UNIVERSITÀ DEGLI STUDI DI PADOVA

SCUOLA DI DOTTORATO IN INGEGNERIA DELL'INFORMAZIONE

University of Padova
Department of Information Engineering

Ph.D. School in Information Engineering
Section: Bioengineering
Series: XXVII

GLUCOSE VARIABILITY ASSESSMENT IN DIABETES MELLITUS MONITORING AND CONTROL

School Director: Ch.mo Prof. Matteo Bertocco

Coordinator: Ch.mo Prof. Giovanni Sparacino

Advisor: Ch.mo Prof. Giovanni Sparacino

Ph.D. Candidate: Ing. Chiara Fabris

A thesis submitted for the degree of
Philosophiæ Doctor (Ph.D.)
2015

To my parents

Acknowledgements

My gratitude goes to all the people who have supported me during this experience.
Thanks to my advisor, Professor Giovanni Sparacino, for his guidance and assistance, and to Professor Andrea Facchinetti for the supervision and collaboration.
Thanks to Professor Marc Breton for the opportunity of working with his group and for his teachings, and to Professor Stephen Patek for his aid and suggestions.
Above all, thanks to my parents for their never-ending unconditional support and help.

C. F.

Abstract

This dissertation is focused on the assessment of glucose variability (GV) in the treatment of the pathology of diabetes mellitus. GV is a risk factor for the development of diabetes complications, and its assessment combined with the evaluation of glycated hemoglobin levels is believed to be useful to characterize the functioning of glucose metabolism. Given the importance of GV in diabetes, a number of indicators to measure it from the retrospective analysis of sparse self-monitoring of blood glucose (SMBG) or continuous glucose monitoring (CGM) recordings have been proposed in the literature, but several issues are still open. For instance, some GV indicators have been developed specifically from SMBG data, and their use on CGM time-series has not been validated yet. Moreover, the availability of a large number of metrics to quantify GV gives rise to problems in terms of redundant conveyed information, and a compact way to extensively characterize GV would be desirable. Finally, the exploitation of CGM signals and GV to classify the metabolic condition of normal and diabetic subjects is a relatively unexplored problem that could deserve an investigation. These three topics are the object of this dissertation, which is specifically made up of six chapters whose content is briefly outlined below.

Chapter 1 will describe the etiology of the different types of diabetes, discuss the development of diabetes complications, and introduce the technologies used to monitor blood glucose levels and the strategies exploited to manage the treatment of type 1 (T1DM) and type 2 (T2DM) diabetes mellitus.

Chapter 2 will focus specifically on GV and its quantification, and, after highlighting the existing open issues, will precisely state the aims of the thesis.

Chapter 3 will consider the problem of adapting some GV indicators originally developed and validated from SMBG, to the use with CGM signals. In particular, we will specifically look at low blood glucose index (LBGI) and high blood glucose index (HBGI), popular metrics that allow to provide a rapid classification of the quality of glucose control in diabetic subjects, and will provide alternate versions of these indicators adapted to the characteristics of CGMs by modeling the relationship between LBGI/HBGI values obtained from SMBG and CGM recordings. A dataset of

28 T1DM subjects monitored with both SMBG and CGM devices will be used to tune and assess the proposed methodology.

Chapter 4 will address the issue of redundant information conveyed by the available GV indices by using the sparse principal component analysis (SPCA) technique as a tool to provide a parsimonious but still comprehensive characterization of GV in both T1DM and T2DM. Specifically, we will consider 25 GV indicators evaluated on CGM profiles acquired from 33 T1DM and 13 T2DM subjects as initial pool of variables. SPCA will be applied to this set of metrics and will be shown to be able to select a small subset of up to 10 indices that can save more than 60% of the original variance in both applications. The subset of metrics provided by SPCA can be used to parsimoniously describe GV in diabetes.

Chapter 5 will be devoted to the assessment of the possibility of using the outputs from SPCA to build GV-based classifiers of the metabolic condition of normal and diabetic subjects. In particular, by resorting to a dataset of 55 T1DM subjects, 34 normal subjects at high risk of developing T2DM, 39 impaired glucose tolerance subjects, and 29 subjects with T2DM diagnosed, we will show that support vector machines are able to successfully classify the quality of glycemic control and the metabolic condition of disordered subjects, allowing to achieve an accuracy of classification always greater than 70%. The investigation will be performed using both the whole initial pool of 25 indicators and the parsimonious set selected by SPCA as features to design the classifiers; the fact that similar results were obtained in the two scenarios strengthens the speculation that the compact description of GV provided by SPCA is effectively comprehensive for characterizing the subjects' metabolic condition.

Chapter 6 will close this dissertation, with a discussion on possible future developments of the presented investigations.

Sommario

L'obiettivo di questa tesi è l'indagine del ruolo della variabilità glicemica (GV) nella patologia del diabete mellito. La GV è un fattore di rischio per lo sviluppo di complicazioni dal diabete, e la sua valutazione combinata con quella dei livelli di emoglobina glicata è ritenuta essere un elemento utile nel caratterizzare il funzionamento del metabolismo del glucosio. Data l'importanza della GV nel diabete, molteplici indicatori che permettono di ottenerne una quantificazione dall'analisi retrospettiva di segnali di self-monitoring of blood glucose (SMBG) o continuous glucose monitoring (CGM) sono stati proposti in letteratura, ma in merito esistono alcune problematiche ancora aperte. Per esempio, alcuni indici sono stati sviluppati specificamente per essere applicati su serie SMBG, ed il loro utilizzo su segnali CGM non è ancora stato validato. Inoltre, il fatto che esistano numerosi indicatori per quantificare la GV dà origine a problemi di ridondanza nell'informazione trasmessa, ed un approccio che permetta di ottenere una descrizione compatta ma esaustiva della GV sarebbe desiderabile. Infine, l'uso di segnali CGM e dell'informazione sulla GV per classificare lo stato metabolico di soggetti normali e diabetici è un problema relativamente inesplorato che potrebbe meritare di essere trattato. Questi tre argomenti sono l'oggetto di questa tesi, che risulta articolata in sei capitoli il cui contenuto è brevemente delineato di seguito.

Il Capitolo 1 descriverà l'eziologia dei differenti tipi di diabete, discuterà lo sviluppo delle complicazioni da diabete, ed introdurrà le tecnologie utilizzate per monitorare la glicemia ed alcune strategie che si possono seguire per trattare il diabete mellito di tipo 1 (T1DM) e 2 (T2DM).

Il Capitolo 2 verterà sulla GV e la sua quantificazione, e, dopo aver evidenziato i problemi aperti esistenti, dichiarerà precisamente gli scopi della tesi.

Il Capitolo 3 considererà il problema di adattare alcuni indicatori di GV originariamente sviluppati e validati su profili SMBG, all'utilizzo su segnali CGM. In particolare, ci concentreremo su low blood glucose index (LBGI) e high blood glucose index (HBGI), indici popolari che permettono di ottenere una rapida classificazione della qualità del controllo glicemico in soggetti diabetici, e forniremo versioni alternative

di questi indicatori adattate alle caratteristiche dei segnali CGM, modellando la relazione tra i valori che LBGI e HBGI assumono quando calcolati da SMBG e CGM. Un dataset di 28 soggetti T1DM monitorati con dispositivi SMBG e CGM sarà utilizzato per mettere a punto la metodologia.
Il Capitolo 4 affronterà il problema della ridondanza nell'informazione fornita dagli indicatori di GV esistenti, utilizzando la sparse principal component analysis (SPCA) come approccio per fornire una descrizione parsimoniosa ma allo stesso tempo esaustiva della GV in popolazioni di soggetti con T1DM e T2DM. In particolare, considereremo 25 indicatori di GV valutati su profili CGM acquisiti da 33 soggetti con T1DM e 13 con T2DM come insieme iniziale di variabili. La SPCA sarà applicata a questo pool di indici e permetterà di selezionare un piccolo sottoinsieme di 10 indicatori che consente di preservare più del 60% della varianza originariamente spiegata dall'insieme di partenza in entrambe le applicazioni. Il sottoinsieme di indicatori fornito dalla SPCA può essere utilizzato per descrivere parsimoniosamente la GV nel diabete.
Il Capitolo 5 sarà dedicato alla valutazione della possibilità di utilizzare gli output della SPCA per costruire classificatori dello stato metabolico di soggetti normali e diabetici basati sulla GV. In particolare, facendo ricorso ad un dataset di 55 soggetti con T1DM, 34 normali a rischio T2DM, 39 con impaired glucose tolerance, e 29 con T2DM diagnosticato, mostreremo che classificatori progettati su support vector machine sono capaci di discriminare con successo la qualità del controllo glicemico e la condizione metabolica di soggetti con disordini, permettendo di raggiungere un'accuratezza di classificazione sempre maggiore del 70%. Lo studio sarà condotto utilizzando sia il pool iniziale di 25 indicatori che il sottoinsieme parsimonioso fornito dalla SPCA come features per costruire i classificatori; il fatto che risultati simili siano ottenuti nei due casi rafforza la speculazione che la descrizione compatta della GV fornita dalla SPCA sia effettivamente esaustiva nel caratterizzare la condizione metabolica dei soggetti.
Il Capitolo 6 chiuderà la tesi, con una discussione su possibili sviluppi futuri degli studi qui presentati.

Contents

Abbreviations

ADRR	Average daily risk range
AP	Artificial pancreas
BG	Blood glucose
BGRI	Blood glucose risk index
CGM	Continuous glucose monitoring
DCCT	Diabetes Control and Complications Trial
GDM	Gestational diabetes mellitus
GRADE	Glycemic risk assessement diabetes equation
GV	Glucose variability
HbA1c	Glycated hemoglobin
HBGI	High blood glucose index
IFG	Impaired fasting glucose
IGC	Index of glycemic control
IGT	Impaired glucose tolerance
IQR	Interquartile range
LASSO	Least amplitude shrinkage and selection operator
LBGI	Low blood glucose index
LLTR	Lower limit of target range
MAGE	Mean amplitude of glycemic excursions
MSE	Mean squared error
PC	Principal component
PCA	Principal component analysis
R^2	Coefficient of determination
SD	Standard deviation
SD_w	Within days standard deviation
SD_{dm}	Standard deviations of daily means
SE	Standard error
SH	Severe hypoglycemia
SMBG	Self-monitoring of blood glucose
SPCA	Sparse principal component analysis
SVM	Support vector machine
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus
ULTR	Upper limit of target range
%CV	Percentage coefficient of variation
$\%GRADE_{eu,hyper,hypo}$	Percentages of GRADE due to eu, hyper, and hypoglycemia

List of figures

List of tables

Chapter 1

The pathology of diabetes and its monitoring

1.1 The pathology of diabetes

Diabetes is a metabolic disease resulting from defects in insulin secretion, insulin action, or both [1]. Insulin is a hormone automatically released by the islets of Langerhans of the pancreatic beta-cells, that regulates the level of glucose concentration in the blood and keeps it within the so-called euglycemic range, a healthy condition where glycemia is between 70 and 180 mg/dL [2]. After a meal, glucose molecules derived from the digestion of carbohydrates enter the bloodstream and rapidly increase the subject's glycemia. If properly secreted by the pancreas, insulin available in the blood stimulates glucose uptake by various body tissues (e.g., muscles and adipose tissue) and inhibits its production by the liver, reducing the level of glucose in the blood and restoring a safe condition of euglycemia. Being the glucose uptake by certain tissues an insulin-dependent process, the appearance of insulin in the bloodstream after a meal is necessary to allow glucose molecules entering the cells, producing energy, and stimulating growth. If no insulin is available in the bloodstream, or tissues are no more sensitive to its action, only insulin-independent glucose uptake (e.g., by the brain) can occur. This entails metabolic damages related to the lack of glucose provided for the feeding of insulin-dependent cells and leads to chronic abnormally high levels of glucose concentration in the blood that end up in a long-term damaging condition known as hyperglycemia [1]. The scenario just described is what characterizes the pathology of diabetes, where, depending on the specific type of disease, the system producing insulin is progressively destroyed or the effectiveness of insulin action on

glucose transportation through the cell membrane is impaired. To compensate the defects in their glucose-insulin regulation system, beyond healthy diet and regular physical activity, diabetic people may have to inject exogenous insulin multiple times a day, especially when they eat, with the risk of overestimating the amount of insulin needed to restore a euglycemic state, thus inducing an unsafe decrease of glucose levels in the bloodstream. This condition, that is the opposite of the previously introduced hyperglycemia, is known as hypoglycemia, and its occurrence represents an extremely serious problem in the pathology of diabetes. Hypoglycemia is a medical emergency where the abnormal reduction of glycemic levels degenerates into an inadequate supply of glucose to the brain (neuroglycopenia) that, if severe and not properly treated, may rapidly progress into coma and eventually into death, also without subject awareness especially at night. Because of the potentially life-threatening damages associated to hypoglycemia, the chance of its occurrence represents one of the most limiting factors in following an intensive therapy to strongly reduce hyperglycemia and achieve a tight glycemic control in the treatment of diabetes [3].

Risk associated with the extreme glycemic conditions experienced by diabetic people and the massive diffusion of this pathology render diabetes an actual emergency of the 2000s, and numbers related to it are effectively alarming. As updated to November 2014, key facts from the World Health Organization [4] report that 347 million people throughout the world suffer from diabetes, with an estimation of 1.5 million deaths directly caused by diabetes in 2012, more than 80% of deaths occurring in low- and middle-income countries, and projection of diabetes being the 7th leading cause of death in 2030. As far as the American community is concerned [5], the prevalence of diabetes in 2012 was 29.1 million people, i.e., 9.3% of the entire population, of whom 8.1 million were undiagnosed. In the same year, the incidence of diabetes was 1.7 million new diagnoses/year, while in 2010 it was 1.9 million. These numbers made diabetes the 7th leading cause of death in the United States in 2010, with 69,071 death certificates listing it as the underlying cause of death, and a total of 234,051 death certificates listing diabetes as an underlying or contributing cause of death.

1.2 Types of diabetes

Three different types of diabetes can be identified [1]. The two major ones are known as type 1 and type 2 diabetes, the third one, less common, is known as gestational diabetes. Moreover, intermediate conditions between normality and diabetes have

been also classified, and are known as impaired glucose tolerance and impaired fasting glucose. They are all detailed below.

1.2.1 Type 1 diabetes mellitus (T1DM)

Type 1 diabetes mellitus (T1DM), also known as insulin-dependent diabetes, is an autoimmune disorder that leads to the complete destruction of the insulin-producing pancreatic beta cells [1]. People with T1DM become totally unable to produce insulin and need to regularly take exogenous insulin to keep their glycemic levels controlled. Approximately 5-10% of diabetic people are affected by T1DM [5]. Although a large number of T1DM patients developed the condition during childhood (reason why T1DM is also called juvenile or childhood diabetes), the pathology can appear in people of any age [6, 7]. The cause of T1DM is not known, the pathology is not preventable with current knowledge, the majority of people developing it are of normal weight, and exercise and diet, though invaluable in the treatment of T1DM, cannot reverse this disorder [4, 7]. Several clinical trials have attempted to find ways to prevent or slow down the progress of T1DM, but with no proven success so far [7].

1.2.2 Type 2 diabetes mellitus (T2DM)

People with type 2 diabetes mellitus (T2DM) are not able to produce enough insulin or suffer from the so-called insulin resistance, condition where cells are not sensitive to insulin action anymore and glucose cannot enter them as much as it would do in normal subjects [1]. T2DM is the most common form of diabetes and about 90% of diabetic people suffer from this type of disorder [4, 6, 7]. The majority of people with T2DM have developed the condition because they are overweight and physically inactive, reason why increased exercise and dietary changes can help in initially managing this disease [7]. Since T2DM can be treated without the injection of exogenous insulin, it is also called insulin-independent diabetes. Nevertheless, if the optimization of lifestyle is not enough to control glycemic levels of T2DM subjects, administering medications or insulin during the day could be necessary to treat T2DM as well [6, 7].

