimental verification has been carried out on the 50 test samples (25 healthy hazelnut oils, 25 unhealthy hazelnut oils) with the aim to confirm the ability of the SVD algorithm for the T_2 distribution calculation to correctly classify the two kind of oils (Fig.III-15).

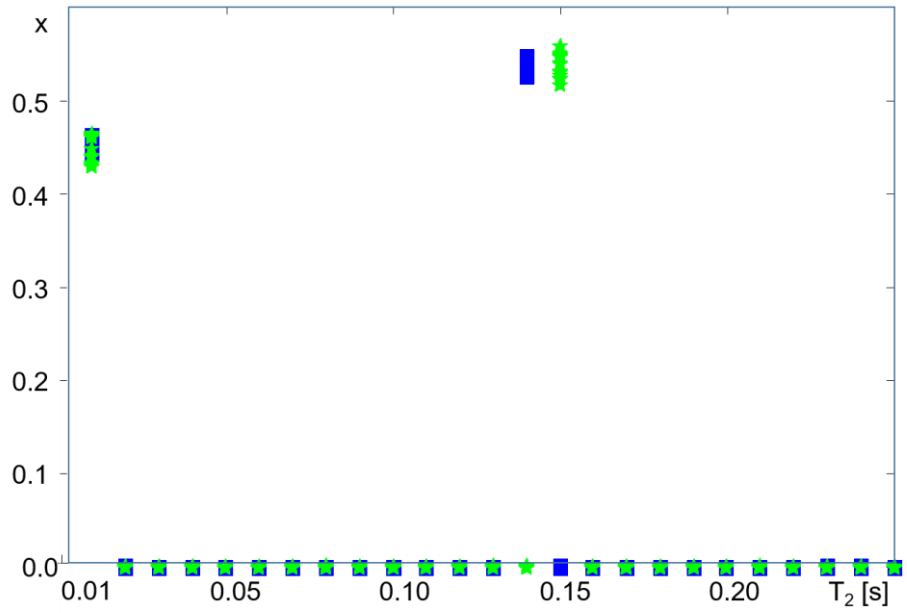

Figure III-15 *T_2 distribution for 25 samples of healthy hazelnut oil (in blue) and 25 samples of unhealthy hazelnut oil (in green)*

Figure III-15 shows the results of the SVD algorithm applied to the signals acquired from 50 samples (25 healthy hazelnut oil and 25 unhealthy hazelnut oil). The T_2 distribution shows that all samples of the two kind of oils are well separated, having a different slower time constant $T_{2,2}$.

In conclusion, comparing the two methods, the T_2 distribution by means of multi-exponential approximation showed a better capability in the oil classification than the ES parameter.

Chapter III

Chapter IV
In-shell hazelnuts characterization

In this chapter, starting from the results related to the CPMG techniques employed in the oil samples classification, the in-shell hazelnuts have been analyzed. Moreover, a more complex classification algorithm, respect to the oil samples, has been developed, due to the presence of more quality properties to detect and more influence factors that affect the system.

As described in the first chapter, hazelnuts quality depends on several parameters, like:

- Moisture content;
- Kernel development;
- Presence of mold or bug disease.

While the analysis on the presence of mold and bug disease can be done both on oil and hazelnuts, the moisture content and the kernel development are only related to the hazelnuts.

The main advantage on carrying on the NMR experiments on oil samples lies on the homogeneous composition of the material. Oil does not contain moisture or any other solid or liquid contaminant. Moreover, it is possible to perform the tests on samples that contain always the same quantity of material.

All these characteristics disappear when working on solid and heterogeneous materials like the whole hazelnuts.

In the following, the data processing on the CMPG signal to determine the quality parameters of the hazelnuts will be depicted (Fig. 4-1).

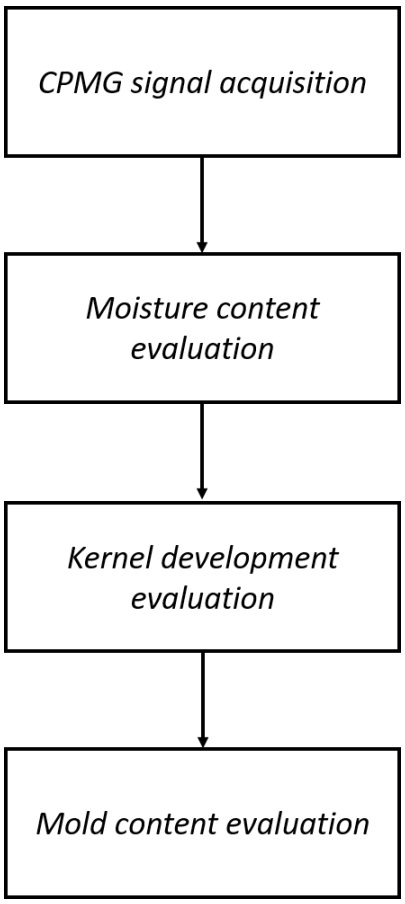

Figure IV-1 *Block diagram of the data processing on the in-shell hazelnuts*

IV.1 CPMG sequence

The CPMG sequence has been carried out with the following parameters:

- 500 kHz sampling frequency;
- π/2 RF pulse: 8 μs;
- π RF pulse: 16 μs;
- Dead time: 16 μs;
- Blanking delay: 3 ms;
- $\tau = 2$ ms;
- $N_{echo} = 60$.

The number of echo is the only parameter that differs from the tests on oil samples. With this value the acquisition time is about 240 ms (Eq. 18).

Discarding the odd echoes, 30 values are available to obtain the transverse relaxation decay (Fig. IV-2)

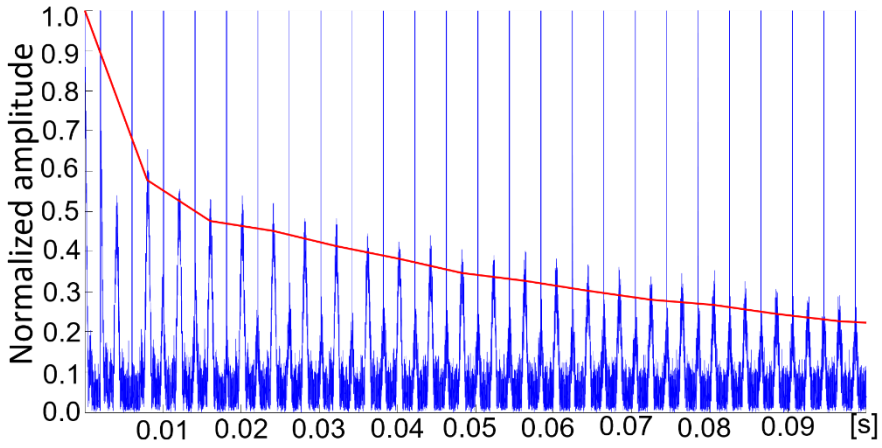

Figure IV-2 *Normalized amplitude of the first 0.10 s of the CPMG signal (in blue) for a nut sample; echo decay envelope (in red).*

IV.2 Moisture content evaluation

The moisture content has been evaluated considering the relationship between the maximum of the FID signal, which is related to the weight of the kernel including the moisture, and the peak of the second echo of the CPMG signal that occurs after about 8 ms (Figure IV-3). At this time, the component of the signal related to the water content is already decayed. This is because bound water (e.g. the moisture inside a solid sample) has a very fast transversal decay, and the contribution to the decay is considered negligible after 7 ms (Todt *et al.*, 2016).

As previously described, the maximum of the FID signal has been determined as the mean value of the first three samples of the acquired signal, while the second echo peak as the maximum of the second order polynomial regression on 160 samples around the position of the peak (Fig. IV-3).

In order to evaluate the relationships between the 2nd echo and the FID signal, well-dried hazelnuts are analyzed to define the condition useful to discriminate dry from wet hazelnuts. In particular, since the maximum allowed moisture content is 6%, 100 samples of hazelnuts have been dried to a value lower than the limit according to the AOAC 925.40 method (Horwitz, 2000).

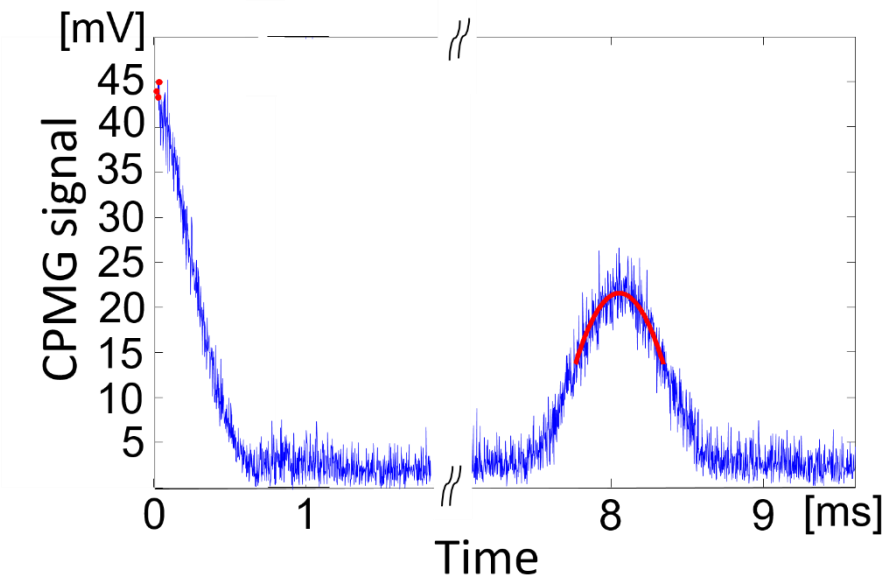

Figure IV-3 *First exponential decay (FID signal) and second echo of the CPMG signal; first three samples of the FID signal second order polynomial regression on the second echo (in red)*