1.2.3 Gestational diabetes mellitus

Gestational diabetes mellitus (GDM) is a condition that occurs when a woman without previously diagnosed diabetes exhibits high levels of glucose concentration in the blood during pregnancy, usually around the 24th week [1, 7]. GDM complicates nearly 4%

of all pregnancies in the United States, resulting in 135,000 cases annually [1]. The prevalence may range from 1 to 14% of pregnancies, depending on the population studied, and nearly 90% of all pregnancies complicated by diabetes are related to the occurrence of GDM [1]. Specialists are not sure about the causes of GDM, but it seems that hormones from the placenta that help the baby develop, also block the action of the mother's insulin in her body [7]. Thus, as for T2DM, women with GDM suffer from insulin resistance, and may need up to three times as much insulin. Mothers with untreated GDM are likely to develop T2DM after pregnancy and can have babies abnormally large for gestational age (condition that may lead to delivery complications) [7].

1.2.4 Pre-diabetes conditions

Impaired glucose tolerance (IGT) and impaired fasting glucose (IFG) are conditions preceding the development of diabetes respectively associated with a state of hyperglycemia due to insulin resistance and a consistently elevated fasting glycemic level, however not enough to prompt a diagnosis of diabetes [1]. Both these conditions increase the risk of developing cardiovascular pathology and T2DM, with a greater risk for IGT than for IFG [4]. Some patients with IFG can also be diagnosed with IGT, but many have normal responses to a glucose tolerance test (medical procedure where glucose is given to the patients and blood samples are taken afterwards to assess how quickly glycemia goes back to normal values). Lifestyle modifications, regular physical exercise, limited intake of sugar and highly processed carbohydrates, and loss of body weight can delay the onset of diabetes [7].

1.3 Development of diabetes complications

The attempt of avoiding hyperglycemia and achieving a tight glycemic control is a major goal in the treatment of diabetes, not only to ensure an adequate delivery of glucose to insulin-dependent cells, but also to decrease the likelihood that body tissues are harmed by prolonged high blood glucose (BG) levels. As detailed in the literature [8–10], hyperglycemia represents a risk factor for the development of serious long-term complications from diabetes, and its direct and indirect effects on the vascular tree have been reported to be the major source of morbidity and mortality in both T1DM and T2DM. The etiologic characteristics of diabetes complications are mostly related to the role played by chronic hyperglycemia in initiating many metabolic and structural

derangements, that include production of advanced glycation end products, abnormal activation of signaling cascades, elevated production of reactive oxygen species (oxygen-containing molecules that can interact with other biomolecules and result in damaging oxidative stress), and abnormal stimulation of hemodynamic regulation systems [11]. These processes, combined with other factors, progress into specific diseases that, depending on their nature, are usually classified as macrovascular or microvascular diabetes complications.

The central pathological mechanism that entails macrovascular diseases is the process of atherosclerosis, where plaques building up along the inner walls of the arteries narrow the lumen of vessels and reduce blood flow to target sites [8]. Atherosclerosis is thought to result from chronic inflammation and injury of the arterial walls in the cardiac, cerebral, and peripheral vascular system, and represents a potential cause of deleterious conditions like coronary artery disease, myocardial infarction, cerebrovascular disease, stroke, and peripheral arterial disease.

Beside damages to large vessels, prolonged hyperglycemia can cause a number of microvascular complications specific to diabetes that include retinopathy, nephropathy, and neuropathy [9]. Among these harmful conditions, retinopathy, a potential cause of visual impairment and blindness, is the most common injury. The majority of diabetic patients are affected by this disease, with approximately 50% of T1DM subjects suffering from some degree of retinopathy after seven years, and 90% after 20 years. Unlike retinopathy, nephropathy develops in only 35 to 45% of subjects with T1DM and less than 20% of subjects with T2DM. This is the diabetes-specific complication with the greatest mortality, and can lead to renal failure with potential progression into end-stage renal disease. The third most common microvascular complication from diabetes is neuropathy, that is associated to a number of clinical manifestations including impairment of gastric or intestinal motility, impotence, and diabetic foot disorders. The prevalence of neuropathy as defined by loss of ankle jerk reflexes is 7% at one year of diabetes, increasing to 50% at 25 years for both T1DM and T2DM.

To treat diabetes and manage the occurrence of sustained hyperglycemia causing the progression of complications, diabetic people have to frequently monitor their BG levels. To this aim, two main strategies exist, which are discussed in the next section.

1.4 Strategies for glucose monitoring

In 1978, a study involving 64 diabetic patients demonstrated that the possibility of measuring BG concentration at home using commercially available glucose meters

could lead to a significant improvement of glycemic control [12]. Since that time, the technology of glucose meters (Section 1.4.1) has evolved towards the nowadays increasingly used glucose sensors (Section 1.4.2).

1.4.1 Self-monitoring of blood glucose (SMBG)

Self-monitoring of blood glucose (SMBG) devices are small, battery-operated glucose meters that allow to measure the concentration of glucose from a drop of capillary blood usually taken from a finger. When the subject wants to test for glucose, he needs to prick his skin with a lancet, place a small sample of his blood on a reagent strip, and determine the glucose concentration by inserting the strip into the meter (usually a reflectance photometer) for an automatic reading [13]. After each measurement, results are stored in the meter's electronic memory. Examples of commercially available SMBG devices are shown in the left panel of Figure 1.1; the right panel of the same figure reports a representative SMBG profile collected from a T1DM subject for three consecutive days together with hypo and hyperglycemic thresholds.

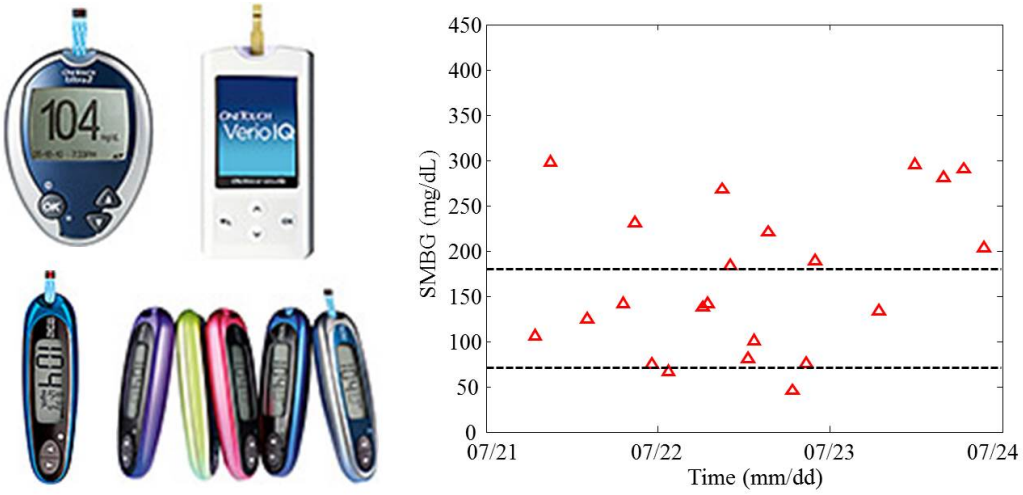

Figure 1.1 Examples of commercially available SMBG devices (left panel - taken from [14]) and representative SMBG profile shown with hypo and hyperglycemic thresholds at 70 and 180 mg/dL, respectively (right panel).

Through the use of SMBG devices, diabetic people have an immediate feedback about their current glycemic level and can correct any deviation out of a desired target range by changing the carbohydrate intake, exercising, or using more or less insulin. In this context, it is clear that the sampling design (when and how frequently subjects measure their glycemia) plays a major role in achieving a good glycemic control, avoiding extreme hyperglycemia, and promptly treating hypoglycemic episodes. The frequency

with which patients sample varies from person to person, but most experts agree that insulin-treated diabetics should monitor their glycemic level at least four times a day, most commonly fasting, before meals, and before bed, with a significant benefit derived from collecting post-prandial BG readings to adjust insulin regimen [13]. In general, a positive correlation between frequency of SMBG and glycemic control among patients with insulin-treated T1DM or T2DM seems to hold [15, 16]. However, even if BG levels are monitored through a full seven-point SMBG profile, the subject is still unable to carefully track the evolution of his glycemic condition across the whole day (see the right panel of Figure 1.1). This can lead to prolonged hyperglycemic states, hypoglycemic attacks not promptly recognized and treated, especially in subjects with hypoglycemia-unawareness, and information about glucose fluctuations totally missed and not considered by the doctor in optimizing the ongoing therapy. To overcome these limitations, a system to allow more-frequent (ideally continuous), less-invasive BG measurements was necessary, and to satisfy these requirements the technology of implantable glucose sensors has been developed since the early 2000s.

1.4.2 Continuous glucose monitoring (CGM)

Continuous glucose monitoring (CGM) systems are portable devices that allow an almost continuous monitoring of interstitial glycemia (sampling time of 1-5 minutes) for up to seven consecutive days [17–21]. The substantially high sampling frequency of these devices is possible thanks to the technology of implantable glucose sensors that, together with a wireless transmitter and a receiver, are the core of CGM systems. The left panel of Figure 1.2 shows an example CGM device; sensor and transmitter can be seen in the upper part of the panel, from left to right, respectively; the receiver is represented in the lower part. In practice, the sensor is inserted into the subcutaneous tissue (most commonly at the abdomen) and a reusable transmitter attached to it relays the glucose data wirelessly to the receiver, where they are displayed. The receiver, that allows to store several days of monitoring in its memory, may be a dedicated handheld device or an insulin pump designed to integrate with CGM (the so-called sensor-augmented insulin pump described in the next section). Either way, looking at the monitor of the receiver, the user can track his glycemic levels in real time and benefits from symbols indicating trends of stable, increasing, or decreasing glucose status, together with a graphical representations of the evolution of his glycemia (see the left panel of Figure 1.2). Beside the visual information, the devices also offer audible alarms to alert the patient when BG levels are below or above a preset safe range,

to aid in identifying impending hypo or hyperglycemia. Together with an example of CGM device, Figure 1.2 shows in its right panel a representative CGM profile collected from the same subject as the SMBG samples of Figure 1.1. It is clear by comparing the time-series that CGM sensors allow to capture glycemic excursions and hypo/hyperglycemic episodes that cannot be seen from sparse SMBG recordings.

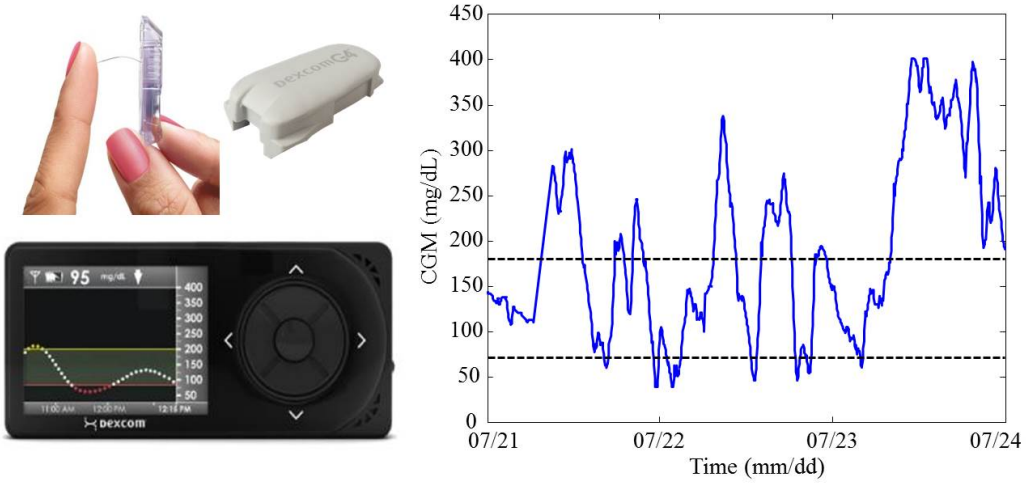

Figure 1.2 Example of commercially available CGM device (left panel - taken from [22]) and representative CGM profile shown with hypo and hyperglycemic thresholds at 70 and 180 mg/dL, respectively (right panel).

The key element within CGM systems is of course represented by the glucose sensor. In most of the currently available minimally-invasive CGM devices, the sensor exploited is electrochemical and consists of a wire inserted subcutaneously into the back, buttocks, or most commonly abdomen, that produces an electrical current correlating with the interstitial glucose level. These sensors are amperometric systems that measure the current flowing from an oxidation (electron producing) reaction at a working electrode to a reduction (electron consuming) reaction at a counter electrode. The working electrode is typically coated with the enzyme glucose oxidase that is necessary to oxidate glucose and start the reaction sequence. The electrical current finally measured by the sensor is proportional to the amount of glucose in the subcutis. Sensors are inserted by the patient and can be immediately replaced after their useful lifetime. To be sure that the sensor maintains its accuracy over time, the device needs to be often calibrated. Sensor calibration is the transformation in real time of the electrical current signal generated by the sensor at a certain time into an estimation of glucose concentration [23]. The reference BG levels used to optimize the transformation are SMBG readings conveniently collected and entered into the system's monitor. All commercial CGM sensors need to be calibrated through two SMBG readings when they are inserted at

the beginning of their lifetime. Then, the minimum frequency of calibration required is once every 12 hours, but calibrating three to four times a day can optimize sensor accuracy. Having accurate glucose sensors within CGM systems is necessary also for the role played by these devices in the development of the so-called artificial pancreas (AP) [24]. The AP is a closed-loop control algorithm that automatically releases insulin based on BG levels and other inputs. BG readings are provided by CGM devices, and inaccurate measurements can lead to wrong insulin administration and occurrence of unsafe conditions of hypoglycemia or hyperglycemia.

Beside their use within the AP, the possibility given by CGM devices to track BG levels in real time has made an important difference in the treatment of diabetes. When compared to traditional SMBG alone, the use of CGM systems has been shown to significantly improve the quality of glycemic control, as quantified by the levels of glycated hemoglobin (HbA1c), a measure related to the average glycemia over the 2-3 months preceding the test, and by the time spent in hypoglycemia [25, 26]. The use of CGMs has been in particular recommended for T1DM adults seeking to improve their HbA1c, control hypoglycemic episodes, and detect nocturnal hypoglycemia, early morning hyperglycemia due to insulin resistance (the dawn phenomenon), and postprandial hyperglycemia [20, 27, 28]. Moreover, because of the high sampling frequency and long-lasting duration of CGM recordings, glucose sensors represent a powerful tool also for the retrospective investigation of the nowadays extremely debated glucose variability, a risk factor for the development of long-term complications from diabetes that will be carefully described in the next chapter and is the focus of this dissertation. Components of glucose dynamics invisible in the SMBG profiles are, in fact, captured in detail by CGM systems (compare the right panel of Figure 1.1 with the right panel of 1.2), and the assessment of the time spent in hypo/hyperglycemia or of the amplitude of glucose fluctuations can be made on more reliable bases than with SMBG, also avoiding the dependence on the specific sampling design followed by the subject in collecting his BG measurements.

As said, glucose variability is matter of Chapter 2. The following of this chapter will provide an overview on the treatment of diabetes, presenting the two major ways exploited by insulin-treated diabetic subjects to inject exogenous insulin.

1.5 Treatment of diabetes

People with diabetes have to follow specific rules in the management of their lifestyle, in order to reduce the risk of developing long-term vascular dysfunctions and be able

to live well with either T1DM or T2DM. The strategy used to treat diabetes has to be designed by the health care team on an individual basis, and needs to address medical, physiological, and lifestyle patient-specific issues. Nonetheless, some general basics for the treatment of the pathology can be outlined, and are defined based on the specific type of diabetes [7].

In the case of T1DM, the complete lack of insulin production requires the patient to undergo a strict regimen that typically includes a carefully calculated diet, planned physical activity, home BG testing a number of times per day, and exogenous insulin administration through multiple daily insulin injections (Section 1.5.1) or continuous subcutaneous insulin infusion (Section 1.5.2). In all cases, T1DM patients have to compensate for the lack of a basal insulin concentration that keeps BG stable through periods of fasting, and for the missed increase in insulin release after meals. This requires them to measure their BG before eating, inject a certain amount of insulin based on the ingested carbohydrates, proteins, and fats, and sample the BG level again after meal to administer further insulin if needed.

Unlike T1DM, a balanced treatment of T2DM is usually achieved through controlled diet, exercise, home BG testing, and, in some cases, oral medication and/or insulin. The prevalence of T2DM people exploiting therapies based on insulin injection is approximately 40% [7].

1.5.1 Multiple daily insulin injections

In the multiple daily insulin injections therapy, the subject injects a long acting insulin once or twice a day as a basal dose and has further injections of rapid acting insulin at each meal time [7]. Insulin is injected through regular syringes or pens, usually at least four times a day. The speed of action of rapid acting insulin allows people to administer correction doses between meals if the BG level rises too high.