Then, the CPMG signal of the dried nuts has been studied; in Figure IV-4 the measured amplitudes of 2nd echo respect to the FID signal are reported. On these data, the linear least square regression method has been applied obtaining the *m* and *b* parameters of the regression line (Fig. IV-4):

$$A_{echo} = m \cdot A_{FID} + b \qquad (35)$$

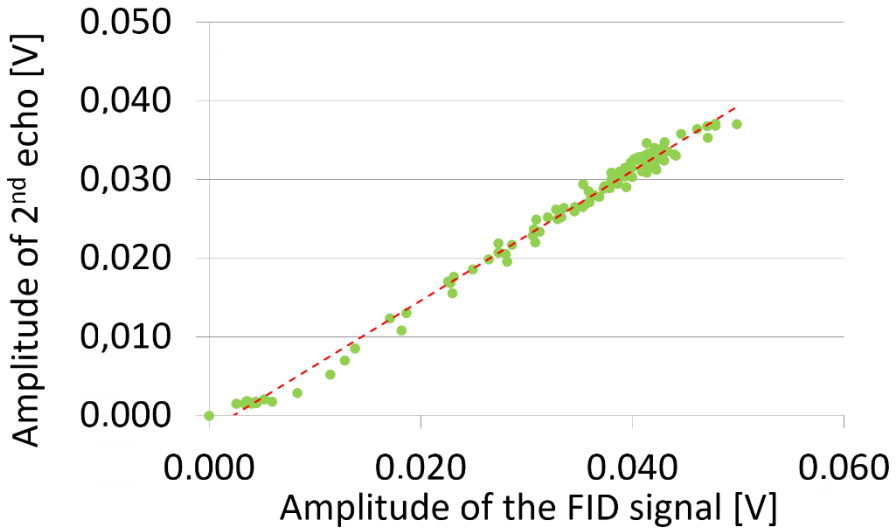

Figure IV-4 *Relation between the amplitude of the FID signal and the 2nd echo peak for 100 dried hazelnuts samples (in green) and regression line (eq.35) (in red)*

A statistical analysis has been carried out calculating the standard deviation σ_m and σ_b related, respectively, to the parameter *m* and *b*. At this point, a dry threshold line has been defined as (Fig. IV-5):

$$A_{echo}^{THR} = (m - 3\sigma_m)A_{FID}^{THR} + (b - 3\sigma_b) \qquad (36)$$

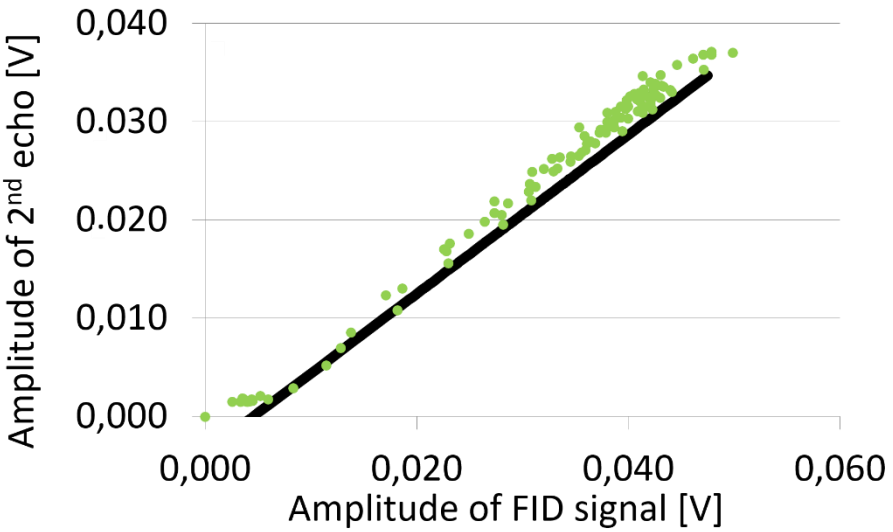

Figure IV-5 *Relation between the amplitude of the FID signal and the 2nd echo peak for 100 dried hazelnuts samples (in green); dry threshold line (in black)*

The threshold line and the relation between the amplitude of the FID signal and the second echo peak for 100 dried samples are reported in fig. IV-5.

IV-3 Kernel development evaluation

According to the OECD standard for in-shell hazelnuts (OECD, 2011), the kernel development is one of the features that determines the nut quality. Only hazelnuts with a well-developed kernel have to be selected. In particular, the hazelnut kernel should fill at least the 50% of the shell cavity and has not to be desiccated. These properties can be related to the weight of the kernel (Fig.IV-6). Unfortunately, the weight of kernel cannot be determined from the weight of the whole hazelnut because the weight of the shell depends on several variables related to the soil condition during the growth.

Figure IV-6 shows the relationship between the weight of the whole hazelnut and its kernel for 100 samples. It can be seen that the weight of the hazelnut does not allow the discrimination between well and not-well developed kernels, while the weight of the kernel is more appropriate. This is because the weight of the shell significantly affects the total weight of the nut.

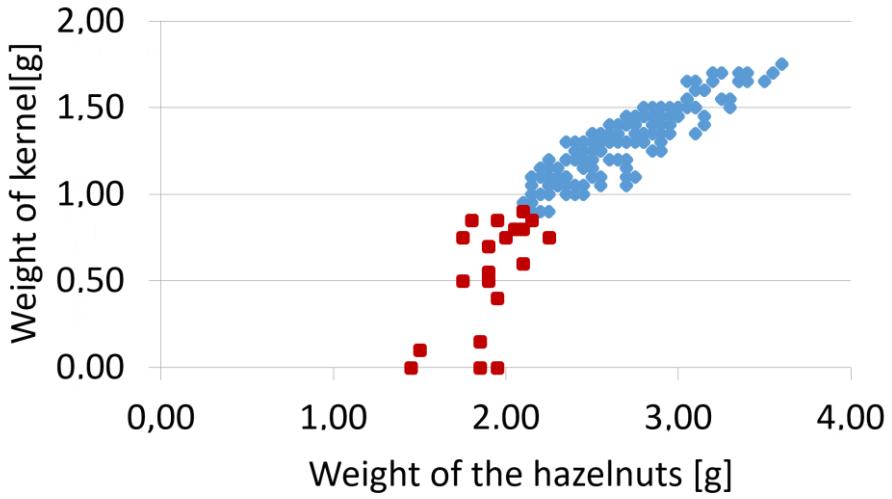

Figure IV-6 *Relation between weight of the hazelnut and weight of the kernel for fully developed kernels (in blue) and not well developed kernels (in red).*

The in-shell hazelnuts have been tested acquiring the CPMG signal and weighting it by means of a scale with a resolution of 0.05 g (Fig. IV-7).

Then, the nuts have been unshelled and classified by means of visual inspection of the kernel. Finally, using the same scale employed for the whole nuts, the hazelnut kernels have been weighted.

Figure IV-7 *Scale with a 0.05 g resolution*

The weight of the hazelnut kernel can be related to the amplitude of the FID signal. This is because the kernel contains ^{1}H molecules that react to the NMR signal, while the shell is mainly composed by inorganic molecules that are not excited at the resonance frequency of the ^{1}H probe. Only the dried hazelnuts have been considered in order to ignore the contribution to the signal due to the ^{1}H molecules related to the moisture.

The relationship between the amplitude of the FID signal and the weight of the kernel on the same 100 samples is shown in Figure IV-8. As can be seen, the two set of data (well-developed and not well-developed kernels) are separated, so a threshold on the amplitude of the FID signal has been defined to discriminate between full and insufficient developed kernels. In particular, a margin of 5% on the minimum value of the amplitude of the FID signal related to the full-developed hazelnuts has been considered to set the threshold. In this way, all the full-developed kernels have been correctly classified. Over the 15 not-well developed kernels, only one has been mis-detected (see Figure IV-8).

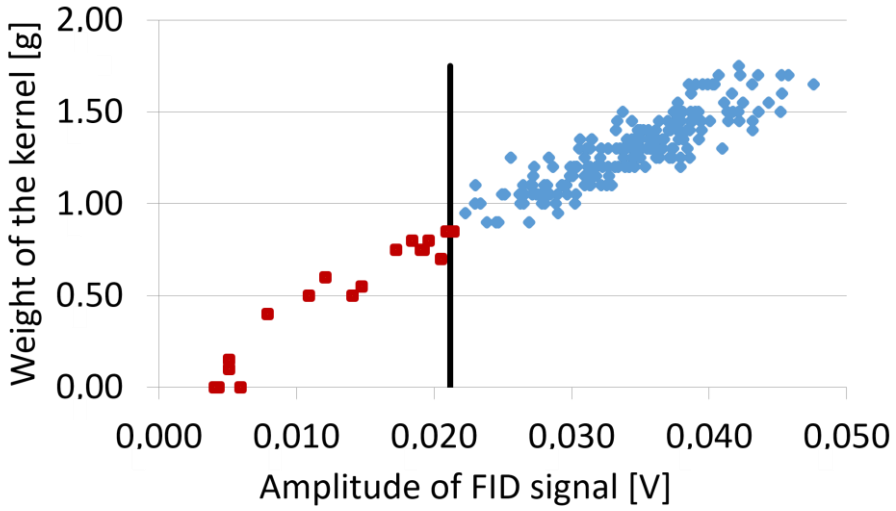

Figure IV-8 *Relation between the amplitude of the FID signal and weight of the kernel for fully developed kernels (in blue) and not well developed kernels (in red). Threshold calculated with a margin of 5% on the lowest amplitude of the FID signal (in black).*

IV-4 T_2 distribution calculation using the SVD algorithm

Following the moisture content and the kernel development evaluation, on the hazelnuts that have not been discarded, the analysis on the T_2 distribution is carried out in order to evaluate the presence of mold

The echo decay envelope is composed by 31 points (n=31 in Eq. 26), corresponding to the amplitude of the fast decay signal and 30 echoes (the even ones). The T_2 range has been chosen from 0.01 s to 0.30 s and it has been divided in 30 logarithmically spaced time constants (m=30 in Eq. 26).

The figure IV-9 shows the T_2 spectrum calculated using the SVD algorithm on several samples of healthy (in blue) and unhealthy (in green) hazelnuts. Both kinds of hazelnuts showed the same three T_2 components with the amplitudes in the same range. Differently of what happened in the case of the oil sample analysis, in the case of the in-shell hazelnuts this method does not allow discriminating between healthy and unhealthy samples.

A reason can be that, because of the presence of the shell and other influence factors, a lower resolution in the T_2 spectrum calculation is required to be able to find out the differences in the spectral components of the two kinds of samples.

Starting from these consideration, a method to improve the resolution in the T_2 distribution has been proposed. Moreover, a classification algorithm has been developed. It is based on several classification steps, in which for

each step a single feature is detected and the hazelnuts affected by defects are discarded. In particular, in the first step the moisture content inside the hazelnuts is evaluated; the second step evaluates the kernel development and, at the end, the T_2 distribution of the well-formed and correctly dried hazelnuts is determined to classify the healthy and unhealthy kernels. In this way, the analysis on the T_2 distribution is made on hazelnuts with more homogeneous characteristics in terms of moisture and kernel dimension, in order to reduce the effect of the influence factors.

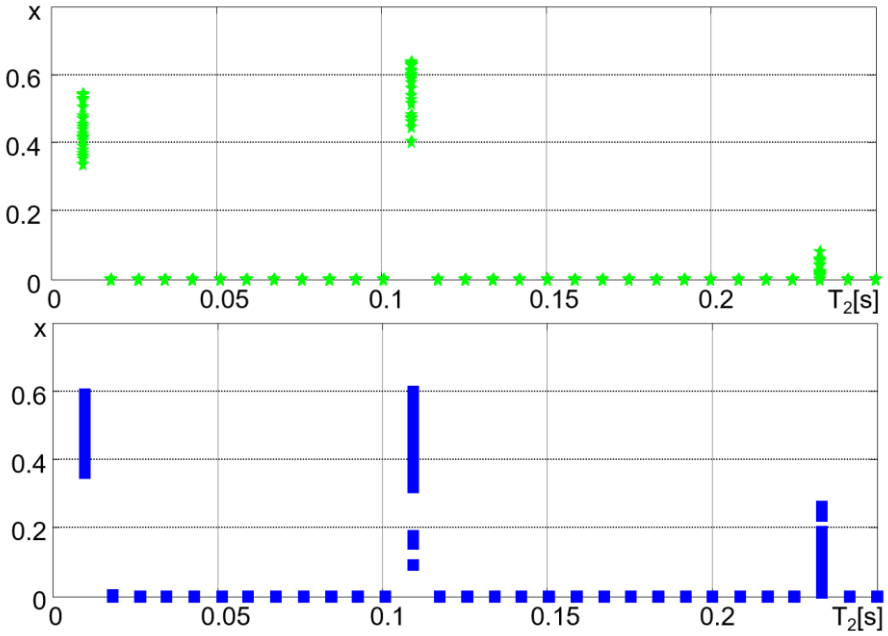

Figure IV-9 T_2 spectrum for healthy hazelnuts (in blue) and unhealthy hazelnuts (in green) calculated by means of the SVD algorithm

IV-5 T_2 distribution

As described in the previous paragraph, the SVD algorithm does not allow discriminating between healthy and unhealthy hazelnuts that show the same three time constant. Using this algorithm, the resolution on the T_2 axis is determined by the number of the pre-assigned time constants in the selected range. This number is limited by the available number of points of the transverse relaxation decay (the vector Y in eq. 27). In order to obtain a solution of eq. 27, if n is the dimension of the Y vector, the number m of the pre-assigned time constants has to satisfy the condition:

$$m \leq n \qquad (37)$$

The way to increase the number of time constants, and then to reduce the T_2 resolution, is increasing the number of points of the acquired transverse relaxation decay. This can be done acquiring a greater number of echo signals, because each echo peak corresponds to a point of the transverse relaxation decay, then using a lower value of τ.

In this work a different approach has been used exploiting the results obtained using the SVD algorithm: both kinds of hazelnuts showed three time constant, so a model with three exponential decays has been applied to represent the transverse relaxation decay related to the hazelnuts (Fig. IV-10):

$$y(t_i) = \sum_{j=1}^{3} x_j e^{-\frac{t_i}{T_{2j}}} \qquad (38)$$

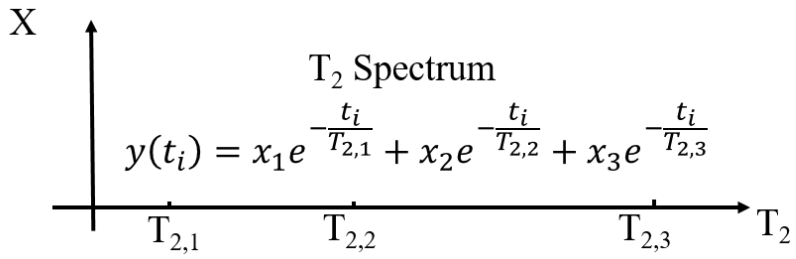

Figure IV-10 *Modeling of the transverse relaxation decay using a three component multi-exponential function*

At this point, a solution of eq. 38, in least square sense, using a nonlinear fitting algorithm has been searched, considering a range of variation of T_{2j} around the values obtained using the SVD algorithm. In this way, a better resolution on the T_2 distribution has been achieved.

More in detail, the ranges of the unknown parameters have been set.

For the three time constants:

- $T_{2,1} \in [0.001\text{s}, 0.010\text{s}]$;
- $T_{2,2} \in [0.055\text{s}, 0.065\text{s}]$;
- $T_{2,3} = 0.27\ s$

The time interval of the time constants has been determined considering a range similar to the T_2 resolution of the SVD algorithm. In this way, the T_2 resolution is considerably reduced; In particular, this has been done for the first two exponential component, while the third component showed a low sensitivity to the time constant variation and, for this reason, $T_{2,3}$ has been forced to a fixed value.

While for the amplitudes the non-negativity condition has been forced:

- $x_i \in [0,1]\ \ i = 1,2,3$

The ranges of the of the time constants have been determined

The non-linear least square solver finds the best solution in the ranges of the parameters with the following condition:

$$\min_x \sum_i (y_i^m - y_i)^2 \qquad (39)$$

Where $y_i^m = y^m(t_i)$ are the measured points of the transverse relaxation decay, y_i are the fitting data.

Using this method, a very low resolution on the T_2 axis has been achieved, only considering the range of interest. This allows avoiding a computational overload due to the acquisition of more points to perform the SVD algorithm with the same resolution.

In figure IV-11 the T_2 spectrum for healthy and unhealthy hazelnuts is shown.

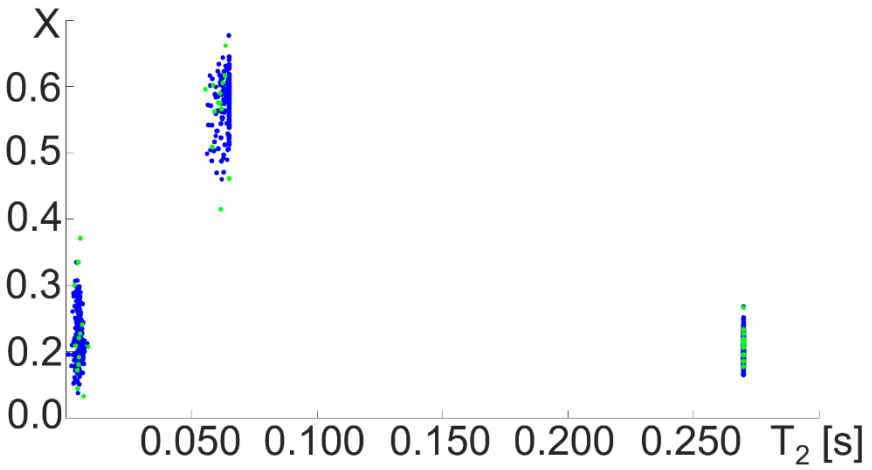

Figure IV-11 *T_2 spectrum for healthy hazelnuts (in blue) and unhealthy hazelnuts (in green) calculated by means of the non-linear least square fitting*

A detailed analysis and the classification algorithm for the in-shell hazelnuts, considering the several properties previously analyzed, will be presented in the next chapter.

Chapter V
Classification algorithm

In this chapter, a classification algorithm for the in-shell hazelnuts, based on the data processing described in the previous chapter, is proposed. It is composed by three steps:

1. Evaluation of the moisture content;
2. Evaluation of the kernel development;
3. Evaluation of the presence of mold.

In the first step, the hazelnuts with a moisture content greater than the limit imposed by the quality standards are discarded; the correctly dried hazelnuts are then evaluated on the basis of the kernel development. In this step the following defects are detected and discarded:

- Empty hazelnuts;
- Not well developed kernels;
- Desiccated kernels.

Finally, in the third step the presence of mold is evaluated and the healthy hazelnuts are selected.

The three steps are carried out on a single signal, acquired by means of the CPMG sequence, so for each sample, only one CPMG sequence is required.

The tests were conducted on hazelnuts supplied by “Consorzio di Tutela Nocciola di Giffoni I.G.P.”. For each sample of hazelnut with shell, the following procedure was used to perform the test:

1. Insertion of the hazelnut with shell inside the probe;
2. Acquisition of the signal by means of CPMG sequence;
3. Signal processing on the data acquired by the instrument and classification of the hazelnut by means of classification algorithm; extraction of the hazelnut from the probe;
4. Breakage of the shell and classification of the hazelnut as healthy or unhealthy by means of vision inspection;

Verification of the classification algorithm comparing the results with the vision inspection method.

In the following figure, a block diagram of the experimental set up is shown. The tests were carried out on 300 samples of hazelnuts, 220 dried samples and 80 fresh samples.