1.5.2 Continuous subcutaneous insulin infusion

Diabetics treated with continuous subcutaneous insulin infusion use insulin pumps attached to the their body that deliver constant amounts of rapid or short acting insulin via a catheter placed under the skin [7]. Insulin pumps are portable devices that consist of a main pump unit holding an insulin reservoir attached to a long, thin piece of tubing with a needle or cannula at one end. These computerized devices allow to release insulin in two ways: in a steady, measured, and continuous dose, to provide

an appropriate basal insulin concentration, and as a surge dose (called bolus) around mealtime, to compensate for carbohydrate intake. Because the insulin pump stays connected to the body, it allows the wearer to modify the amount of insulin he needs to take by pressing a few buttons at any time of the day, or to program a higher or lower rate of insulin delivery to occur at a chosen time, e.g., during the night. Around 6% of adults and 19% of children with T1DM use an insulin pump [7]. Insulin pumps have also recently progressed into the so-called sensor-augmented insulin pumps, where the insulin pump acts as the receiver of the CGM data sent wirelessly from the sensor by the transmitter. In both adults and children with inadequately controlled T1DM, sensor-augmented pump therapy resulted in a significant improvement of HbA1c levels, as compared with the traditional therapy based on multiple daily insulin injections [29].

1.6 Conclusions

The development of vascular complications from diabetes can be managed and limited through the achievement of a tight glycemic control based on strict insulin therapy. Intensive treatment with three or more insulin administrations a day by injection or insulin pump was shown to lower BG concentration (as monitored by HbA1c levels) and delay the onset of retinopathy, nephropathy, and neuropathy in insulin-treated T1DM patients [30]. In the evaluation of the development of diseases from diabetes, however, not only hyperglycemia but also glycemic fluctuations and swings have to be considered [31]. Given that, the next chapter will focus on glucose variability, on its role in the development of complications from diabetes, and on the algorithms available to quantitatively measure it.

Chapter 2

The role of glucose variability in diabetes

2.1 The concept of glucose variability

To properly address the issue of developing long-term complications from diabetes, a sustained high BG level is not the only risk factor that has to be considered, and the occurrence of abnormal glucose variability (GV) needs to be taken into account as well. Before considering GV as a relevant element in the assessment of the metabolic situation of diabetic subjects, the major role in the evaluation of glycemic control was played by HbA1c, whose levels provide an integrated measure of glycemic exposure during the 2-3 months preceding the examination [32]. In 1993, landmark results from the Diabetes Control and Complications Trial (DCCT) demonstrated that lower HbA1c levels achieved in T1DM subjects under intensive insulin treatment led to a slower progression of diabetic complications with respect to conventionally treated patients [30]. These findings confirmed the association between hyperglycemia and the development of vascular diseases from diabetes, establishing HbA1c as the principal indicator of successful therapy and good control, with levels $\leq$7% deemed appropriate for reducing the development of diabetic complications [30, 33]. Later investigations made by the DCCT investigators themselves, however, showed that the risk of retinopathy progression associated with a given HbA1c level was significantly different between conventionally and intensively treated patients, suggesting that mean HbA1c could not be the most complete expression of the degree of glycemia and diabetic complications might be highly dependent on other factors [31]. Specifically, a greater increase in the risk of progression to retinopathy over time in participants under conventional

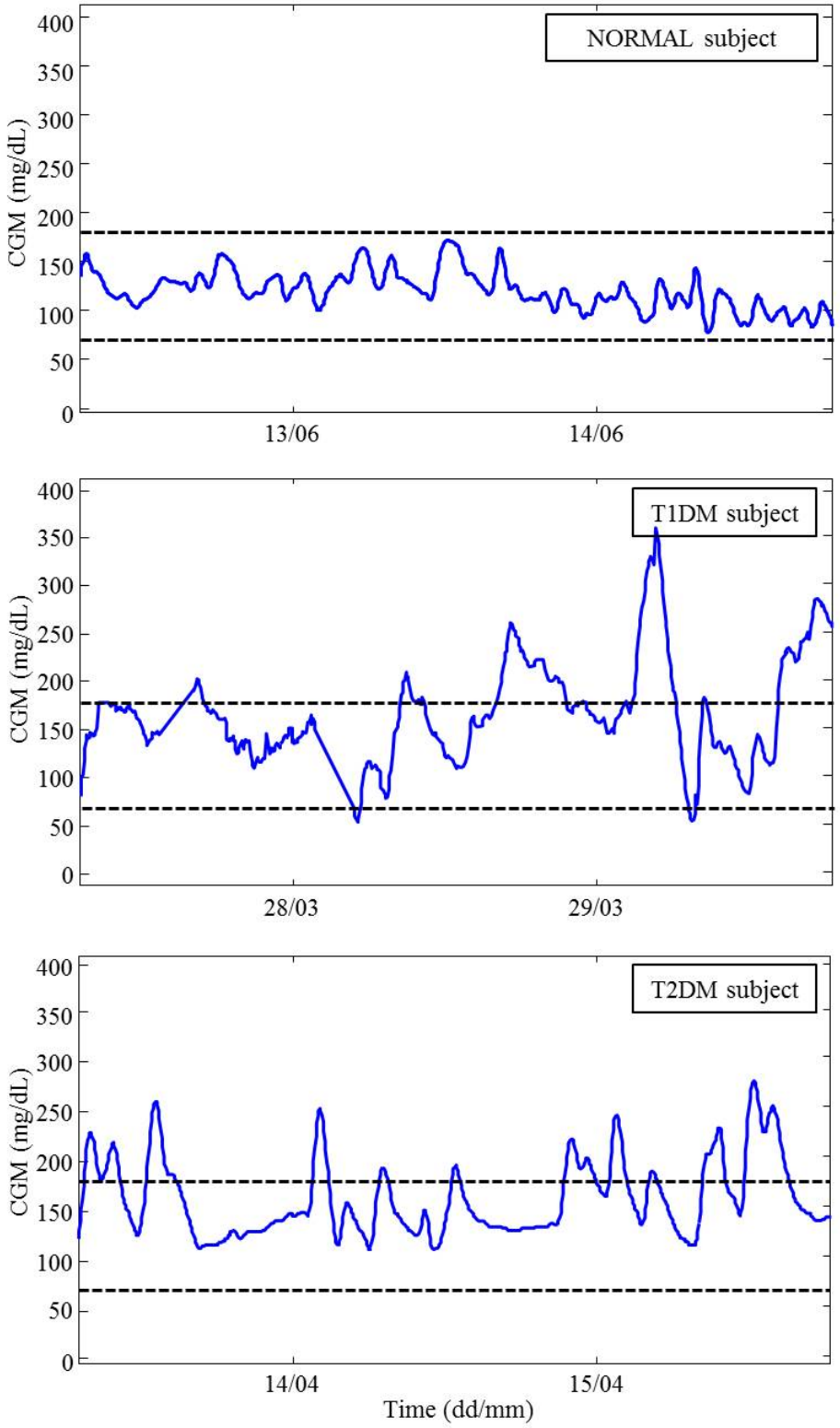

Figure 2.1 CGM profiles acquired from normal (top panel), T1DM (middle panel), and T2DM (bottom panel) subjects; black dotted lines at 70 and 180 mg/dL represent hypo and hyperglycemic thresholds, respectively. These examples are drawn from the datasets exploited in the study described in Chapter 5, where details about data collection can be found.

treatment was found, and this phenomenon was thought to be related to the more frequent high-amplitude fluctuations experienced by these patients, that might lead to increased generation of reactive oxygen species and oxidative stress [34, 35]. These early findings about a significant role of GV in the treatment of diabetes were progressively supported by a number of other investigations, which together developed a growing evidence that GV should be regarded as an actual HbA1c-independent risk factor for the development of long-term vascular complications from diabetes [36–42]. In view of this and of the need of reducing GV in order to achieve desired levels of glycemic control [43–45], tens of methods to provide an objective evaluation of GV from BG profiles have been designed. Before detailing some of them in the next section, it could be worth providing a qualitative description of the meaning of abnormal GV by resorting to representative CGM time-series collected from a normal, a T1DM, and a T2DM subject and reported in Figure 2.1. Abnormal GV typically refers to the occurrence of repeated high-amplitude glycemic fluctuations and frequent hypo/hyperglycemic events, that, as can be seen from Figure 2.1, might be experienced by people suffering from either type of diabetes. On the other hand, a well-controlled normal subject usually has his glycemia within the safe euglycemic range, without significant excursions (see the top panel of Figure 2.1).

2.2 Indices to measure glucose variability

The assessment of GV has been a widely investigated research field, and several indicators to retrospectively quantify GV from SMBG and CGM time-series have been proposed in the literature. Extensive reviews where all metrics are detailed are available, e.g., in [44–50]; here, we limit the presentation to 25 indicators that we exploited in our investigations, detailed in two different subsections to better organize their description within the text. Among them, some indices are more trivial, like sample mean and standard deviation (SD) of BG readings, while some others have been specifically designed for the quantification of GV. In all definitions, the glycemic time-series on which the index is computed is indicated as $\{\mathrm{BG}(i)\}$, $i = 1, \ldots, N$, and when present, N_{days} is the duration of the recording in terms of days.

2.2.1 Indices based on the distribution of glycemic levels

This first subsection provides a list of 13 GV indices whose formulations are mostly derived from the distribution of BG levels and the amplitude of glycemic excursions.

These metrics are:

- *Mean of BG readings* (Mean)

$$\text{Mean} = \frac{1}{N}\sum_{i=1}^{N}\text{BG}(i) \tag{2.1}$$

  simple and classic indicator mostly influenced by hyperglycemia which shows the best correlation with HbA1c levels [43, 51, 52];

- *SD of BG readings* (SD)

$$\text{SD} = \sqrt{\frac{\sum_{i=1}^{N}\left(\text{BG}(i) - \text{Mean}\right)^2}{N}} \tag{2.2}$$

  simple statistical method exploited, e.g., in [43, 53, 54], which provides a measure of variability but is influenced by non-Gaussian distributions and outliers [45];

- *Percentage coefficient of variation of BG readings* (%CV)

$$\%\text{CV} = 100\,\frac{\text{SD}}{\text{Mean}} \tag{2.3}$$

  variability measure used in statistics which normalizes sample SD to sample mean and is subject to the same limitations as SD [45];

- *Mean over all days of within-day SD of BG readings* (SD_w)

$$\text{SD}_w = \frac{1}{N_{days}}\sum_{j=1}^{N_{days}}\text{SD}(j) \tag{2.4}$$

  with $\text{SD}(j)$ being the SD of all measurements collected during the j^{th} day; defined in [44] and highly correlated with SD [45];

- *SD over all days of within-day mean of BG readings* (SD_{dm})

$$\text{SD}_{dm} = \sqrt{\frac{\sum_{j=1}^{N_{days}}\left(\text{Mean}(j) - \text{Mean}_{AllDays}\right)^2}{N_{days}}} \tag{2.5}$$

  with $\text{Mean}(j)$ being the mean of all measurements collected during the j^{th} day and $\text{Mean}_{AllDays}$ the average value of $\text{Mean}(j)$ over j (that can be different with

respect to Mean because the number of BG readings per day may vary from a day to another); as SD_w, defined in [44] and correlated with the overall SD [45];

– *Median of BG readings* (Median)

$$\text{Median} = 50^{\text{th}} \text{ percentile} \tag{2.6}$$

with X^{th} percentile being the value below which the X% of observations in a group of observations fall; simple statistical measure usually more suitable to the analysis of glycemic time-series than the sample mean because of the skewness of BG distribution, exploited, e.g., in [53–56];

– *Interquartile range of BG readings* (IQR)

$$\text{IQR} = 75^{\text{th}} \text{ percentile} - 25^{\text{th}} \text{ percentile} \tag{2.7}$$

difference between 75^{th} and 25^{th} percentiles, used, e.g., in [53–56];

– *Range of BG readings* (Range)

$$\text{Range} = \max(\text{BG}(i)) - \min(\text{BG}(i)) \tag{2.8}$$

difference between higher and lower BG reading, exploited, e.g., in [45];

– *Percentages of BG readings within* (%BG in target)*, below* (%BG below target)*, and above* (%BG above target) *the target range 70-180 mg/dL*

$$\%\text{BG in target} = \frac{\#\text{ of BG}(i):\ 70 \leq \text{BG}(i) \leq\ 180}{N} \tag{2.9}$$

$$\%\text{BG below target} = \frac{\#\text{ of BG}(i):\ \text{BG}(i) < 70}{N} \tag{2.10}$$

$$\%\text{BG above target} = \frac{\#\text{ of BG}(i):\ \text{BG}(i) > 180}{N} \tag{2.11}$$

indicators used, e.g., in [57–59], simple and direct, but accounting only for the number of extreme BG fluctuations and not for their amplitude;

– *J-index* (J-index)

$$\text{J-index} = 0.001\ (\text{Mean} + \text{SD})^2 \tag{2.12}$$

combination of Mean and SD defined in [60], exploited, e.g., in [45], and shown to be highly correlated with %BG in target [47];

– *Mean amplitude of glycemic excursions* (MAGE)

$$\text{MAGE} = \frac{1}{N_{excursions}} \sum_{k=1}^{N_{excursions}} \Delta\text{BG}(k) \tag{2.13}$$

where $N_{excursions}$ is the number of glucose excursions having amplitude greater than one time the SD of the considered segment and $\Delta\text{BG}(k)$ is the amplitude of the k^{th} considered excursion; index defined in [61], typically computed for each day of monitoring and then averaged over all days, and shown to be highly correlated with the overall SD [45]; an algorithmic implementation of MAGE for the smart calculation on CGM time-series was provided by Baghurst in [62]; examples of use of this metric can be found in [44, 45].

2.2.2 Indices based on risk and quality of glycemic control

Twelve indices derived from nonlinear transformations of the BG range are presented in this section. These metrics are:

– *M_R with $R = 100$ mg/dL* (M_{100})

$$\text{M}_{100} = \frac{1}{N} \sum_{i=1}^{N} 1000 \left| \log_{10} \left(\frac{\text{BG}(i)}{100} \right) \right|^3 \tag{2.14}$$

indicator defined in [63, 64] and exploited, e.g., in [45];

– *Low blood glucose index* (LBGI)*, high blood glucose index* (HBGI)*, average daily risk range* (ADRR)*, and blood glucose risk index* (BGRI)

These metrics were designed to quantify the risk of hypoglycemia (LBGI), hyperglycemia (HBGI), and the overall quality of glycemic control (ADRR and BGRI) from SMBG time-series. Their formulation was derived from a transformation of the BG scale detailed by Kovatchev *et al.* in [65], aimed at symmetrizing the range of all possible BG values. The reason behind the transformation is the significantly different size of hypoglycemic (BG<70 mg/dL) and hyperglycemic (BG>180 mg/dL) ranges, condition that leads to a target BG range (70-180 mg/dL [30]) not centered in the range of all possible BG levels (assuming 20-600 mg/dL), causes the clinical center of the BG range (around 112.5 mg/dL) not to match the numerical one (around 300 mg/dL), and makes the distribution of BG measurements skewed and asymmetric. Based on this transformation of the BG

scale, a BG risk function that assigns a certain degree of risk to each BG reading was designed and used to subsequently define LBGI and HBGI. Symmetrization and risk functions are detailed below (Equation 2.15 and 2.16, respectively)

$$\mathrm{f(BG)} = 1.509\,([\log(\mathrm{BG})]^{1.084} - 5.381) \tag{2.15}$$

$$\mathrm{r(BG)} = 10\,\mathrm{f(BG)}^2\,. \tag{2.16}$$

The effect of the symmetrization function on the original BG range can be immediately drawn from its graphical representation reported in the left panel of Figure 2.2. The transformation, that makes the BG readings normally distributed, stretches the hypoglycemic range, shrinks the hyperglycemic range, and leads to a BG scale symmetric around zero. In the transformed range, thus, the euglycemic value of 112.5 mg/dL becomes the clinical and numerical center of the BG scale, and the target BG range results to be centered in the range of all possible BG values. The BG risk function then defined, shown in the right panel of Figure 2.2 as plotted against the original BG scale, ranges from 0 to 100, assigning zero risk to BG readings at 112.5 mg/dL and maximum risk to severe hypoglycemia (BG around 20 mg/dL) and extreme hyperglycemia (BG around 600 mg/dL).