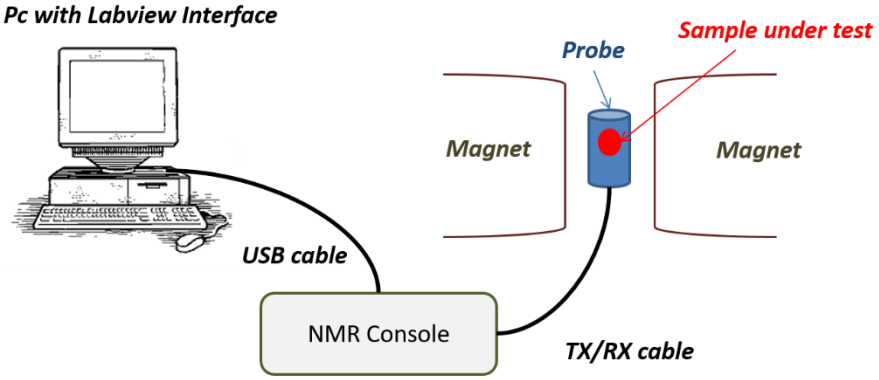

Figure V-1 *Experimental setup*

In the next paragraphs, each steps of the classification algorithm will be described highlighting the classification performance.

V.1 Moisture content evaluation

As described in the previous chapter, the moisture content evaluation is made analyzing the relationship between the amplitude of the FID signal (A_{FID}) and the amplitude of the 2nd echo of the CPMG sequence (Fig. IV-3).

According to the threshold line defined in (eq. 36), the well-dried hazelnuts have to satisfy the following condition:

$$A_{echo} > (m - 3\sigma_m)A_{FID} + (b - 3\sigma_b) \qquad (40)$$

In the following figures, the results obtained on 220 dried hazelnuts and 80 fresh hazelnuts are shown.

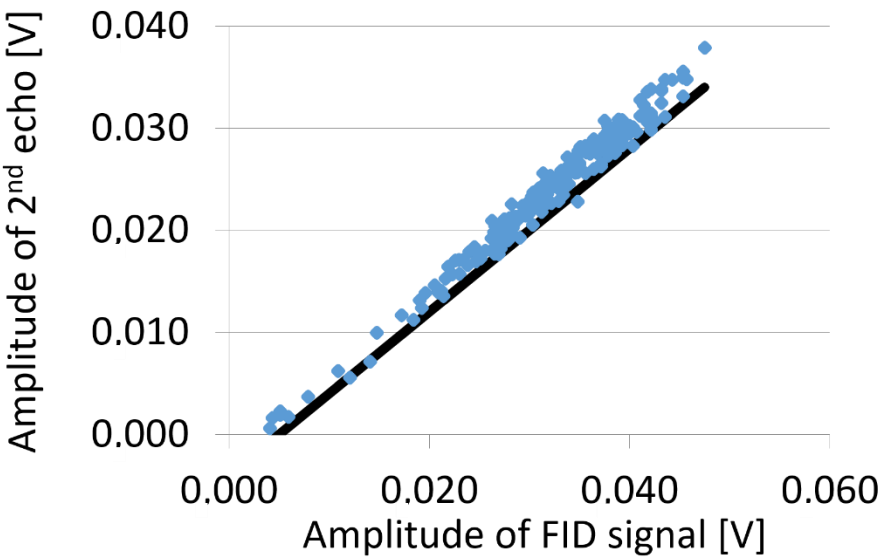

Figure V-2 *Comparison among the dry threshold line (in black) and the amplitudes of the FID signal and the 2nd echo peak for 220 dried hazelnuts samples (in blue).*

As can be seen from the figure V-2, one in 220 samples of dried hazelnuts has been mis-detected.

Figure V-3 shows the results on the 80 fresh hazelnuts. All the samples have been correctly detected. Moreover, the fresh samples with a value close to the threshold line have a lower moisture content than others. This is due to the drying effect of the sun, so the moisture content can vary on the basis of the weather condition during the harvest and the storage.

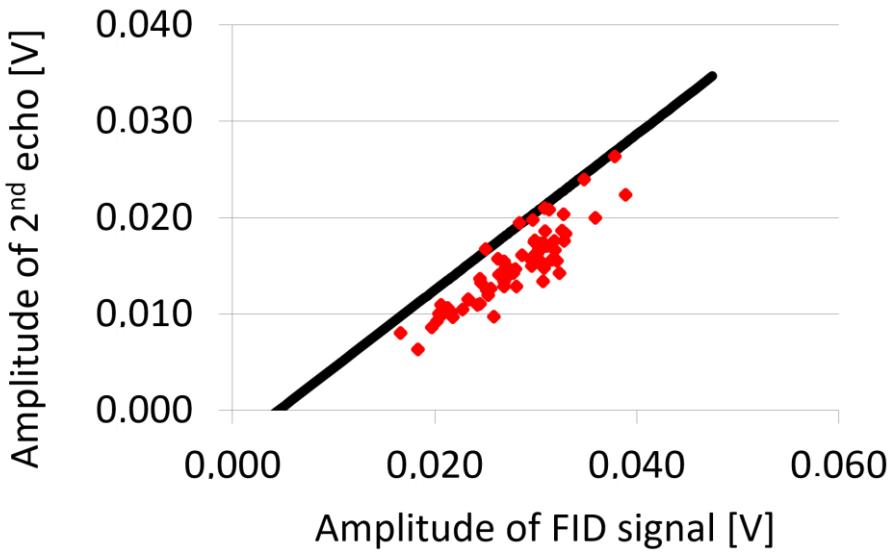

Figure V-3 *Comparison among the dry threshold line (in black) and the amplitudes of the FID signal and the 2nd echo peak for 80 fresh hazelnuts samples (in red).*

V.2 Kernel development evaluation

In this paragraph, the analysis on the kernel development on the 219 hazelnuts correctly classified as dried in the previous step, is described. It is made exploiting the threshold on the amplitude of the FID signal defined in the paragraph IV-3.

In particular, three category of defects: the well-formed nuts and the blank, insufficiently developed and desiccated have been included in the not well developed kernel

The well-formed hazelnuts have to satisfy the condition:

$$A_{FID} > threshold \qquad (41)$$

In the following figure (V-4), the results related to the 219 samples selected from the previous step are shown.

As can be seen, all the well-developed kernels have been detected and only one in 25 not-well developed kernels has been not correctly classified.

Classification algorithm

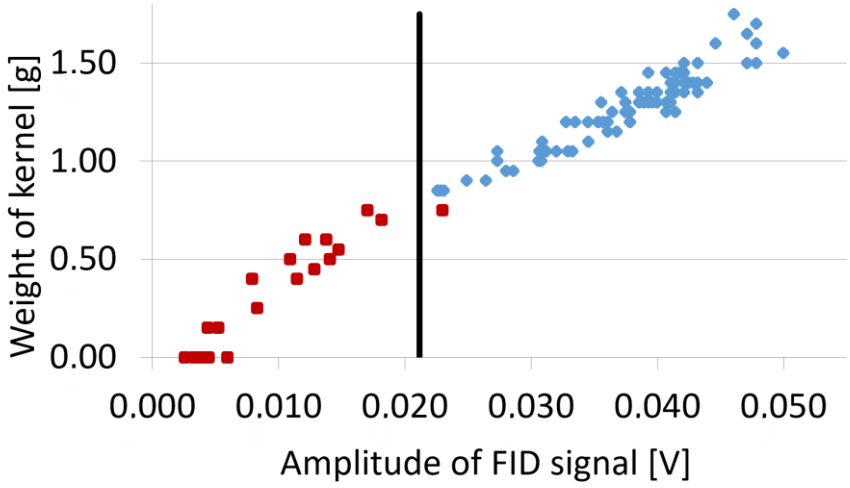

Figure V-4 *Relation between the amplitude of the FID signal and weight of the kernel for 219 hazelnuts samples: fully developed kernels (in blue) and not well developed kernels (in red)*

V.3 Mold presence evaluation

The analysis of the mold presence is made on the dried and well-developed hazelnuts. It exploits the differences in the T_2 spectrum between the unhealthy (with presence of mold) and healthy hazelnuts. The T_2 spectrum is obtained as described in the paragraph IV-5.

Analyzing the T_2 spectrum in figure IV-11, the healthy hazelnuts exhibited a greater value of the second time constant than the unhealthy ones, but the two categories are not well separated.

Moreover, the unhealthy hazelnuts are characterized by two different behaviors, compared to the healthy nuts:

1. The unhealthy hazelnuts that are not completely developed, because the damage appeared during the kernel growth, show an higher amplitude (x_1) of the first exponential component and, at the same time, a lower amplitude (x_2) of the second exponential component;
2. The unhealthy hazelnuts that are completely developed, for example because the mold developed during the storage due to an high level of moisture, exhibit a higher amplitude (x_2) of the second exponential component with a lower time constant (T_2).

Starting from these considerations, the two kinds of unhealthy hazelnuts have been analyzed.

V.3.1 Unhealthy hazelnuts not completely developed

The T_2 spectrum related to the first two components of the exponential decay is shown in figure V-5.

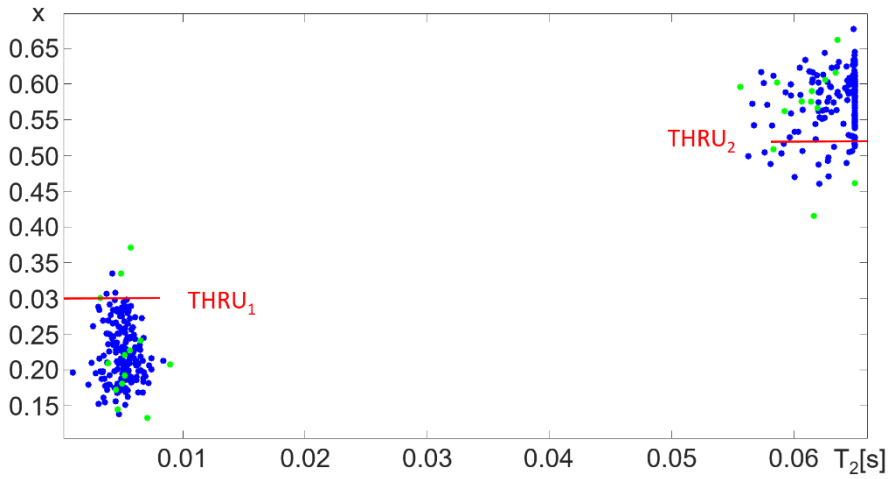

Figure V-5 *T_2 spectrum related to the first two components of the transverse relaxation decay. Unhealthy thresholds ($THRU_1$ and $THRU_2$) in red.*

In order to discriminate the unhealthy hazelnuts not completely developed, an unhealthy condition, based on the amplitude of the first and the second exponential components, has been defined:

$$(x_1 > THRU_1)\ and\ (x_2 < THRU_2) \qquad (42)$$

Where $THRU_1$ and $THRU_2$ are the thresholds on the amplitude x_1 and x_2, respectively.

The unhealthy condition relies on the property of the unhealthy hazelnuts that are not completely developed to have a fat content lower than the typical value contained in the healthy hazelnuts (Memoli *et al.*, 2017). The effect on the relaxation curve is a faster decay that results in a greater amplitude of the first exponential decay and a lower amplitude of the second exponential decay in the T_2 spectrum.

The unhealthy condition allowed discarding all the unhealthy hazelnuts not completely developed, with lower than 2% of healthy hazelnuts wrongly classified as unhealthy.

V.3.2 Unhealthy hazelnuts completely developed

The analysis of the T_2 spectrum of the hazelnuts with this kind of defect revealed that the significant difference with the healthy nuts can be observed on the second component of the transverse relaxation decay. In particular, the unhealthy hazelnuts tend to exhibit a lower time constant (T_2) and, at the same time, a higher amplitude (x_2) than the healthy hazelnuts. Moreover, most of the healthy hazelnuts exhibited a time constant very close to the upper bound of the range. Starting from this consideration, a two steps procedure was set up: (i) a healthy condition has been defined to detect the healthy hazelnuts:

$$T_{2,2} > THRH \qquad (43)$$

The threshold has been chosen considering the maximum value of the time constant related to the unhealthy hazelnuts adding a margin of 5% on that value.

The figure V-6 shows the time constant and the amplitude related to the second exponential component for the healthy and unhealthy hazelnuts and the healthy threshold.

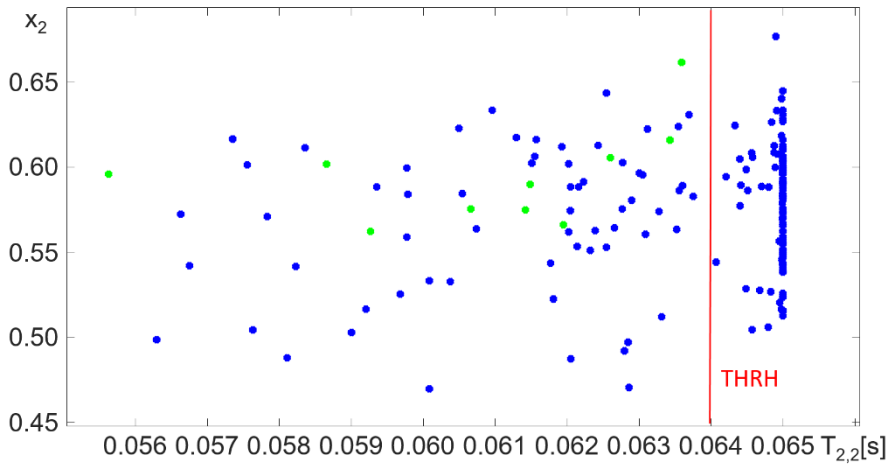

Figure V-6 *Healthy condition on the T_2 value of the second exponential component of the transverse relaxation decay for healthy (in blue) and unhealthy (in green) hazelnuts.*

As can be seen from the previous figure, the healthy threshold allows correctly classifying most of the healthy hazelnuts, but several samples fall in the unhealthy area ($T_{2,2}$ < THRH). In this case, it is not possible to discriminate the two kind of nuts on the basis of the amplitude or the time constant. Then, in order to try highlighting the differences among the two classes of nuts

included in the unhealthy area, (ii) an analysis on the relation between amplitude and time constant has been carried out. In particular, the best selectivity was obtained considering the ratio between the amplitude and the square value of the time constant. In Fig. V-7 the ratio between the amplitude and the square and the time constant of the second component versus the ratio between the amplitude and the square and the time constant of the first component are reported for the healthy (in blue) and unhealthy (in green) hazelnuts.

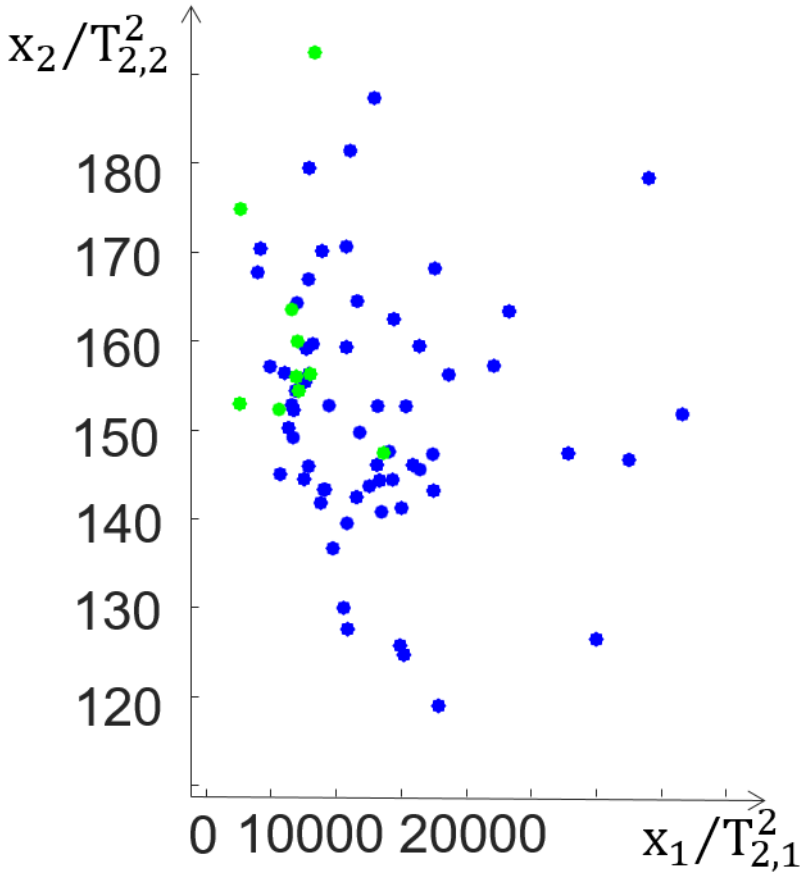

Figure V-7 *Relation between ratio of the amplitude and the square of the time constant [s^{-2}] of the first two exponential components for the healthy (in blue) and unhealthy (in green) hazelnuts*

As can be seen from the figure V-7, the two categories of nuts are still not completely separated, but the greatest sensitivity is observed on the $x_2/T_{2,2}^2$

parameter. In order to set a threshold on this parameter, an approach based on ROC curve has been adopted. In detail, in order to define the threshold for the healthy area (eq. 44) the global performances of the classification procedure have been taken into account, as described in V.3.3.

$$\frac{x_2}{T^2_{2,2}} < THR \qquad (44)$$

V.3.3 Threshold setting by means of ROC curve analysis

The ROC curves (Fawcett, 2006) have been carried out considering as unhealthy hazelnuts all the samples with defects: empty, not well developed, desiccated and with mold. Only the moisture content has not been considered because, even if it is a quality parameter, it is not a defect of the hazelnut.

The analysis has been carried out considering two classes:

- Unhealthy hazelnuts: positive test;
- Healthy hazelnuts: negative test.

The parameters used to evaluate the classification errors are described in the following table:

Table V-1 *Classification error parameters*

Classification parameter	Description
True positive, TP	Unhealthy sample correctly classified
False positive, FP	Healthy sample classified as unhealthy
True negative, TN	Healthy sample correctly classified
False negative, FN	Unhealthy sample classified as healthy

From the parameters defined in table V-1, the following measures can be defined:

$$TPR = \frac{TP}{TP+FN} = \frac{unhealthy\ samples\ correctly\ classified}{total\ number\ of\ unhealthy\ samples} \qquad (45)$$

$$FPR = \frac{FP}{FP+TN} = \frac{healthy\ samples\ classified\ as\ unhealthy}{total\ number\ of\ healthy\ samples} \qquad (46)$$

$$TNR = \frac{TN}{FP+TN} = \frac{healthy\ samples\ correctly\ classified}{total\ number\ of\ healthy\ samples} \qquad (47)$$

The true positive rate (TPR) (eq. 45), also named sensitivity or recall, represents the capability of the system to detect the unhealthy samples. The false positive rate (FPR) (eq.46) is a measure of the false alarms, namely the healthy samples wrongly classified as unhealthy. The true negative rate (TNR) (eq.47) is the specificity, which represents the capability of the system to detect the healthy hazelnuts.

The ROC curve is a two-dimensional graph in which the y-axis is represented by the TPR and x-axis is represented by the FPR (Fawcett, 2006).

It allows depicting the tradeoff between the benefits, represented by the true positive rate, and the costs, represented by the false positive rate (Fawcett, 2006). Each point of the ROC curve is related to a value of the threshold to set, and it is obtained calculating the TPR and FPR using that value of the threshold.