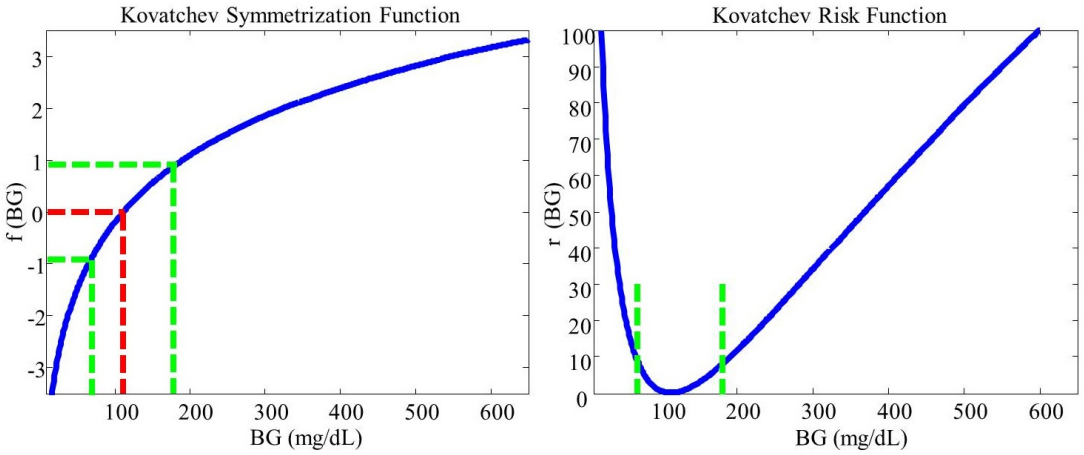

Figure 2.2 Symmetrization and risk functions proposed by Kovatchev *et al.* in [65]. Red lines indicate the clinical center at 112.5 mg/dL and its transformed value; green lines indicate the target range 70-180 mg/dL and its transformed value.

Based on the BG risk function, two further functions were designed. They represent the risk of low BG (rl(BG)) and high BG (rh(BG)), are specified as follows

$$\mathrm{rl(BG)} = \begin{cases} \mathrm{r(BG)}, & \text{if BG} < 112.5 \text{ mg/dL} \\ 0, & \text{otherwise} \end{cases} \tag{2.17}$$

$$\mathrm{rh(BG)} = \begin{cases} \mathrm{r(BG)}, & \text{if BG} > 112.5 \text{ mg/dL} \\ 0, & \text{otherwise} \end{cases} \tag{2.18}$$

and allow the definition of the final LBGI and HBGI [66, 67]

$$\mathrm{LBGI} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{rl(BG)} \tag{2.19}$$

$$\mathrm{HBGI} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{rh(BG)}\,. \tag{2.20}$$

As far as the indices assessing the overall quality of glycemic control are concerned, the ADRR was defined as the sum of the risk associated to the lower and higher BG readings of the j^{th} day averaged over all days [68]

$$\mathrm{ADRR} = \frac{1}{N_{days}}\sum_{j=1}^{N_{days}} \left[\max(\mathrm{rl}^j(\mathrm{BG})) + \max(\mathrm{rh}^j(\mathrm{BG}))\right], \tag{2.21}$$

and the BGRI was designed as the sum of LBGI and HBGI [46]

$$\mathrm{BGRI} = \mathrm{LBGI} + \mathrm{HBGI}\,. \tag{2.22}$$

A description of the clinical relevance of LBGI and HBGI, and of their use in the literature is discussed in the next chapter.

– *Hypoglycemic index* (Hypo Index), *hyperglycemic index* (Hyper Index), *and index of glycemic control* (IGC)

Hypo Index, Hyper Index, and IGC are metrics assessing the risk of hypoglycemia, hyperglycemia, and the overall quality of glycemic control, as the previous ones, but are derived from different transformation functions. They

were introduced by Rodbard in [45] and their formulations are detailed below

$$\text{Hypo Index} = \frac{\sum_{i=1}^{N} (\text{LLTR} - \text{BG}(i))^b}{N \times d} \tag{2.23}$$

$$\text{Hyper Index} = \frac{\sum_{i=1}^{N} (\text{BG}(i) - \text{ULTR})^a}{N \times c} \tag{2.24}$$

with IGC being the combination of the previous two

$$\text{IGC} = \text{Hypo Index} + \text{Hyper Index} \tag{2.25}$$

and LLTR, ULTR standing for lower and upper limit of target range with default values at 80 and 140 mg/dL, and a, b, c, and d parameters generally set to 1.1, 2.0, 30, and 30.

– *Glycemic risk assessment diabetes equation* (GRADE)*, and percentage contributions due to euglycemia* (%GRADE_{eu})*, hypoglycemia* (%$\text{GRADE}_{\text{hypo}}$)*, and hyperglycemia* (%$\text{GRADE}_{\text{hyper}}$)

These indicators have been proposed by Hill in [69, 70], and provide a measure of the quality of glycemic control summarized in the GRADE score, and the contribution to this score respectively due to euglycemia, hypoglycemia, and hyperglycemia. The GRADE score is defined as the average over all BG readings of the following transformed value

$$\text{GRADE} = 425 \left(\log[\log(\text{BG}/18)] + 0.16\right)^2 \tag{2.26}$$

and the three contributions due to the different glycemic conditions are

$$\%\text{GRADE}_{\text{eu}} = 100 \, \frac{\sum_{\text{BG in euglycemia}} \text{GRADE}}{\sum_{\text{all BG}} \text{GRADE}} \tag{2.27}$$

$$\%\text{GRADE}_{\text{hypo}} = 100 \, \frac{\sum_{\text{BG in hypoglycemia}} \text{GRADE}}{\sum_{\text{all BG}} \text{GRADE}} \tag{2.28}$$

$$\%\text{GRADE}_{\text{hyper}} = 100 \, \frac{\sum_{\text{BG in hyperglycemia}} \text{GRADE}}{\sum_{\text{all BG}} \text{GRADE}} \, . \tag{2.29}$$

Examples of their use and characterization can be found in [45].

2.3 Open issues and aim of the thesis

The review of techniques for the quantification of GV proposed in the previous sections provides the state of the art about the assessment of GV in diabetes, and simultaneously rises some issues still unsolved within this research field. To give an example, certain GV indices have been defined and validated only on SMBG profiles, and thus their use on CGM time-series is not immediate. This is the case, e.g., of LBGI and HBGI, popular metrics that can be exploited to provide a rapid classification of the overall quality of a subject's glucose control if they are computed on SMBG time-series, but have not a validated physiological meaning when applied to CGM profiles. Moreover, if we think of all metrics available for GV quantification, we can observe that the information conveyed by them is highly redundant and a compact way to describe GV is still missing in the literature. Finally, the exploitation of CGM signals and GV to classify the metabolic condition of normal and diabetic subjects is a relatively unexplored problem that could deserve an investigation. These three topics are the object of this dissertation. In particular, in the following chapters we will:

- provide alternate versions of the well-known LBGI and HBGI that are adapted to the characteristics of CGM time-series, thus allowing to extend the computation of these risk indices from SMBG to CGM profiles (Chapter 3);
- explore the use of the sparse principal component analysis (SPCA) technique to provide a parsimonious but still comprehensive description of GV in both T1DM and T2DM (Chapter 4);
- explore the possibility of using the outputs obtained from SPCA to develop GV-based classifiers of the quality of glycemic control in T1DM subjects and of the metabolic condition of normal subjects at risk of developing T2DM, IGT subjects, and subjects with T2DM diagnosed (Chapter 5).

Chapter 3

Adapting SMBG-based indices to CGM: the case of LBGI/HBGI [1]

3.1 Aim of the investigation

LBGI and HBGI are popular metrics used to quantify the risk of hypo and hyperglycemia from sparse SMBG profiles. As discussed in the previous section, they were designed based on a symmetrization of the BG range [65] to summarize the number and extent of extreme BG fluctuations into single numbers, with LBGI accounting for hypoglycemic episodes and HBGI for hyperglycemic ones. Therefore, a higher LBGI may indicate a large number of mild hypoglycemic events, a small number of severe hypoglycemic events, or a combination of both, and the same can be said for HBGI with regard to hyperglycemia.

The clinical relevance of these risk indicators has long been established as LBGI was shown to predict the occurrence of future severe hypoglycemia (SH) episodes [65, 66] and HBGI demonstrated a positive correlation with HbA1c levels [65]. Moreover, significant cutoff values that allow to promptly classify the quality of the overall glycemic control of a patient based on his LBGI and HBGI values were also identified [66, 67]. Specifically, the risk of SH can be determined by exploiting LBGI, and is considered low if LBGI < 2.5, moderate if LBGI is between 2.5 and 5, and high if LBGI > 5, and the same can be said for HBGI concerning the risk of hyperglycemia, with cutpoints

[1] The work presented in this chapter was developed under the supervision of Professor Marc Breton and represents part of the research activity performed during a period of six months spent at the Center for Diabetes Technology, University of Virginia (Charlottesville, VA, USA).

identified at HBGI = 4.5 and HBGI = 9.

Because LBGI and HBGI enable a rapid classification of the actual glycemic condition of a patient, their use in the quantification of GV was largely exploited in the literature [32, 38, 46, 48, 68, 71–78]. However, using these indicators and their cutoff values is possible only from SMBG profiles and has never been validated on CGM data stream. Given the growing use of CGM systems, thus, the aim of this work is to describe the relationship between SMBG-based and CGM-based LBGI/HBGI, and provide guidelines to apply when the indices are computed from CGM time-series. Specifically, based on a dataset of 28 T1DM subjects monitored with both SMBG and CGM for up to four weeks, we identified transformation functions for LBGI and HBGI that adapt these indicators to the characteristics of CGM profiles and enable references to previous works and clinically relevant thresholds [79].

3.2 Dataset

3.2.1 Data collection

Data used in this work were collected at the University of Virginia (Charlottesville, VA, USA) and are baseline data from a study aimed at investigating how BG changes in response to insulin and what the body does to counteract low BG in people with T1DM. Twenty-eight T1DM subjects were involved in the analysis. All subjects were under insulin pump therapy, and monitored for up to four weeks with both SMBG and CGM systems. After two weeks, they received a liquid mixed-meal in an inpatient setting, and had their BG monitored in order to study insulin sensitivity; then, they received additional insulin injections to cause a condition of low BG, to understand how the body responds to hypoglycemia. All subjects were closely monitored during the time insulin was given, by frequent checks of BG and constant medical and nursing supervision. Beyond the inpatient day, subjects were in daily-life conditions. During the study, participants were asked to use their own insulin pump and glucometer, and the same glucometer had to be used for the entire study. CGM profiles were collected using the DexCom® G4® glucose sensor [22]. Demographic characteristics of the subjects expressed as mean±SD were age 43±11 years, duration of diabetes 23±11 years, duration of pump therapy 11±7 years, and HbA1c 8±1%. The study was approved by the local ethical committee and registered at ClinicalTrials.Gov with the identifier number NCT01835964.

3.2.2 Data preprocessing and characterization

To avoid any influence of the results on the protocol used for data collection, the inpatient day was removed from SMBG and CGM profiles. Days where sensor issues occurred and no CGM points were stored were removed also from the SMBG time-series. Beyond this, all datapoints collected were kept in the study. The subject storing the little number of SMBG samples had an SMBG profile of 2.61 points/day on average with 73 measurements acquired over the entire four weeks. The longest profile was made up of 315 datapoints, with 11.25 points/day on average. Mean value over the whole population of the average length per day was 5.65 samples and of the entire length was 158.26. CGM data were collected with a sampling time of five minutes. Figure 3.1 shows an example of data collected from a representative subject; Figure 3.2 shows histograms of the distribution of SMBG and CGM datapoints across the time of the day, where it can be seen that CGM data are acquired almost constantly throughout the day, while SMBGs are mostly acquired around meal times.

3.3 Design of the strategy to correct LBGI/HBGI

The relationship between CGM-based and SMBG-based LBGI/HBGI was described using linear and nonlinear models. The acceptance/rejection of a model was made through the assessment of the obtained fit, statistical significance of parameter estimates, residuals, QQ-plot of residuals, and performances in terms of classification of the subjects' risk of hypo/hyperglycemia as compared to the SMBG-based one.
Defining the index computed on CGM as the independent variable or predictor (x) and that computed on SMBG as the dependent one (y), the linear model was the canonical straight line described as

$$y = mx + q\,. \tag{3.1}$$

When a liner description was not satisfactory to match CGM-based and SMBG-based LBGI/HBGI, the following three-parameter nonlinear model was considered

$$y = ax + b - \frac{1}{c + x}\,. \tag{3.2}$$

The number of model parameters in 3.2 was also reduced by imposing the reasonable constraint of having a curve meeting the origin of the Cartesian coordinate system.

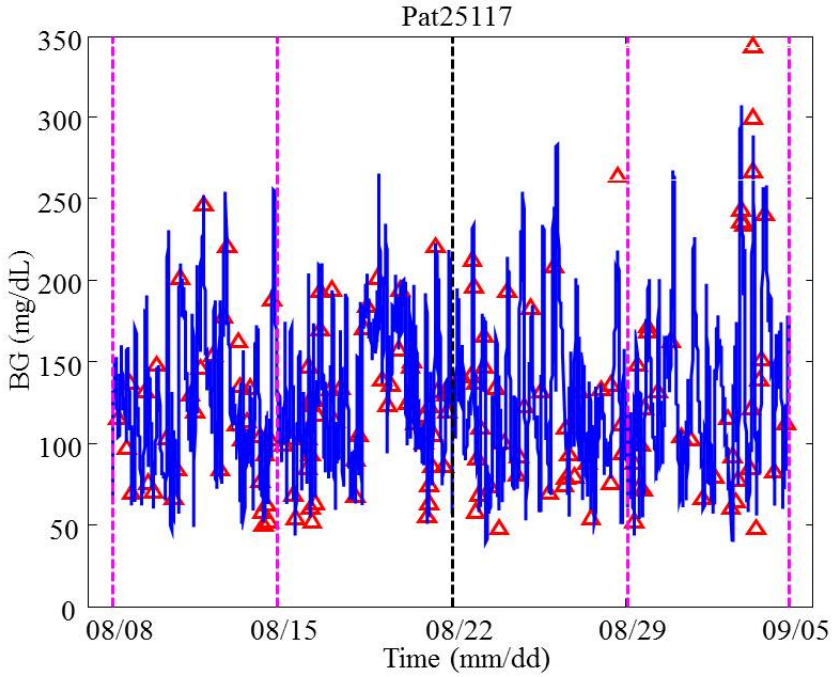

Figure 3.1 Example of data used to tune the LBGI/HBGI transformation functions. SMBG and CGM data are shown as red triangles and blue profile, respectively; pink lines identify the four week of monitoring; black line is at noon of the inpatient day.

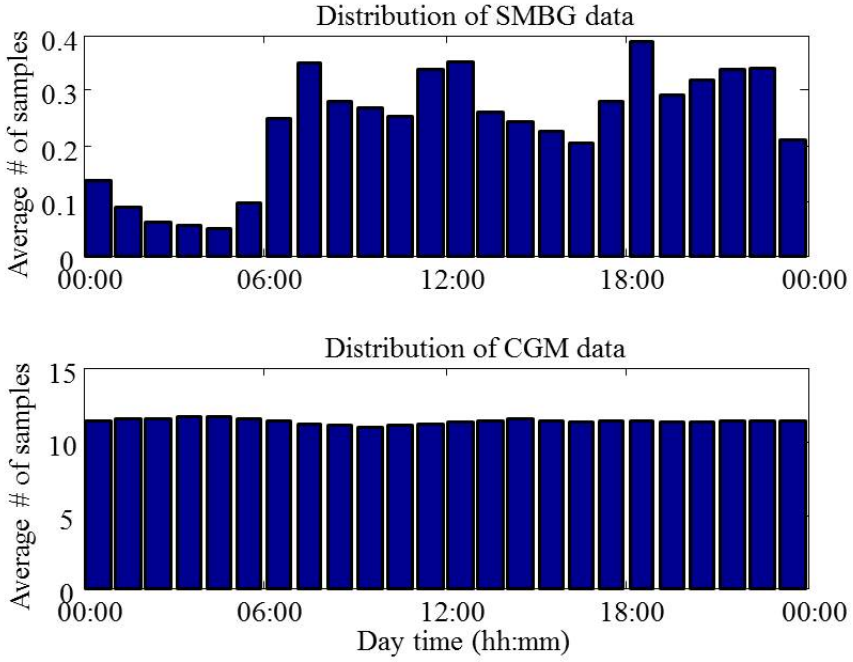

Figure 3.2 Distribution of SMBG (top panel) and CGM (bottom panel) datapoints across the day. Values were obtained as average over all days for each subject, and then average over all subjects.