Figure V-8 shows the ROC curve related to the classification of the hazelnuts obtained varying the threshold THR.

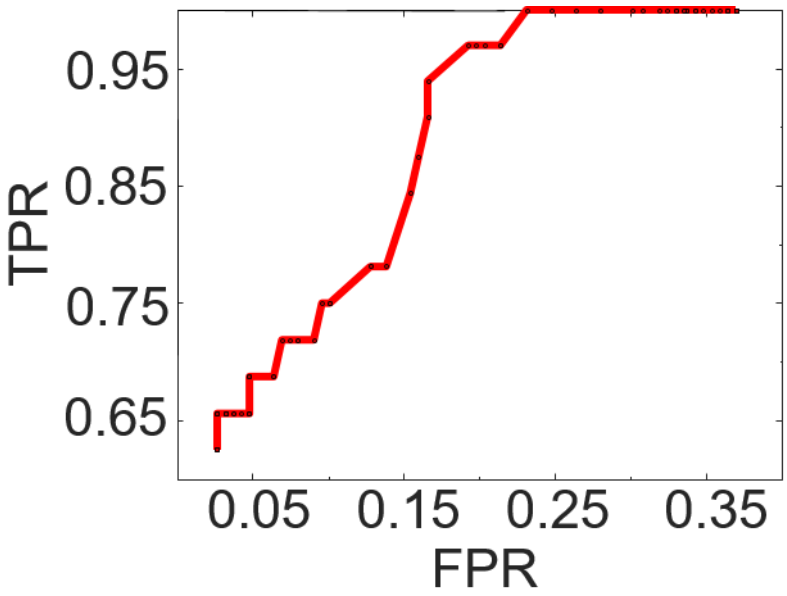

Figure V-8 *ROC curve related to the threshold THR*

The best value of the threshold (THR) is the value corresponding to the highest TPR considering the maximum acceptable FPR.

Figure V-9 shows the point of the ROC curve corresponding to the selected value of the threshold, while the result of the classification using the selected threshold is depicted in figure V-10.

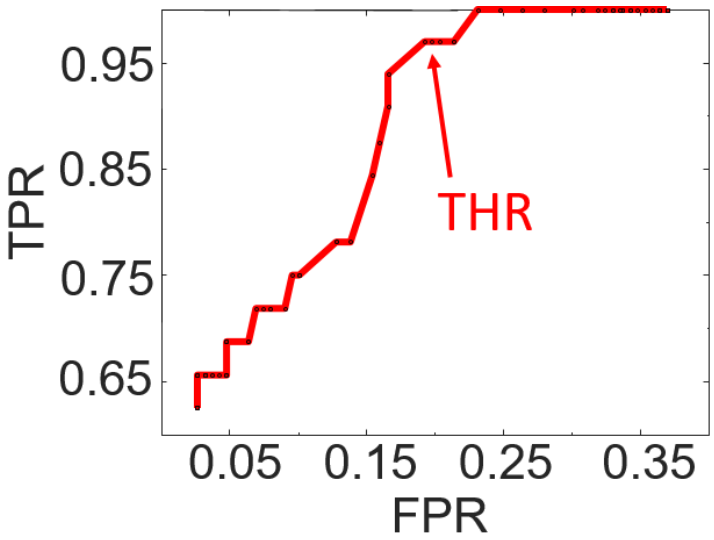

Figure V-9 *ROC curve and the selected threshold THR*

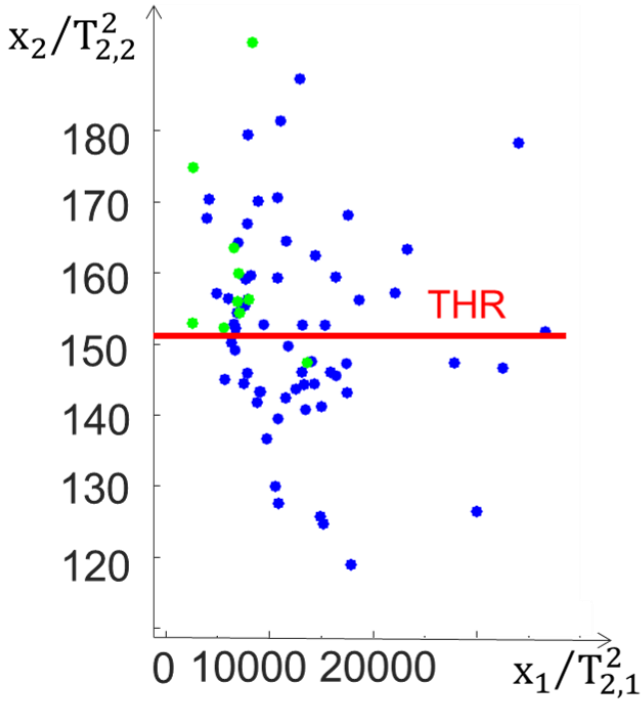

Figure V-10 *Relation between the amplitude and the square of the time constant [s^{-2}] for the healthy (in blue) and unhealthy (in green) hazelnuts. Threshold (THR) in red.*

In this way, a sensitivity (TPR) of 97.0% and a specificity (TNR) of 80.8% has been obtain. The classification algorithm exhibited a high sensitivity, so a high capability to detect the unhealthy hazelnuts, with a specificity around the 80%.

A further analysis has been made on the threshold, in order to verify the effectiveness. A bootstrap resampling method has been used to obtain new datasets starting from the initial dataset. A description this method and the obtained results can be found in the next paragraph.

V.3.4 Threshold validation using a bootstrap method

A further analysis has been made on the whole procedure including the selected final threshold, in order to verify its effectiveness; because of the reduced size of the whole sample, a bootstrap resampling method has been used to obtain new datasets starting from the initial dataset. The bootstrap is a computationally intensive statistical technique that allows one to make

inferences from data without making strong distributional assumptions about the statistic that is calculated and/or the data.
The main idea is that a number of new data sets, which are referred to as bootstrap samples, can be generated from the initial data set by sampling with replacement. With this resampling scheme, these distributions can be seen as approximations to the true distributions of the estimators, and then a good estimate can be obtained of the distribution of a statistics of interest, such as bias, standard deviation and so on (Liguori *et al.*, 2017). In this work, the bootstrap method has been used to obtain 100 datasets starting from the initial dataset and computing the classification performance in order to verify the effectiveness of the selected threshold (THR). A block diagram of the adopted procedure is shown in the following figure.

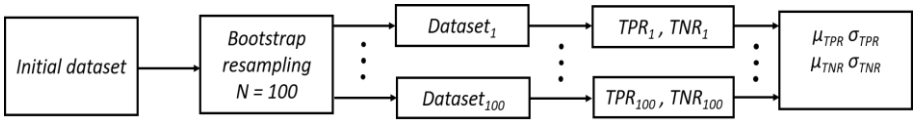

Figure V-11 *Block diagram of the bootstrap resampling method*

The adopted procedure (Fig. V-11) is composed by the following steps:

1. Bootstrap resampling on the initial dataset (N = 100); in this way, 100 new datasets are generated from the initial one, with a different distribution of the classes (healthy, unhealthy):
2. Using the classification algorithm with the previously described thresholds: calculation of the classification performance on the datasets generated using the bootstrap (TPR_i and TNR_i for i=1,…100);
3. Calculation of the mean value (μ) and the standard deviation (σ) of the TPR and TNR indexes.

Following this approach each generated dataset has a very different distribution of the several classes of samples; starting with a number of hazelnuts with mold equal to 13 in the initial dataset, a mean value of 15.8 and a standard deviation of 4.6 have been observed in the 100 generated datasets.

The results of the estimated TPR and TNR indexes are shown in the table V-2.

Table V-2 *Classification performances using 100 datasets generated with bootstrap resampling*

	TPR	**TNR**
μ	97.6%	80.8%
σ	2.7%	2.9%

As can be seen from the table V-2, the mean value of the sensitivity (TPR) and the specificity (TNR) calculated on the 100 datasets generated using the

bootstrap resampling are similar to the values obtained on the initial database and the standard deviation is lower than 3% for both. According to these results, despite a large variation of the number of unhealthy samples in the datasets, the classification performances remain similar to the initial dataset. For this reason, the threshold THR has been considered reliable.

V.4 Analysis of the algorithms execution time

An analysis of the execution time of the algorithms employed in the classification procedure has been carried out in order to verify that they are compatible with the requirements of an in-line application. For each acquisition, the CPMG sequence is carried out and the data processing is performed on the acquired signal. As described in detail in chapter II, the CPMG sequence starts after the blanking delay and it is composed by a train of TX pulses. After the first TX pulse, a FID signal is received, while an echo signal is received after each following TX pulses. The time interval between the echoes is determined by the τ parameter, and it is equal to 2τ.

In particular, in the proposed system the blanking delay is equal to 3 ms; the FID signal is acquired during the time τ (2 ms) and then the echoes are acquired after each TX pulses. The time interval between two TX pulses or, similarly, between two echoes is 2τ (4 ms). Sixty echoes are acquired for each sequence, therefore, the total time to perform a single CPMG sequence is around 245 ms (Fig. V-12).

The processing is composed by two algorithms: (i) echo peak detection by means of second order polynomial regression; (ii) T_2 distribution calculation by means of non-linear mean square exponential fitting. In order to estimate the processing time 100 runs were made on a personal computer with CPU Intel Core i5 2.7 GHz and Window 10 64 bit operating system.

Table V-3 shows the mean value and the standard deviation of the execution time (T_e) related to the three algorithms, calculated over 100 acquisitions.

Table V-3 *Mean value and standard deviation of the execution time of the algorithms calculated over 100 acquisitions*

Algorithm description	$\overline{\mathbf{T_e}}$ **[ms]**	**s(T_e) [ms]**
Echo peak detection with polynomial regression	0.55	0.22
T_2 distribution calculation with least square fitting	20.7	2.6

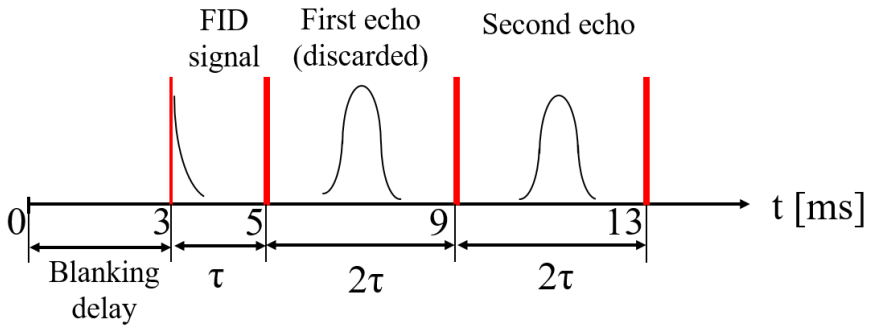

Figure V-12 *Time diagram of the CPMG sequence (first two echo signal).TX pulses in red.*

Comparing the acquisition and execution time related to the echo peak detection, the time interval between two echoes is 8 ms (considering that the odd echoes are discarded); it is much higher than the execution time of the polynomial regression algorithm, so a real-time detection of the echo peaks is allowed.

The non-linear least square fitting is very time consuming but compared with the total acquisition time is an order of magnitude lower, then it not represent a limit.

Chapter V

Chapter VI
Conclusions

In this work, a method for the quality detection of the in-shell hazelnuts has been proposed. It is based on the low field NMR, which is a non-destructive technique employed in the quality control of food, especially in offline laboratory applications. This research project focuses on an in-line industrial application of the NMR system that has to be able to detect the hidden defects of the hazelnuts in order to help the farmers and the producers to meet the quality requirements imposed by the international standards and rules. For this aim, an analysis of the main NMR techniques and systems has been carried out, highlighting the influence parameters that affect the measurements and the issues related to the execution time of the data processing. Moreover, a permanent magnet and probe specifically designed for this application have been used in order to be able to host a whole nut. The work then focused on the signal processing, introducing pre-processing and processing algorithms in order to obtain execution times suitable to an in-line application. The proposed classification algorithm is based on the analysis of the transverse relaxation decay, which is obtained with the CPMG sequence. In particular, three different steps of the algorithm have been defined. In each step a single quality property of the hazelnuts is analyzed: the kernel development, that allows detecting the empty, not well developed and desiccated nuts, is evaluated considering the maximum amplitude of the acquired signal; the moisture content, which is related to the mold development, is evaluated analyzing the relation between the maximum amplitude of the CPMG signal and the second echo peak. Finally, the mold development is evaluated analyzing the T_2 spectrum of the CPMG signal, which consists of the amplitudes and the time constants related to the multi-exponential decomposition of the transverse relaxation decay. The first two features can be detected in a very fast way, only 12 ms are needed to obtain the results, because the information are contained in the first part of the acquired signal; for the third feature, the total acquisition time is needed, which is around 245 ms.

The classification algorithm showed a high sensitivity (97%) with a specificity around the 80%, so it is capable detecting the unhealthy hazelnuts but with a relatively high number of discarded healthy hazelnuts. A bootstrap method has been employed during performance evaluation in order to increase the sample population. The loss of specificity is mainly due to the third step of the algorithm, while in first two steps this value is similar to the sensitivity. This does not represent a limit because the defects related to the kernel development are well detected and they outnumber other kinds of defects. Moreover, the moisture content is also well detected by the classification algorithm, and this allows discarding the hazelnuts that contain a high level of moisture, that causes the mold development during the storage.

A more accurate unhealthy detection could be obtained with a strictly control of the influence parameters or using more complex techniques, like the Magnetic Resonance Imaging (MRI), but they are not suitable solutions for an industrial application, both for the cost and for the complexity of the systems.

Finally, the study has been carried out on the hazelnuts, but it can be easily extended to other kinds of fruit in shell, as well as the measurement algorithms can be applied in other kinds of NMR applications.

References

Aboonajmi, M. and Faridi, H., 2016. Nondestructive quality assessment of Agro-food products. In *Iranian International NDT Conference. Tehran* (pp. 1-9).

Amaral, J.S., Casal, S., Citová, I., Santos, A., Seabra, R.M. and Oliveira, B.P., 2006. Characterization of several hazelnut (Corylus avellana L.) cultivars based in chemical, fatty acid and sterol composition. *European Food Research and Technology*, *222*(3-4), pp.274-280.

Borgia, G.C., Brown, R.J.S. and Fantazzini, P., 1998. Uniform-penalty inversion of multiexponential decay data. *Journal of Magnetic Resonance*, *132*(1), pp.65-77.

Borgia, G.C., Brown, R.J.S. and Fantazzini, P., 2000. Uniform-penalty inversion of multiexponential decay data: II. Data spacing, T2 data, systematic data errors, and diagnostics. *Journal of Magnetic Resonance*, *147*(2), pp.273-285.

Carosio, M.G., Bernardes, D.F., Andrade, F.D., Moraes, T.B., Tosin, G. and Colnago, L.A., 2016. Measuring thermal properties of oilseeds using time domain nuclear magnetic resonance spectroscopy. *Journal of Food Engineering*, *173*, pp.143-149.

Chayaprasert, W. and Stroshine, R., 2005. Rapid sensing of internal browning in whole apples using a low-cost, low-field proton magnetic resonance sensor. *Postharvest biology and technology*, *36*(3), pp.291-301.

Chen, S.S., Li, R., Yu, J., Wang, H.Z. and Zhang, X.L., 2009, October. Program Algorithm Research of T2 Spectrum in NMR and MATLAB Realization. In *Image and Signal Processing, 2009. CISP'09. 2nd International Congress on* (pp. 1-4). IEEE.

Claridge, T.D., 2016. High-resolution NMR techniques in organic chemistry (Vol.27). Elsevier.

Colnago, L.A., Andrade, F.D., Souza, A.A., Azeredo, R.B., Lima, A.A., Cerioni, L.M., Osán, T.M. and Pusiol, D.J., 2014. Why is inline NMR rarely used as industrial sensor? Challenges and opportunities. *Chemical Engineering & Technology*, *37*(2), pp.191-203.

Commission Regulation (EC) No 1257/2006 of 21 August 2006 approving amendments to the specification for a protected geographical indication listed in the register of protected designation of origin and protected geographical indications (Nocciola di Giffoni) (PGI).

Dalitz, F., Cudaj, M., Maiwald, M. and Guthausen, G., 2012. Process and reaction monitoring by low-field NMR spectroscopy. *Progress in nuclear magnetic resonance spectroscopy*, *60*, pp.52-70.

Di Caro, D., Liguori, C., Pietrosanto, A. and Sommella, P., 2016, May. Investigation of the NMR techniques to detect hidden defects in hazelnuts. In *Instrumentation and Measurement Technology Conference Proceedings (I2MTC), 2016 IEEE International* (pp. 1-5). IEEE.

Di Caro, D., Liguori, C., Pietrosanto, A. and Sommella, P., 2017, May. Using a SVD-based algorithm for T 2 spectrum calculation in TD-NMR application to detect hidden defects in hazelnuts. In *Instrumentation and Measurement Technology Conference (I2MTC), 2017 IEEE International* (pp. 1-6). IEEE.

Di Caro, D., Liguori, C., Pietrosanto, A. and Sommella, P., 2017. Hazelnut Oil Classification by NMR Techniques. *IEEE Transactions on Instrumentation and Measurement*, *66*(5), pp.928-934.

Donis-Gonzalez, I.R., Guyer, D.E., Pease, A. and Barthel, F., 2014. Internal characterisation of fresh agricultural products using traditional and ultrafast electron beam X-ray computed tomography imaging. *Biosystems engineering*, *117*, pp.104-113.

FAOSTAT (2017). Food and Agriculture Organization of the United Nations. Available: http://faostat.fao.org/

Fawcett, T., 2006. An introduction to ROC analysis. *Pattern recognition letters*, *27*(8), pp.861-874.

FDA Technical Bulletin Number 5 Macroanalytical Procedures Manual 1984; Electronic Version 1998

Gianferri, R., D'Aiuto, V., Curini, R., Delfini, M. and Brosio, E., 2007. Proton NMR transverse relaxation measurements to study water dynamic states and age-related changes in Mozzarella di Bufala Campana cheese. *Food chemistry*, *105*(2), pp.720-726.

Horwitz, W. "Moisture in nuts and nut products (AOAC Official Method 925.40)." Officials methods of analysis of AOAC International, 17th edn. AOAC International, Gaithersburg, MD (2000)

Jibia, A.U. and Salami, M.J.E., 2012. Parameter Estimation of Transient Multiexponential Signals Using SVD-ARMA and Multiparameter Deconvolution Techniques. *International Journal of Computer Theory and Engineering*, *4*(5), p.751.

Jibia, A.U. and Salami, M.J.E., 2012. Analysis of transient multiexponential signals using exponential compensation deconvolution. *Measurement*, *45*(1), pp.19-29.

Jibia, A.U. and Salami, M.J.E., 2012. An appraisal of Gardner transform-based methods of transient multiexponential signal analysis. *International Journal of Computer theory and engineering*, *4*(1), p.16.

Kalkan, Habil, and Bayram Çetisli. "Online feature selection and classification." In Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, pp. 2124-2127. IEEE, 2011.

Kirtil, E., Cikrikci, S., McCarthy, M.J. and Oztop, M.H., 2017. Recent advances in time domain NMR & MRI sensors and their food applications. *Current Opinion in Food Science*, *17*, pp.9-15.

Liguori, C., Ruggiero, A., Sommella, P. and Russo, D., 2017. Choosing Bootstrap Method for the Estimation of the Uncertainty of Traffic Noise Measurements. *IEEE Transactions on Instrumentation and Measurement*, *66*(5), pp.869-878.