This led to the following two-parameter formulation

$$y = ax + b - \frac{1}{b + x} \,. \tag{3.3}$$

Models were identified on the available data using the least squares estimator.

3.4 CGM-adapted risk indices

3.4.1 HBGI for CGM

The linear model detailed by Equation 3.1 was used to describe the relationship between SMBG-based and CGM-based HBGI. Fit, residuals, and QQ-plot of residuals are shown in Figure 3.3, upper, lower-left, and lower-right panel, respectively; results from model identification are summarized in the upper rows of Table 3.1. The table shows the estimates of parameters, their precision expressed as standard error (SE), t statistics for testing the null hypothesis that parameters are zero, and p-values associated with t statistics. A p-value lower than 0.05 was considered to be statistically significant and allowed the rejection of the null hypothesis. As apparent from the top panel of Figure 3.3, fit is extremely satisfactory; the mean squared error (MSE) for this model was equal to 6.2 and a value of 75% was obtained for the coefficient of determination (R^2), revealing that the model can account for 75% of the variance of the original SMBG-based HBGI values. Looking at residuals, they appear sufficiently uncorrelated and, except from some outliers probably related to biasing SMBG sampling designs, QQ-plot suggests to consider them as drawn from a normal distribution. The last column of Table 3.1, however, shows that the p-value associated with q is much greater than 0.05, indicating that the parameter is not statistically significant within the model. Because of this finding, a linear model with slope and intercept forced to zero was then identified on the experimental data. Fit, residuals, and QQ-plot of residuals are shown in Figure 3.4, upper, lower-left, and lower-right panel, respectively, and results are summarized in the lower rows of Table 3.1. Fit and residuals are as satisfactory as for the previous model, with MSE equal to 6, and the obtained estimate for m is extremely close to one, suggesting the possibility of using no transformation to compute HBGI from CGM time-series.

To assess the reliability of this result, the CGM-based classification of the patients' risk of hyperglycemia was compared with the gold-standard SMBG-based one. Table 3.2 shows how subjects were assigned to the different risk classes in the two cases (left

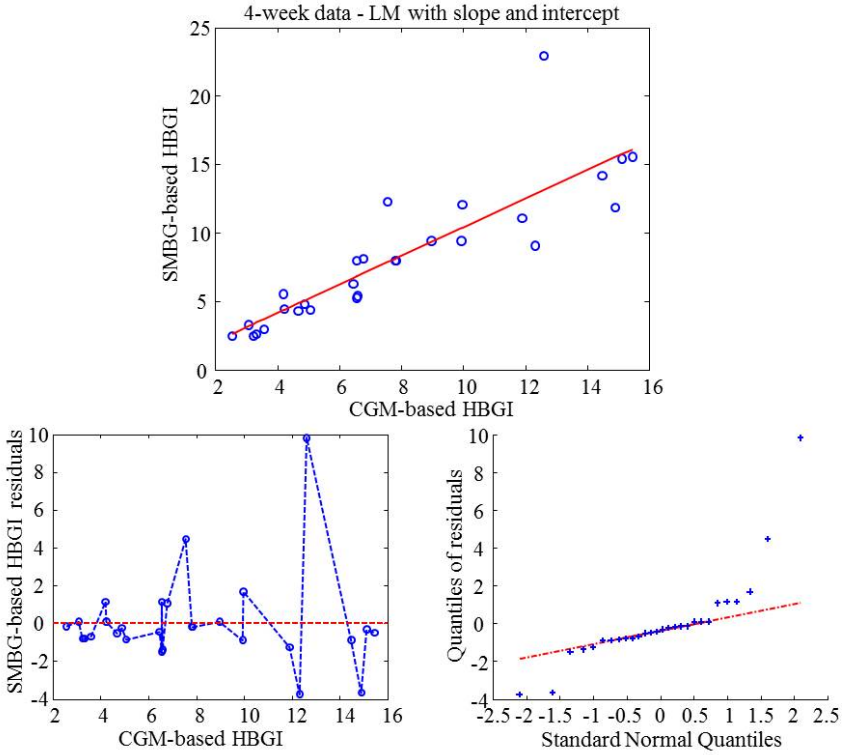

Figure 3.3 Model prediction (red line) and experimental data points (blue circles) for HBGI correction as described with a full linear model (upper panel), together with residuals (lower-left panel) and their QQ-plot (lower-right panel).

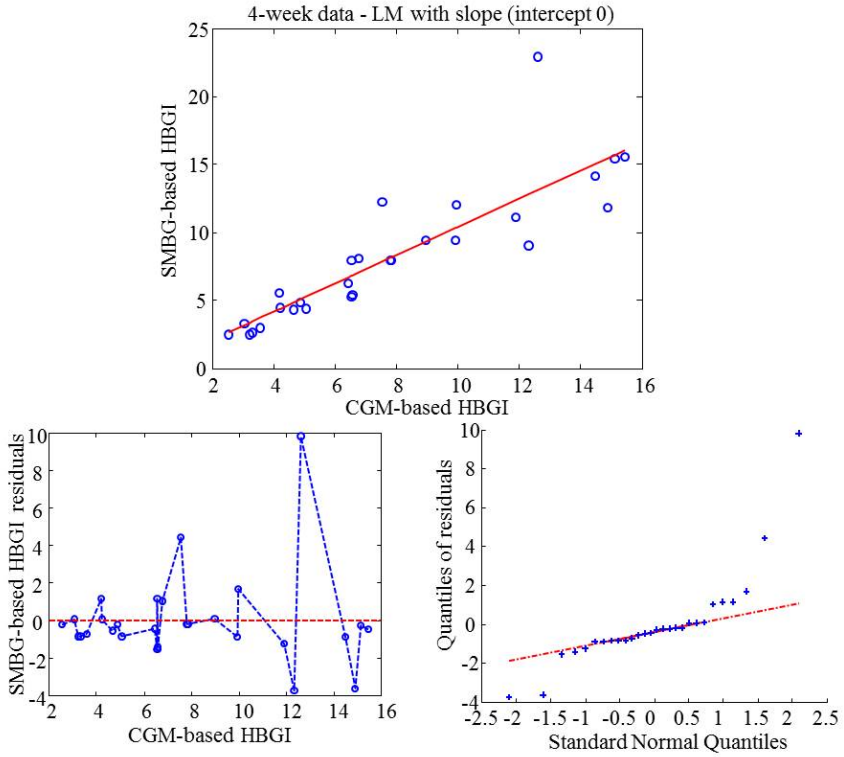

Figure 3.4 Model prediction (red line) and experimental data points (blue circles) for HBGI correction as described with a linear model with slope and intercept forced to zero (upper panel), together with residuals (lower-left panel) and their QQ-plot (lower-right panel).

columns), and what kind of errors were computed in the CGM-based classification (right columns). These results are satisfactory not only because 82% of patients were classified correctly, but also observing that misclassification was of only one class (no subject at low risk was classified at high risk, and vice versa) and extremely balanced in over and underestimation. This suggests that there is no systematic error that needs to be corrected, and classification errors are random and most likely linked to the potentially biasing designs used to collect SMBG readings and/or to possible miscalibration of the sensors.

Given these results, the one-parameter linear model was considered reliable to describe the relationship between SMBG-based and CGM-based HBGI, and we could conclude that HBGI as derived from CGM is interchangeable with SMBG-based HBGI.

HBGI				
Linear model with slope and intercept				
	estimate	**SE**	***t***	**p-value**
m	1.0436	0.1182	8.8319	$2.6354 \cdot 10^{-9}$
q	-0.0619	1.0451	-0.0593	0.9532
Linear model with slope (intercept forced to zero)				
	estimate	**SE**	***t***	**p-value**
m	1.0374	0.0523	19.8356	$3.2566 \cdot 10^{-17}$
q	0	–	–	–

Table 3.1 Results from the identification of a linear model with and without intercept for the correction of HBGI.

low risk	**moderate risk**	**high risk**	**underest errors**	**UNDEREST errors**	**overest errors**	**OVEREST errors**
SMBG based classification						
8	9	11	–	–	–	–
CGM based classification						
7	11	10	3	0	2	0

Table 3.2 Classification of the risk of hyperglycemia as assessed by HBGI computed on SMBG and CGM. On the left side of the table, subjects assigned to each risk group are reported; on the right side of the table, under and overestimation errors due to the assessment of HBGI on CGM are listed (underest/overest errors stand for misclassification of one class; UNDEREST/OVEREST errors stand for misclassification of two classes).

3.4.2 LBGI for CGM

The first model identified on SMBG-based LBGI values was the linear one (Equation 3.1), that provided parameter estimates summarized in the upper rows of Table 3.3. Along with parameter values, as for HBGI, Table 3.3 shows their precision expressed in terms of SE, t statistics to test the null hypothesis that parameters are equal to zero, and p-values to be considered in accepting or rejecting that hypothesis. For the linear model, both parameters were statistically significant, MSE of 0.68 was obtained, and R^2 resulted equal to 76%. However, despite acceptable residuals (lower-left and lower-right panel of Figure 3.5) it appears clear that a straight line cannot provide a reliable description of the relationship between SMBG-based and CGM-based LBGI at lowest LBGI values (upper panel of Figure 3.5). We, thus, moved to the identification of the nonlinear model 3.3. Parameter estimates obtained in this case are reported in the lower rows of Table 3.3; the prediction is plotted against experimental data in the upper panel of Figure 3.6, and residuals and QQ-plot are shown in the lower-left and lower-right panel of Figure 3.6, respectively. Again both parameters were statistically significant and estimated with good precision. The MSE obtained for this model was equal to 0.62, and residuals are satisfactory. Nonetheless, in choosing the best model to be used to transform CGM-based LBGI, it is worth observing that the nonlinear model shows better performance than the linear one only at very low LBGI values, where there is no significant glycemic risk and a lower degree of precision in describing the data could be still adequate.

To assess the actual need for a correction and to investigate if the nonlinear model could make any significant difference, the comparison of the classification of subjects' risk of hypoglycemia was performed as done for HBGI. Table 3.4 shows the assignment of subjects to the corresponding risk group based on LBGI computed from SMBG, CGM, linearly corrected CGM, and nonlinearly corrected CGM. We can observe from the table that the application of the linear transformation to correct the CGM-based LBGI values dramatically changed the assessment of the risk of hypoglycemia, with classification performances largely improved after the correction. Errors in identifying the risk zone decreased from 7 to 5 (with 82% of subjects correctly classified), became random and balanced in over and underestimation, and the two-class error performed without correcting LBGI was avoided. The same results were obtained applying the nonlinear transformation to CGM-based LBGI values.

From this analysis, we can conclude that the computation of LBGI from CGM time-series is biased, with a clear underestimation of the actual risk of hypoglycemia, and

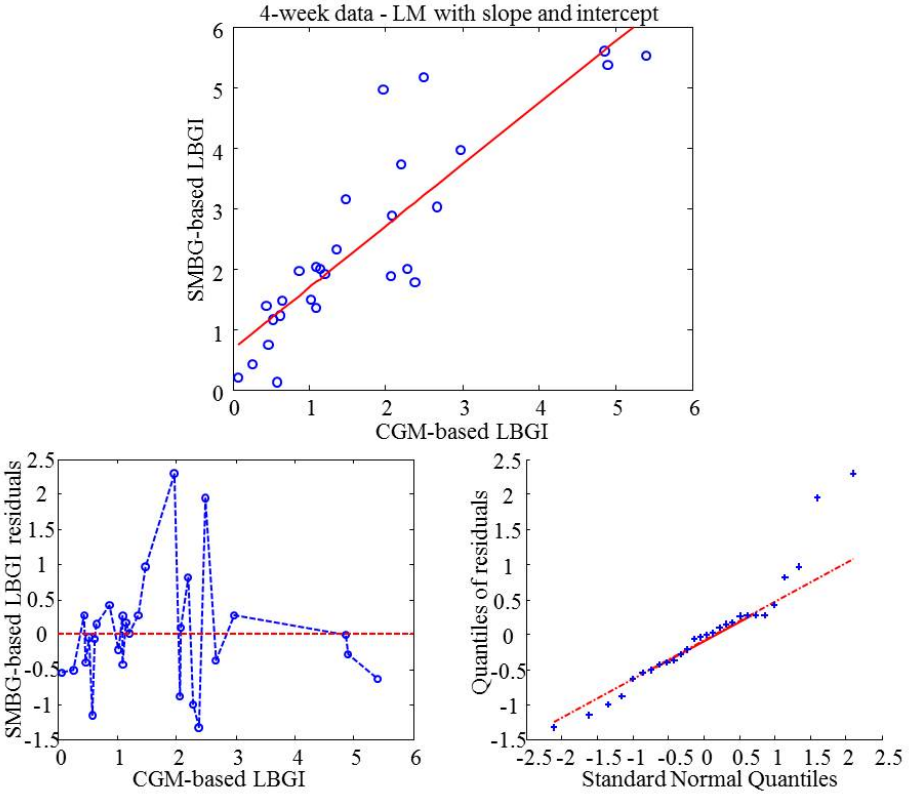

Figure 3.5 Model prediction (red line) and experimental data points (blue circles) for LBGI correction as described with a full linear model (upper panel), together with residuals (lower-left panel) and their QQ-plot (lower-right panel).

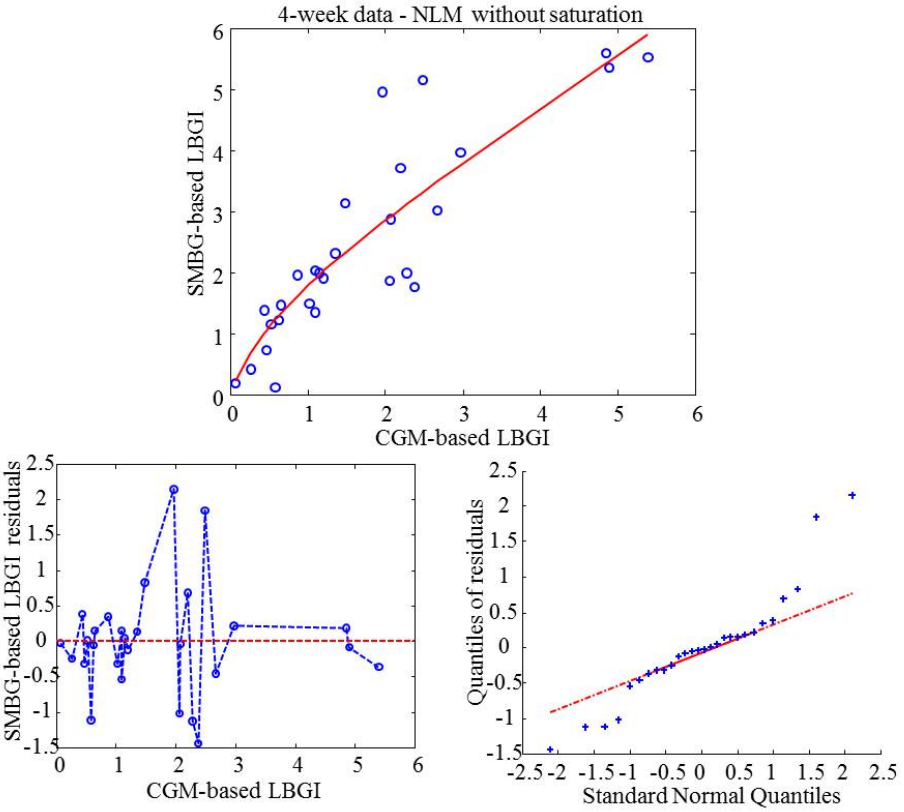

Figure 3.6 Model prediction (red line) and experimental data points (blue circles) for LBGI correction as described with a nonlinear model (upper panel), together with residuals (lower-left panel) and their QQ-plot (lower-right panel).

a correction is needed. Though the nonlinear transformation is more accurate at low LBGI values, no advantage was visible in terms of classification. As already mentioned, this is due to the fact that the nonlinear model works better in a very low risk zone, where a high degree of precision is not needed. Given that, the linear model can be considered satisfactory in describing the relationship between SMBG-based and CGM-based LBGI, turning out to be the transformation function that better adapts LBGI to the characteristics of CGM signals.