Mahdavi-Jafari, Somaieh, Hojjat Salehinejad, and Siamak Talebi. "A pistachio nuts classification technique: An ann based signal processing scheme." In Computational Intelligence for Modelling Control & Automation, 2008 International Conference on, pp. 447-451. IEEE, 2008.

Mancin, P. and Pilo, A., 1970. A computer program for multiexponential fitting by the peeling method. *Computers and Biomedical Research*, *3*(1), pp.1-14.

Mannina, L., Sobolev, A.P. and Viel, S., 2012. Liquid state 1 H high field NMR in food analysis. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *66*, pp.1-39.

Marcone, M.F., Wang, S., Albabish, W., Nie, S., Somnarain, D. and Hill, A., 2013. Diverse food-based applications of nuclear magnetic resonance (NMR) technology. *Food Research International*, *51*(2), pp.729-747.

Memoli, A., Albanese, D., Esti, M., Lombardelli, C., Crescitelli, A., Di Matteo, M. and Benucci, I., 2017. Effect of bug damage and mold contamination on fatty acids and sterols of hazelnut oil. European Food Research and Technology, 243(4), pp.651-658.

Meng, Z., Wu, Z. and Gray, J., 2017. Microwave sensor technologies for food evaluation and analysis: Methods, challenges and solutions. *Transactions of the Institute of Measurement and Control*, p.0142331217721968.

Mitchell, J., Gladden, L.F., Chandrasekera, T.C. and Fordham, E.J., 2014. Low-field permanent magnets for industrial process and quality control. *Progress in nuclear magnetic resonance spectroscopy*, *76*, pp.1-60.

Moody, J.B. and Xia, Y., 2004. Analysis of multi-exponential relaxation data with very short components using linear regularization. *Journal of Magnetic Resonance*, *167*(1), pp.36-41.

Nelson, S. O., C. V. K. Kandala, and K. C. Lawrence. "Single-kernel moisture determination in grain and nuts by RF impedance measurements." In Instrumentation and Measurement Technology Conference, 1992. IMTC'92., 9th IEEE, pp. 422-424. IEEE, 1992

Nicolotti, L., Cordero, C., Bicchi, C., Rubiolo, P., Sgorbini, B. and Liberto, E., 2013. Volatile profiling of high quality hazelnuts (Corylus avellana L.): chemical indices of roasting. *Food chemistry*, *138*(2), pp.1723-1733.

OECD (2011), International Standards for Fruit and Vegetables: Inshell hazelnuts and hazelnut kernels, International Standards for Fruit and Vegetables, OECD Publishing. http://dx.doi.org/10.1787/9789264166721-enfr

Onaran, Ibrahim, Berkan Dulek, Tom C. Pearson, Yasemin Yardimci, and A. Enis Cetin. "Detection of empty hazelnuts from fully developed nuts by impact acoustics." In Signal Processing Conference, 2005 13th European, pp. 1-4. IEEE, 2005

Pannico, A., Schouten, R.E., Basile, B., Romano, R., Woltering, E.J. and Cirillo, C., 2015. Non-destructive detection of flawed hazelnut kernels and lipid oxidation assessment using NIR spectroscopy. *Journal of Food Engineering*, *160*, pp.42-48.

Parker, T., Limer, E., Watson, A.D., Defernez, M., Williamson, D. and Kemsley, E.K., 2014. 60MHz 1 H NMR spectroscopy for the analysis of edible oils. *TrAC Trends in Analytical Chemistry*, *57*, pp.147-158.

Provencher, S.W., 1982. CONTIN: a general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. *Computer Physics Communications*, *27*(3), pp.229-242.

Salami, M.J.E., Tijani, I.B. and Jibia, A.U., 2012. Matlab-Based Algorithm for Real Time Analysis of Multiexponential Transient Signals. In *MATLAB-A Fundamental Tool for Scientific Computing and Engineering Applications-Volume 1*. InTech.

Solar, Mitja, and Anita Solar. "Non-destructive determination of moisture content in hazelnut (Corylus avellana L.)." Computers and Electronics in Agriculture 121 (2016): 320-330

Speir, R.A. and Haidekker, M.A., 2017. Onion postharvest quality assessment with x-ray computed tomography–A pilot study. *IEEE Instrumentation & Measurement Magazine*, *20*(3), pp.15-19.

Spincore Technologies Inc., 2014. iSpinNMR™ Owner's Manual.

Teng, Q., 2012. *Structural biology: practical NMR applications*. Springer Science & Business Media.

Todt, H., Guthausen, G., Burk, W., Schmalbein, D. and Kamlowski, A., 2006. Water/moisture and fat analysis by time-domain NMR. *Food chemistry*, *96*(3), pp.436-440.

Trabelsi, S. and Nelson, S.O., 2016. Microwave sensing of quality attributes of agricultural and food products. *IEEE Instrumentation & Measurement Magazine*, *19*(1), pp.36-41.

Wang, H., Liu, J., Xu, X., Huang, Q., Chen, S., Yang, P., Chen, S. and Song, Y., 2016. Fully-Automated High-Throughput NMR System for Screening of Haploid Kernels of Maize (Corn) by Measurement of Oil Content. *PloS one*, *11*(7), p.e0159444.

Weekley, A.J., Bruins, P., Sisto, M. and Augustine, M.P., 2003. Using NMR to study full intact wine bottles. *Journal of Magnetic Resonance*, *161*(1), pp.91-98.

Zhang, J., Liu, Y., Wang, N. and Ruan, R., 2012. NMR Technique Application in Evaluating the Quality of Navel Orange During Storage. *Procedia Engineering*, *37*, pp.234-239.

Appendix
Multi-exponential inversion algorithms

The analysis of the multi-exponential transient signals is particularly important for their occurrence in several fields, mainly applied physics and chemistry. Examples of problems involving this kind of signals are the fluorescence decay analysis in biophysics, the radioactive decay in nuclear physics, the reaction kinetics in chemistry and the study of the relaxation decay in the Nuclear Magnetic Resonance (Jibia and Salami, 2012).
A multi-exponential transient signal can be expressed as a linear combination of exponentials:

$$S(\tau) = \sum_{i=1}^{M} A_i e^{-\lambda_i \tau} + n(\tau) \qquad (48)$$

Where M is the number of components, A_i is the amplitude of the i-th component, λ_i is the decay rate constant of the i-th component and $n(\tau)$ is the additive white Gaussian noise. The solution of eq.(48) consist of the determination of the parameters M, A_i and λ_i of the measured signal $S(\tau)$. The main problem in the analysis of this kind of signals is the non-orthogonality of the exponential decay functions. This means that if an attempt is made to determine the unknown parameters A_i, λ_i and M from finite time samples of the signal in eq.(48), the distribution function of the decay rates will not be unique (Jibia and Salami, 2012b).

Several techniques have been proposed for the analysis of the multi-exponential transient signals. They can be classified as time-domain or frequency-domain, and parametric or non-parametric (Salami *et al.*, 2012). Time-domain techniques are the oldest methods of multi-exponential signal analysis. One of the first example of time-domain technique is the peeling method, which attempts to separate the various components based on the different slopes of the components at infinity (Mancin and Pilo, 1970). This method produces poor results when the signal contains more than two components. The main application of a frequency-domain technique has been the Gardner transformation. Several algorithms have been proposed, based on

this approach. It is based on the inversion of an integral form of eq.(48) into a convolutional integral and then the deconvolution is obtained applying the Fourier transform (Jibia and Salami, 2012c). Another category of techniques, exploited in different algorithms, is the linear regularization. An example of algorithm based on this technique is the uniform penalty inversion of multi-exponential decay data (UPEN) that is used for the inversion of the relaxation decay of NMR signals in order to give a quasi-continuous distribution of the relaxation times. Usually, some smoothing of the distributions is implemented to avoid excess variation. When the same distribution has a sharp peak and a much broader peak or a "tail," as for many porous media saturated with liquids, an inversion program using a fixed smoothing coefficient may broaden the sharp peak and/or break the wide peak or tail into several separate peaks, even if the coefficient is adaptively chosen in accord with the noise level of the data. The UPEN algorithm deals with this problem by using variable smoothing, determined by iterative feedback in such a way that the smoothing penalty is roughly constant (Borgia *et al.*, 1998, Borgia *et al.*, 2000). Another algorithm based on linear regularization is CONTIN (Provencher, 1982). It exploits several constraints into the regularization in order to obtain a stable solution. In particular, it makes use of a combination of three strategies to determine the constraints: (1) Absolute prior knowledge, (2) statistical prior knowledge, and (3) the principle of parsimony. A comparison between the UPEN and CONTIN algorithm can be found in (Moody and Xia, 2004), in which the ability of detecting a decay component lower than the time of the first acquired echo in the CPMG sequence, is analyzed. In this work, as described in chapter 3 and 4, an algorithm based on singular value decomposition (SVD) has been used to determine the main components of the T_2 distribution in the oil samples, and then a non-linear least square fitting considering a three component model of the signal.

All the multi-exponential techniques have some advantages and disadvantages. The main parameters to take in account are:

- Noise level of the acquired signal;
- Number of components of the multi-exponential decay;
- Time distribution of the components;
- Time of execution of the algorithm.

All the mentioned techniques suffer a low S/N ratio of the acquired signal. In particular, when the S/N ratio decreases, the solution of eq. (48) tends to be instable. Some techniques work better than other with signal that exhibit a few number of components, or when it is not required a small time resolution. Moreover, most of the algorithms are iterative procedures, and they are often computational inefficient.

In the present work, the SVD algorithm has been adopted because it exhibited a stable solution in few iterations, so it proved to be suitable for the specific application. Moreover, the non-linear least square fitting, although is considered a computational inefficient technique, allowed obtaining the small

resolution required to discriminate between the similar components related to healthy and unhealthy hazelnuts, limiting the execution by means of the time constraints in the model of the signal, composed by three components with the time constants in predefined ranges.