LBGI				
Linear model with slope and intercept				
	estimate	**SE**	***t***	**p-value**
m	1.0199	0.1125	9.0661	$1.5674 \cdot 10^{-9}$
q	0.6521	0.2519	2.5893	0.0156
Nonlinear model without saturation				
	estimate	**SE**	***t***	**p-value**
a	0.8346	0.1284	6.5026	$6.8204 \cdot 10^{-7}$
b	0.6476	0.1412	4.5876	$1.1107 \cdot 10^{-5}$

Table 3.3 Results from the identification of a linear model with intercept and a nonlinear model without saturation for the correction of LBGI.

low risk	**moderate risk**	**high risk**	**underest errors**	**UNDEREST errors**	**overest errors**	**OVEREST errors**
SMBG based classification						
8	9	11	–	–	–	–
CGM based classification						
23	4	1	6	1	0	0
LINEARLY corrected CGM based classification						
16	9	3	2	0	3	0
NONLINEARLY corrected CGM based classification						
16	9	3	2	0	3	0

Table 3.4 Classification of the risk of hypoglycemia as assessed by LBGI computed on SMBG and CGM. On the left side of the table, subjects assigned to each risk group are reported; on the right side of the table, under and overestimation errors due to the assessment of LBGI on CGM are listed (underest/overest errors stand for misclassification of one class; UNDEREST/OVEREST errors stand for misclassification of two classes). On the bottom of the table, we can observe the performance obtained when LBGI from CGM was conveniently corrected.

3.5 Conclusions

The risk of hypo and hyperglycemia can be assessed in a straightforward way by computing LBGI and HBGI on collections of SMBG readings. These indicators have been shown to be predictive of future glycemic events (e.g., LBGI of the frequency of future SH episodes) and clinically relevant cutoff values have been defined to classify the current condition of a patient based on his SMBG-based LBGI/HBGI. Because of their easy and fast implementation and immediate interpretation, LBGI and HBGI are powerful tools to assess glycemic variability and risks, and their applicability to CGM signals could be of significant impact both from a scientific and a clinical viewpoint. The aim of this study was, thus, defining alternate versions of LBGI and HBGI that fit better the characteristics of CGMs, providing transformations to apply when these indices are computed from CGM time-series and enabling references to previously published works and relevant cutoff values.

In doing that, a number of models were exploited to describe the relationship between SMBG-based and CGM-based LBGI/HBGI, and the optimal parameters to correct the indicators were identified. It is worth mentioning that, however, transformations and correction coefficients of LBGI/HBGI do not account only for the different sampling rate of SMBG and CGM systems, but also for some intrinsic characteristics of the specific sensor used for the continuous monitoring of glucose levels. This means that studies developed on data acquired with a CGM device different from the DexCom® G4® could provide additional information as to the optimal correction of LBGI and HBGI. Thus, further work shall consider the correction coefficients from a wider range of CGM systems and the potential need for a sensor-specific correction.

Beyond this, other analyses to systematically assess the influence of the sampling frequency on LBGI/HBGI values are worth to be computed. Simulation studies could be used to recreate CGM time-series and randomly downsample them up to recreating a full seven-point SMBG profile, with the aim of identifying the maximum number of BG samples that allows references to previous works without correction coefficients. Also, the effect of a significant decrease of the number of SMBGs should be investigated, to provide an inferior limit to be respected in order to obtain physiologically meaningful values of LBGI and HBGI. As an additional variable to be considered in all cases, the effect of the placement of SMBG samples across the day is likely to have a significant influence on the values obtained for LBGI and HBGI, and defining an optimal scenario for SMBG measurements could be really interesting and useful to get a uniformity, e.g., in clinical trials. These open issues will be matter of future works.

Chapter 4

Parsimonious description of glucose variability by SPCA

4.1 Aim of the investigation

As outlined in Chapter 2, a number of indicators for GV quantification have been proposed in the literature [43, 44, 46–50]. These metrics include indices derived from the distribution of glucose readings and the amplitude of glycemic excursions, measures of the risk of hyper and hypoglycemia, and indicators of the overall quality of glycemic control. The issue with the large pool of defined GV metrics, however, is that a gold-standard technique to assess GV has not been identified yet, the information carried by the indices is highly redundant, and some indicators may be of relatively minor added value in the characterization of GV within a diabetic population.

To manage this situation, we propose the use of SPCA, a technique introduced by Zou and co-workers in 2006 [80] that allows to extract a small combination of GV indices which preserves a large part of the variance originally conveyed by a wider pool of indicators. Based on this method, we can select a reduced subset of metrics that are able to effectively characterize the whole GV of a diabetic population, and thus we end up providing a parsimonious description of GV in diabetes.

In our investigation, SPCA was applied to the 25 GV indices introduced in the previous chapter, evaluated on two T1DM CGM datasets made up of 17 and 16 time-series, respectively, and on the one obtained by merging them. The dependency of the extracted GV indices from the specific T1DM dataset was assessed, and the application of the technique was then extended to a 13 CGM time-series T2DM dataset.

The applicability of LBGI and HBGI to CGM time-series was possible thanks to the

transformation functions previously described [79].
In the following of this chapter, Section 4.2 describes the SPCA technique, Section 4.3 provides details about data collection, and Sections 4.4 and 4.5 present the results obtained from the application of SPCA to T1DM [81] and T2DM [82], respectively, discussing the choice of parameters needed for the implementation of the technique.

4.2 SPCA

SPCA is a two-step data processing technique of general applicability introduced by Zou *et al.* in [80]. In our implementation, assuming that the n observations of the m GV metrics are stored along the columns of the $n \times m$ matrix $\mathbf{X}$, the two steps are:

A. apply the canonical principal component analysis (PCA) [83] to the matrix $\mathbf{X}$ to retrieve the principal components (PCs), and select a reduced number p ($p \ll m$) of them, saving the majority of the total original variance;

B. since PCs are linear combinations (regressions) of all GV indices and have not a direct physiological meaning, apply the least absolute shrinkage and selection operator (LASSO) constraint to each selected PC, to obtain sparse estimates of regression coefficients and maintain in the PC regressor a reduced number of metrics (those forming the parsimonious set of indices to describe GV).

Before entering SPCA, a normalization of the indices within each dataset was performed. Specifically, the indices were mean centered and scaled (i.e., divided by their sample SD), and this preprocessing allowed us to avoid any bias in the SPCA results and compare the outputs obtained from the application of SPCA to different datasets (the importance of this normalization is discussed in the Appendix to our paper [81]).
To have a quick overview of the implementation of SPCA in our problem, a block diagram highlighting the main steps of the procedure is shown in Figure 4.1.
Calculation of GV indices and implementation of SPCA were performed with software developed by us in MATLAB® R2012a.

A. PCA and formulation of the regression problem

Let the covariance matrix of the data $\mathbf{X}$ be a general full matrix (i.e., the m variables are correlated among them). Then, the aim of PCA is to find a linear transformation

$$\mathbf{Y} = \mathbf{X}\,\mathbf{S} \tag{4.1}$$

that allows to reduce the correlation of the original data. Formally, we are seeking for an orthogonal matrix $\mathbf{S}$ to transform the original data $\mathbf{X}$ into the new data $\mathbf{Y}$, under the constraint that the covariance matrix of $\mathbf{Y}$ is a diagonal matrix.
To solve the problem and find the transformation matrix $\mathbf{S}$, PCA resorts to the singular value decomposition of the data matrix, that allows to express $\mathbf{X}$ as

$$\mathbf{X} = \mathbf{U}\,\mathbf{\Sigma}\,\mathbf{V}^{\mathrm{T}} \tag{4.2}$$

with $\mathbf{U}$ being an $n \times n$ orthogonal matrix collecting the vectors $\mathbf{u}_i = \mathbf{X}\mathbf{v}_i/\sigma_i$, Σ an $n \times m$ rectangular diagonal matrix with nonnegative diagonal entries $\sigma_i = \sqrt{\lambda_i}$, known as singular values, and $\mathbf{V}$ an $m \times m$ orthogonal matrix whose columns are the eigenvectors of the matrix $\mathbf{X}^{\mathrm{T}}\,\mathbf{X}$, with associated eigenvalues λ_i. Given the orthogonality of $\mathbf{V}$, Equation 4.2 can be rewritten as

$$\mathbf{X}\,\mathbf{V} = \mathbf{U}\,\mathbf{\Sigma} \tag{4.3}$$

and from this statement it is straightforward to see that the transformation matrix $\mathbf{S}$ we are looking for is exactly the matrix $\mathbf{V}$ collecting the eigenvectors of $\mathbf{X}^{\mathrm{T}}\,\mathbf{X}$. As requested by the algorithm, in fact, $\mathbf{V}$ is orthogonal, and the new data expressed as $\mathbf{U}\,\mathbf{\Sigma}$ have a diagonal covariance matrix (because $\mathbf{U}$ is orthogonal and $\mathbf{\Sigma}$ is diagonal). Being $\mathbf{S} = \mathbf{V}$, the transformed data collected in the $n \times m$ matrix $\mathbf{Y}$ can be written as

$$\mathbf{Y} = \mathbf{X}\,\mathbf{V} \tag{4.4}$$

where the columns of $\mathbf{Y}$ are m new uncorrelated variables known as PCs of $\mathbf{X}$.
The main advantages of data decomposition through PCA is that PCs capture the maximum variability among the columns of $\mathbf{X}$ [80] and are sorted in decreasing order in terms of explained variance of the original data (i.e., the first greatest variance is explained by the first PC, the second greatest variance by the second PC, and so on); furthermore, PCs are uncorrelated and thus can be considered separately one from another. Typically, the application of PCA leads to the selection of p PCs, and data dimensionality can be reduced from m to $p \ll m$, with minimum loss of information.
Looking specifically at PCs, they are obtained as linear combinations of all the m initial variables, with nonzero coefficients (also called loadings, and referred to as $\alpha_j^{(i)}$ in Figure 4.1) collected along the columns of the matrix $\mathbf{V}$. Denoting now $\mathbf{V} = \mathbf{B}$, the problem of searching for the unknown loading matrix $\mathbf{B}$ can be seen as a regression

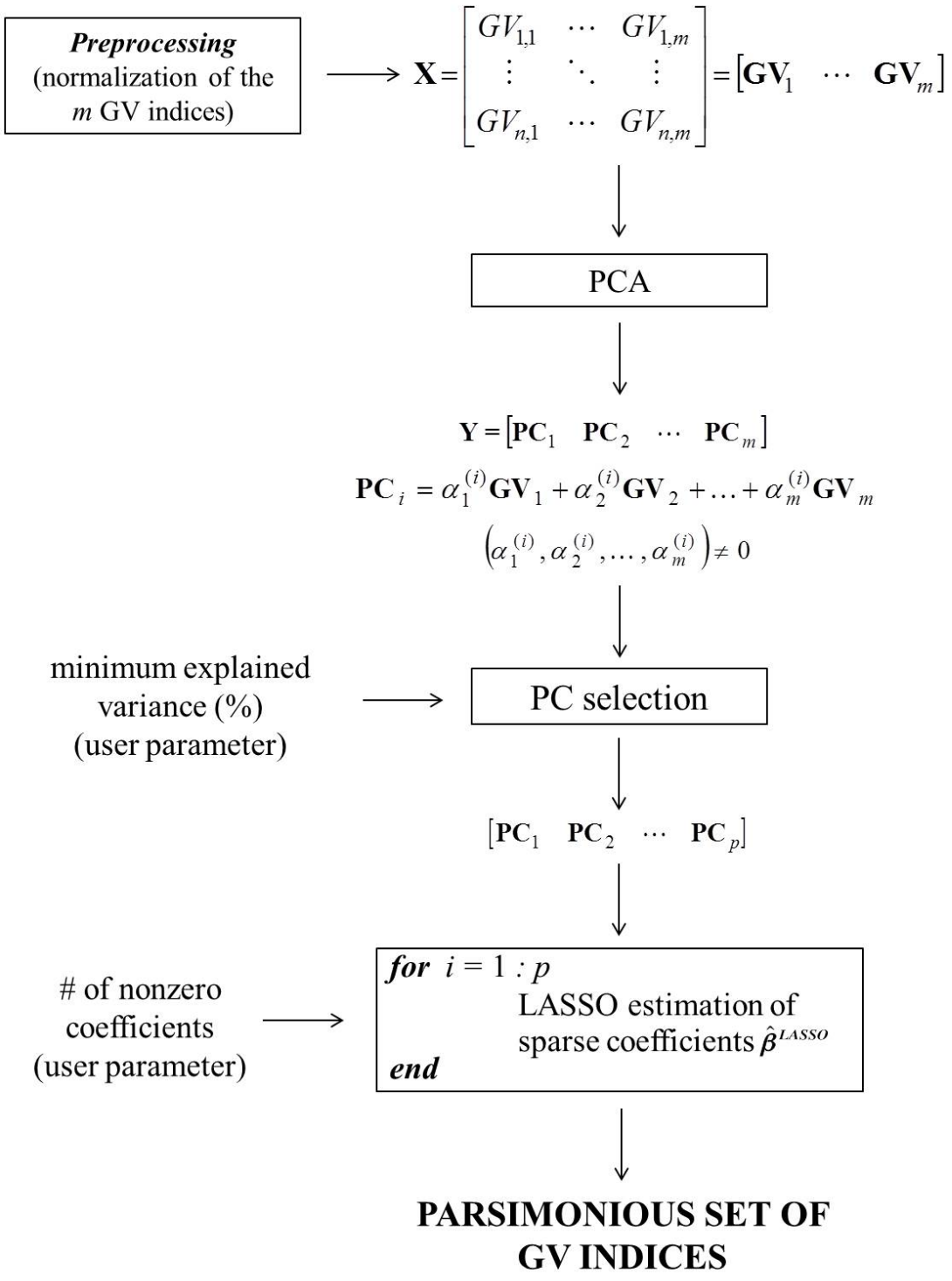

Figure 4.1 Schematic representation of the implementation of SPCA in our problem. The block diagram highlights the need of two user parameters, i.e., the number of PCs selected after PCA and the number of predictors kept in the regressors by the LASSO constraint. Also, inputs and outputs of each step are detailed.

problem. In particular, the i^{th} column of $\mathbf{Y}$ can be written as

$$\mathbf{y}_i = \mathbf{X}\,\boldsymbol{\beta} \tag{4.5}$$

and each column vector $\boldsymbol{\beta}$ of the loading matrix $\mathbf{B}$ can be estimated from the data $\mathbf{y}_i$ using the LASSO constraint.

B. LASSO estimation of sparse loadings

As just seen, PCA allows to reduce data dimensionality through the selection of p uncorrelated PCs. Each PC, however, results from a linear combination of all the original variables with coefficients $\alpha_j^{(i)}$ that are typically nonzero. The aim of the LASSO estimation is to reduce the number of explicitly used variables, through the calculation of sparse loadings. Defining $\widehat{\boldsymbol{\beta}}^{\mathbf{LASSO}}$ the estimated vector of sparse coefficients, it is obtained from the solution of the following optimization problem

$$\widehat{\boldsymbol{\beta}}^{\mathbf{LASSO}} = \underset{\beta}{\operatorname{argmin}} \left(\left|\left| \mathbf{y}_i - \sum_{j=1}^{m} \mathbf{X}_i\,\beta_j \right|\right|^2 + \lambda \sum_{j=1}^{m} |\beta_j| \right) \tag{4.6}$$

where $\mathbf{X}_j$ is the j^{th} column of the data matrix $\mathbf{X}$ collecting the n observations of the same variable. As one can see from Equation 4.6, the cost function here defined is made up of two different terms, i.e., the residual sum of squares and the sum of coefficient absolute values; the role of each term within the optimization problem is determined by the nonnegative parameter λ. Specifically, high values of λ lead to a large number of coefficients shrunk to zero and a small number of variables kept in the PC regressor, while λ values close to zero lead to a full estimated loading vector. In practice, λ is set based on the requirements of the considered problem, and, once specified, if large enough, allows to select a reduced number of variables from the larger initial pool; the ensemble of the selected variables is usually able to preserve a great part of the variance originally explained by the whole set of metrics.

4.3 Datasets

T1DM data considered for the analysis are two CGM datasets collected within the EU FP7 project Diadvisor (2008-2012) [84] using the Dexcom® SEVEN® Plus CGM System (Dexcom, San Diego, CA, USA) [22]. In this work, dataset 1 is made up of CGM time-series acquired from 17 T1DM males studied at the Department of Medicine,

University of Padova (Padova, Italy), while dataset 2 comprises 16 T1DM CGM profiles collected at the Department of Endocrinology, Diabetes, Nutrition, University of Montpellier (Montpellier, France). A third dataset of 33 CGM time-series obtained by combining datasets 1 and 2 was also considered in the analysis.
The T2DM CGM dataset, on the other hand, is made up of CGM time-series collected from 13 males under normal life conditions using the Medtronic® Guardian REAL-Time® CGM System (Medtronic, Northridge, CA) [85]. Mean±SD demographic characteristics of the subjects are age 54.4±10.0 years, duration of diabetes 10.7±7.0 years, HbA1c 7.4±1.4 %, and BMI 29.4±3.6 kg/m^2. Data have been collected within the ongoing EU FP7 project MOSAIC [86].

4.4 Parsimonious set of indices in T1DM

The application of SPCA to T1DM involved the assessment of the technique on two different CGM datasets, and on the one obtained by merging them. We provide results obtained in the three scenarios, discuss the choice of the two SPCA parameters (i.e., the number of selected PCs and the number of GV indices kept in the regressor), and compare the parsimonious sets of indicators finally extracted by SPCA.
The first T1DM dataset we considered in the study was made up of $n = 17$ CGM profiles. The $m = 25$ GV indices were assessed on the time-series, normalized, and the $n \times m$ data matrix $\mathbf{X}$ was built up. To determine the number of PCs selected after the application of PCA to $\mathbf{X}$, a threshold was set in terms of explained variance; then, the number of selected PCs could be increased by one if the gained explained variance was large enough. In particular, 85% of the variance of the original data was set as minimum value to be preserved, and an additional PC was selected only if the increase of explained variance was at least of 10%. The reason of such a choice, which is somewhat arbitrary, is that 85% of the initial variance is a sufficiently large value to be saved, and a relevant benefit has to derive from the further increase of the number of PCs (since it implies an increased number of finally selected GV indices). As far as dataset 1 is concerned, a plot of the percentage explained variance as a function of the number of selected PCs is reported in the top-left panel of Figure 4.2. It is easy to appreciate from it that, to explain at least 85% of the variance originally present in the data, two PCs were necessary; this number of PCs was also sufficient, since passing from two to three PCs, less than 6% of explained variance was gained. For these reasons, $p = 2$ PCs were finally selected for dataset 1, saving 87% of the variance originally present in the data. The PCA step, thus, enabled to achieve a great

reduction in data dimensionality (from $m = 25$ to $p = 2$) with a minimum loss of information. As described in Section 4.2, however, each PC was obtained as a linear combination of all GV indices. Then, to reduce the number of variables contributing to the regression, the LASSO estimation of sparse loadings was computed for each PC, setting to five the maximum number of nonzero coefficients; the GV indices selected by SPCA were those corresponding to nonzero loadings. The number of nonzero coefficients, and thus of selected GV metrics for each PC, was the second input required to the user. Again, a certain degree of subjectivity is involved in the choice. In the present study, we wanted to have the same number of selected GV indices per PC for both datasets, and we decided to set it as the smallest number of variables that allowed to explain at least 60% of the initial variance in both datasets. In Figure 4.2, top and middle right panels, the percentage explained variance is plotted against the number of selected GV indices per PC, for the previously determined number of selected PCs. As the reader can appreciate from the inspection of top-right (dataset 1) and middle-right (dataset 2) panels, the minimum number of GV metrics that allowed to preserve at least 60% of the original variance in both cases was equal to five. Thus, five was the chosen number of nonzero loadings per PC. Of note is that imposing a maximum number of nonzero coefficients is the same as setting a certain value for the parameter λ. Specifically, there exist a number of λ values that satisfy the requirement of having not more than five loadings different from zero; between them, we chose the smallest one, as it represents the first solution to the optimization problem. For dataset 1, with five (out of 25) GV indices for each of the two (out of 25) PCs, SPCA finally allowed to explain 77% of the variance originally explained by the whole pool of GV indices. The selected metrics are, thus, sufficient for a comprehensive characterization of GV in the analyzed population of 17 T1DM subjects. Results obtained from dataset 1 are summarized in Table 4.1 (top panel), where the indices selected for each PC are listed. Moving on to the second dataset, the data matrix $\mathbf{X}$ was made up of $n = 16$ rows and again $m = 25$ columns, and the choice of SPCA parameters was developed as described for dataset 1. Again, two PCs turned out to be sufficient to explain most of the variance of the initial data, with 86% of the original variance saved, and the LASSO constraint allowed to select five metrics for each of the two PCs. With five indices for each PC, SPCA finally explained 61% of the variance originally conveyed by the 25 considered GV indices. Results obtained from this dataset are summarized in the middle panel of Table 4.1, and plots of the percentage explained variance as a function of SPCA parameters are shown in the middle panels of Figure 4.2, below those referred to dataset 1.

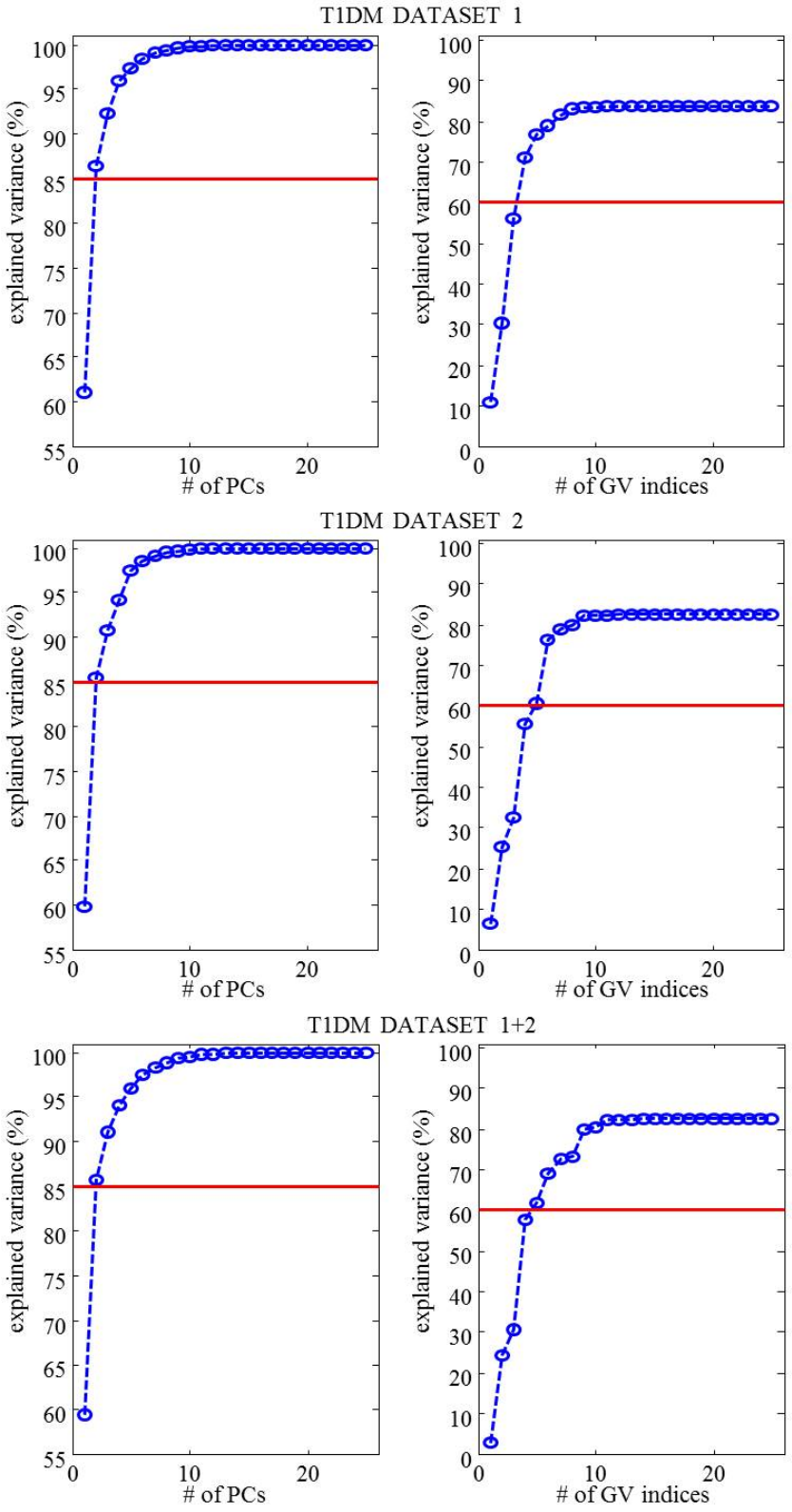

Figure 4.2 Plots of the explained variance (%) vs the number of selected PCs (left panels) and, for the chosen number of PCs (i.e., 2), vs the number of selected GV metrics (right panels) in T1DM. Top panels refer to dataset 1, middle panels to dataset 2, and bottom panels to dataset 1+2. Thresholds fixed by the user are also shown as red horizontal lines at 85% for left panels and 60% for right ones.

T1DM dataset 1 (77%)	
First PC	**Second PC**
J-index	%CV
MAGE	Range
ADRR	LBGI
IGC	ADRR
$\%GRADE_{eu}$	Hypo Index
T1DM dataset 2 (61%)	
First PC	**Second PC**
SD	%CV
BGRI	Median
IGC	%BG below target
$\%GRADE_{eu}$	%BG above target
–	LBGI
T1DM dataset 1+2 (62%)	
First PC	**Second PC**
J-index	%CV
ADRR	Median
BGRI	%BG above target
IGC	LBGI
$\%GRADE_{eu}$	$\%GRADE_{hypo}$

Table 4.1 GV indices selected by SPCA for each PC for the T1DM datasets (1 - top, 2 - middle, and 1+2 - bottom).

Results obtained so far allow us to make some observations and remarks.

Focusing on the first step of the algorithm, the application of PCA to GV indices showed that two out of 25 PCs are sufficient to preserve more than 85% of the total variance originally present in the data. This means that, if analyzed in terms of PCs, GV is mainly a two-dimensional phenomenon. Moreover, the LASSO step confirmed that a subset of up to five metrics extracted for each of the two PCs was enough to save a still large percentage of the original variance (77% in dataset 1 and 61% in dataset 2), suggesting that the two dimensions that well describe GV can be effectively characterized by five predictors each, instead of 25.

Regarding specifically the obtained subsets of metrics in the two datasets, we can observe that most of the selected GV indices result to be measures related to the

dispersion of glucose distribution (e.g., SD, %CV, and J-index), the overall quality of glycemic control (e.g., ADRR, BGRI, and IGC), and the hypoglycemic state (e.g., LBGI, Hypo Index, %BG below target, and $\%GRADE_{hypo}$), and it may happen that two indicators well correlated are selected within the same PC. This is the case, e.g., of LBGI and Hypo Index in dataset 1, and BGRI and IGC in dataset 2. To prove that, even if correlated, both indicators selected by SPCA in the two scenarios are effectively necessary to explain the final percentage of the original variance, we removed LBGI from the selected indices in dataset 1 while maintaining Hypo Index, and removed BGRI in dataset 2, while keeping IGC. In the first case, the percentage explained variance decreased from 77% to 64%; in the second case, it decreased from 61% to 28%. This confirms that even if two indices are well correlated and seem to measure almost the same entity, both of them may be necessary to explain a large percentage of the variance of the original data.

In addition to identifying the parsimonious set of GV indices for each dataset, it is worthwhile investigating how many and which indices are selected in all cases, thus seeming able to well characterize GV in T1DM, independently from the specific dataset. In particular, referring the reader to Table 4.1 where all the obtained results are summarized, we can observe that, for the first PC, two selected GV indices are shared by both datasets 1 and 2. These metrics are IGC and $\%GRADE_{eu}$. On the other hand, J-index, MAGE, and ADRR extracted for dataset 1 are replaced by SD and BGRI in dataset 2. As far as the second PC is concerned, again two GV indices are in common to datasets 1 and 2 (i.e., %CV and LBGI), while Range, ADRR, and Hypo Index, selected for dataset 1, are substituted by Median and %BG below/above target in dataset 2. Considering specifically the pool of GV metrics shared by the two datasets, thus, it comprises GV metrics (i.e., IGC, $\%GRADE_{eu}$, %CV, and LBGI) that evaluate the overall quality of glycemic control (IGC), the euglycemic state ($\%GRADE_{eu}$), the dispersion of glucose distribution (%CV), and the hypoglycemic state (LBGI). From our results, this pool of four GV indices appears independent from the specific T1DM dataset and could be regarded as particularly representative of GV in T1DM. However, it is important to stress that there is still no evidence of this result, since more datasets, comprising a larger number of subjects, are needed to investigate the determination of the most representative subset of GV indices in T1DM on more solid grounds.

As said before, SPCA was also applied to the dataset obtained by merging datasets 1 and 2. Dataset 1+2 was made up of 33 CGM time-series, and the application of SPCA to it followed the same strategy used for the previous scenarios, selecting two PCs and saving five GV indices for each of them. As expected, two PCs were able to

explain more than 85% of the initial variance (in particular, 86% was saved) and five nonzero loadings for each PC allowed finally to preserve 62% of the variance originally conveyed by all metrics. The parsimonious set of indices for dataset 1+2 is reported in the bottom panel of Table 4.1, and plots of the explained variance vs SPCA parameters are shown in the bottom panels of Figure 4.2. Looking at the parsimonious set of indicators, it is possible to see that the four indices shared by datasets 1 and 2 were selected for dataset 1+2 as well, as expected from a T1DM dataset.

4.5 Parsimonious set of indices in T2DM

Results described in the previous section demonstrate that SPCA is a valuable tool to provide a parsimonious but still comprehensive description of GV in T1DM, selecting a small subset of up to 10 indicators and saving more than 60% of the original variance conveyed by 25 GV indices. Moreover, our analyses showed that a small group of four metrics (i.e., IGC, %GRADE$_{\text{eu}}$, %CV, and LBGI) is shared by all the considered T1DM datasets, suggesting the existence of a subset of indicators particularly representative of GV in T1DM regardless of the specific considered dataset. Given these findings, we thought that extending the technique to T2DM could be an interesting investigation, to further confirm the capability of SPCA to parsimoniously describe GV and to assess whether any selected metric was shared by the two types of diabetes. As done in T1DM, the $m = 25$ GV indices were evaluated on the $n = 13$ CGM profiles of the T2DM dataset, and SPCA was applied to the data matrix $\mathbf{X}$ built up after the normalization of the indices. The first step of SPCA led again to the selection of two out of 25 PCs, since with two PCs it was possible to go beyond the fixed 85% threshold (see the left panel of Figure 4.3) and save 88% of the original variance. The LASSO step was then performed, and the trend of the percentage explained variance as a function of the number of GV indices selected per PC is shown in the right panel of Figure 4.3. From the plot, it can be seen that with a pool of indices of the same dimension as that selected for the T1DM datasets, i.e., made up of 10 metrics, the variance finally saved for the T2DM dataset was equal to 83% of the total original one. To cross the 60% threshold of explained variance for this specific dataset, on the other hand, an even smaller subset of indices was sufficient. In particular, selecting two GV indices for each of the two PCs, 67% of the original variance was explained, suggesting that T2DM may need a lower number of indicators than T1DM to explain the same percentage of the original variance (to confirm the last speculation, however, other T2DM datasets have to be analyzed). The parsimonious set of GV indices selected by

SPCA for the T2DM dataset is reported in Table 4.2. The table shows the 10-index pool (five indices for each PC), and the smaller group of four metrics, that is a subset of the former, is made up of the indicators reported as underlined.

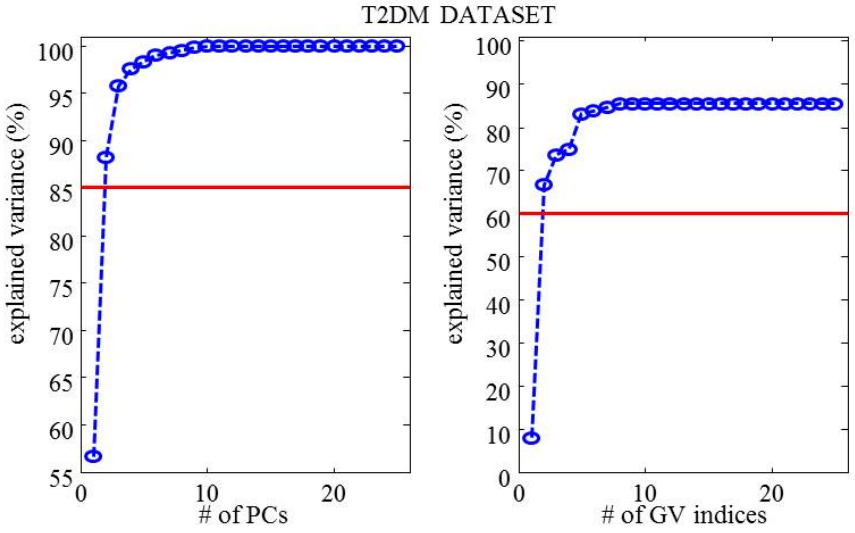

Figure 4.3 Plots of the explained variance (%) vs the number of selected PCs (left panels) and, for the chosen number of PCs (i.e., 2), vs the number of selected GV metrics (right panels) in T2DM. Thresholds fixed by the user are also shown as red horizontal lines at 85% for left panels and 60% for right ones.

T2DM dataset (83%)	
First PC	**Second PC**
J-index	%CV
ADRR	Median
BGRI	%BG below target
$\%GRADE_{eu}$	LBGI
$\%GRADE_{hyper}$	Hypo Index

Table 4.2 GV indices selected by SPCA for each PC for the T2DM dataset.

What is of course possible to state from these results is that SPCA confirms its capability to parsimoniously describe GV in diabetes. Then, we can see that the subset of four indicators shared by all the T1DM datasets is largely shared by the T2DM dataset too. Three out of four indicators (i.e., $\%GRADE_{eu}$, %CV, and LBGI), in fact, belong to the T2DM parsimonious set of indices, and this finding could suggest that these metrics are particularly informative of GV in diabetes, regardless of the specific

type of disease. As said before, however, this result needs to be confirmed by a larger number of datasets.

4.6 Conclusions

GV is believed to be a risk factor for the development of long-term complications from diabetes, and tens of different indicators for its quantification from SMBG and CGM profiles have been proposed in the literature. The issue with the quantification of GV, however, is that a gold-standard approach for its evaluation has not been identified yet, and, even if a combination of indices is probably necessary to well describe GV [87, 88], some indicators can provide redundant information and can be of relatively minor added value in the characterization of GV within a diabetic population.
To identify a reduced subset of indicators that can provide a parsimonious but still comprehensive description of GV in diabetes, we proposed the use of a technique known in the literature as SPCA. If applied to a pool of GV indices evaluated on a dataset of glycemic profiles, SPCA allows to determine a subset of indicators able to explain a large part of the variance originally revealed by the whole pool of metrics. In other terms, SPCA allows the identification of a combination of a reduced number of indices that is still effective for describing the variance of a larger pool of GV indicators within the considered group of subjects.
The use of SPCA was assessed on both T1DM [81] and T2DM [82] CGM datasets, and the obtained results confirmed its usability as a tool to parsimoniously describe GV. Also, our findings showed that this technique can be exploited to determine a small subset of indicators particularly representative of GV in the overall diabetes, given that some indicators emerged to be selected in all scenarios, regardless of the specific dataset or type of disease.
To use SPCA as a tool to identify the gold-standard combination of indicators to quantify GV, however, a larger number of CGM datasets need to be considered, so that the whole diabetic population could be represented and the final parsimonious set of GV metrics can be identified. In addition, we need also to observe that the number of available CGM time-series determines the ratio between number of observations (n) and number of variables (m) used to build the data matrix $\mathbf{X}$. An insufficiently high ratio can result in high dimension low sample size problems, with potential inconsistency of PCA in estimating the covariance matrix of the GV indices and determining the PCs [89–91]. Unfortunately, when the study was developed, only the data documented in this chapter were available for testing the SPCA methodology. Further development of

the work, thus, should start from the extension of the datasets. With larger datasets, in high sample size contexts, then, we could also consider to extend the initial pool of indicators given as input to SPCA by inserting other well established GV metrics, as, e.g., the mean of daily differences [92] and the continuous overlapping net glycemic action [93] indices, as well as new indicators, e.g., those related to the concept of dynamic risk [94] or those derived in the nonlinear time-series analysis context [95–97]. As far as the SPCA algorithm is concerned, future works will consider the use of other regression techniques to estimate the vector of sparse coefficients after the PCA step; the elastic net approach [80], e.g., could be exploited instead of the LASSO constraint.

Chapter 5

Classification of glycemic control and metabolic condition

5.1 Aim of the investigation

In Chapter 4, we used SPCA to select a small number of indicators able to well characterize GV in diabetes. In the present chapter, we investigate if the linear combinations obtained after PCA (i.e., PCs) and after LASSO (i.e., sparse PCs) can be exploited in two classification problems, to discriminate between subjects with different quality of glycemic control (Application 1 - Section 5.5) and between subjects with different metabolic conditions (Application 2 - Section 5.6). In the first investigation, in particular, we had 55 T1DM subjects divided into three groups depending on their glycemic control as assessed by an expert clinician from visual inspection of CGM profiles; in the second investigation, we exploited CGM time-series acquired from 34 normal subjects at high risk of developing T2DM, 39 IGT, and 29 subjects with diagnosed T2DM. Both applications were managed by using linear and nonlinear support vector machine (SVM) classifiers, an approach widely employed in machine learning [98, 99] and summarily described in Section 5.2, which was recently used in the diabetes technology area for improving hypoglycemia detection from CGM [100–102] and detecting incorrect measurements of CGM systems [103, 104]. To build the classifiers, we considered PCs, sparse PCs, and single GV metrics as possible feature configurations. We performed the investigation using both PCs and sparse PCs to assess if sparse PCs, that allow to describe GV in a parsimonious fashion through a small number of indicators, could provide an accuracy of classification comparable to that obtained with full PCs. Moreover, we evaluated the performance obtained by considering single GV metrics

to test the hypothesis that a combination of indicators was effectively necessary to perform well in the classification problems.

5.2 SVM classifiers

This section is devoted to an introductive presentation of SVMs, that provides a qualitative description of their design and use. A more detailed discussion can be found in textbooks [98, 99], to which we refer the reader interested in this topic.

5.2.1 Formulation of the classification problem

Before entering the description of SVMs, we report some insights from statistical learning theory that help us in setting the problem of pattern recognition or classification. In general, who resorts to learning algorithms is facing situations where two classes of objects are given and a new object needs to be assigned to one of the two classes. This means that we have a group of n data

$$(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n) \in \mathcal{H} \times \{\pm 1\} \tag{5.1}$$

with $\mathbf{x}_i$ observations or patterns represented as vectors in the feature space $\mathcal{H}$ and y_i labels of the observations identifying the class to which each observation belongs, and we want to generalize the unseen data points through a certain decision function

$$f : \mathcal{H} \to \{\pm 1\} \tag{5.2}$$

that predicts the class $y \in \{\pm 1\}$ of all the new unclassified patterns $\mathbf{x} \in \mathcal{H}$. The mathematical formulation of the decision function depends on the specific method adopted in the analysis, and the unknown parameters in it are optimized on a set of pattern-label points called training set. After determining the optimal values of the parameters in f, a different dataset, called test set, is used to asses the classification performance, expressed, e.g., in terms of percentage of correctly classified points (classification accuracy).

5.2.2 Optimal margin hyperplanes

In this section, we will show how to obtain the optimal decision function in the case of linearly separable two-class training set. If we are given a feature space $\mathcal{H}$ (in which

the dot product is defined), and a set of pattern vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathcal{H}$ belonging to two classes labeled by $+1$ and -1, any hyperplane in $\mathcal{H}$ separating our data can be written as

$$\{\mathbf{x} \in \mathcal{H} \mid \langle \mathbf{w}, \mathbf{x} \rangle + b = 0\}, \quad \mathbf{w} \in \mathcal{H}, b \in \mathbb{R} \tag{5.3}$$

where $\mathbf{w}$ is a vector orthogonal to the hyperplane, commonly referred to as the weight vector in the machine learning literature. Our aim is finding the best hyperplane $(\mathbf{w}, b)$ that separates the observations in the training set belonging to different classes, based on a certain optimization criterion. To determine the optimal hyperplane, a decision function $f_{\mathbf{w},b} : \mathcal{H} \to \{\pm 1\}$ needs to be defined, and the pair $(\mathbf{w}, b)$ in it is determined so that the labeled training examples are correctly classified and the margin of the hyperplane is maximized. For a point $(\mathbf{x}, y) \in \mathcal{H} \times \{\pm 1\}$, the margin is defined as

$$\rho_{(\mathbf{w},b)}(\mathbf{x}, y) = y \left(\langle \mathbf{w}, \mathbf{x} \rangle + b \right) / ||\mathbf{w}|| \tag{5.4}$$

and represents the distance from the point to the the separating hyperplane, with a positive value if the point is correctly classified and a negative value otherwise. When the entire training set is concerned, the margin of the hyperplane is the margin of the closest observation correctly classified, and, as can be seen from Equation 5.4, it depends on the length of the weight vector $\mathbf{w}$. Loosely speaking, if we can separate the data in our training set with a large margin (i.e., if the vector patterns are far enough from the hyperplane), then we achieve a high generalization ability and have reason to believe that we will do well on the test set. To confirm this idea, a number of explanations exist [98]; an example is the insensitivity to pattern noise represented in Figure 5.1, where it can be seen that if the optimal hyperplane has a margin ρ and some noise bounded by $r < \rho$ is added to each pattern, then the hyperplane will correctly separate even the noisy patterns.

The hyperplane decision functions, thus, should be constructed such that they maximize the margin, and at the same time separate the training data with as few exceptions as possible. Given these requirements, the optimization problem can be formalized.

We can assume that we are managing nonempty classes and that the point closest to the hyperplane has a distance equal to $1/||\mathbf{w}||$, i.e., the margin of the hyperplane is $1/||\mathbf{w}||$. The condition of classifying correctly the training points implies that we have to find a map of the form $f_{\mathbf{w},b} : \mathcal{H} \to \{\pm 1\}$ satisfying

$$f_{\mathbf{w},b}(\mathbf{x}_i) = y_i \tag{5.5}$$

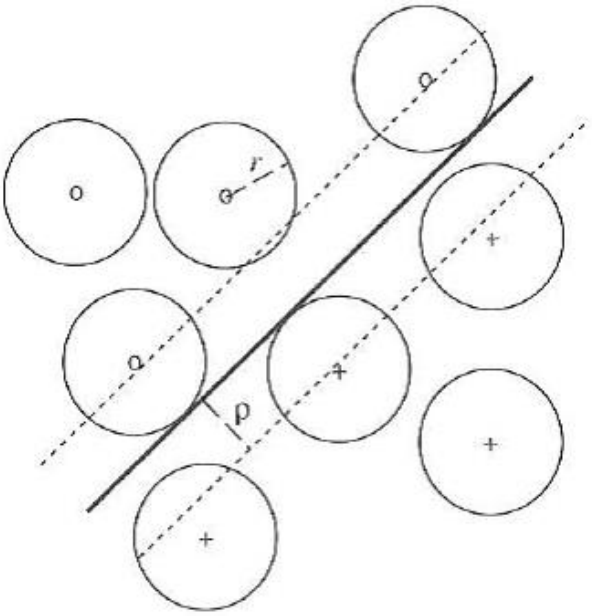

Figure 5.1 Two-dimensional toy example of a classification problem [98].

which amounts to saying that the output of the decision function evaluated on each pattern $\mathbf{x}_i$ equals the real class to which the pattern belongs. If such a function exists, then each pattern $\mathbf{x}_i$ is correctly classified and its distance from the hyperplane is a positive value equal to or greater than $1/||\mathbf{w}||$. This means that we can write the following constraints

$$y_i(\langle \mathbf{x}_i, \mathbf{w} \rangle + b) \geq 1, \quad i = 1, \ldots, n \tag{5.6}$$

and use them in the formulation of the optimization problem to summarize the requirement of having the training observations correctly classified. Beyond this, we need to say that we want to choose the separating hyperplane that has the greatest margin. The objective function of our optimization problem, thus, should maximize the quantity $2/||\mathbf{w}||$, that equals twice the margin of the hyperplane because we are considering its symmetric version around the hyperplane. Maximizing $2/||\mathbf{w}||$ is the same as minimizing $||\mathbf{w}||/2$, and this leads to obtaining the optimal separating hyperplane as the solution of the following constrained quadratic optimization problem

$$\underset{\mathbf{w}\in\mathcal{H}, b\in\mathbb{R}}{\text{minimize}} \quad \tau(\mathbf{w}) = \frac{1}{2}||\mathbf{w}||^2 \tag{5.7}$$

subject to the constraints 5.6. This is called the primal optimization problem, and a graphic representation of its implementation to determine the optimal margin hyperplane is shown in Figure 5.2. Points represented in red, that lie exactly on the margin of the hyperplane, are known as support vectors and are those where the constraints

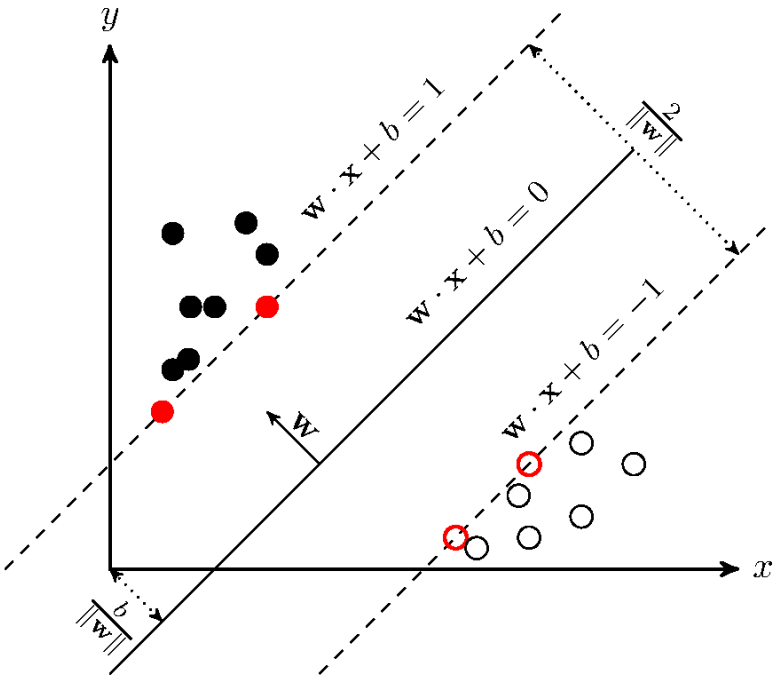

Figure 5.2 Optimal margin hyperplane classification for linearly separable two-class training set. Support vectors are shown as red points.

5.6 are precisely met. Getting the solution of the primal optimization problem is possible, but not always easy. In the context of SVMs, the so-called dual problem is usually derived, which is a problem that can be shown to provide the same solutions as the primal problem, but turns out to be more convenient to deal with (we address the interested reader to the textbook [98]).

5.2.3 Soft margin hyperplanes

In some scenarios, we would rather have an algorithm which can tolerate a certain fraction of classification errors in separating the training data, instead of necessarily looking for an optimal margin hyperplane. The approach that allows to implement this kind of separation is based on the concept of soft margin hyperplane, and specifically introduces some wights to the training data, called slack variables and defined as

$$\xi_i \geq 0, \quad i = 1, \ldots, n, \tag{5.8}$$

to relax the constraints 5.6 and get the following conditions

$$y_i(\langle \mathbf{x}_i, \mathbf{w} \rangle + b) \geq 1 - \xi_i, \quad i = 1, \ldots, n. \tag{5.9}$$

Clearly, by making ξ_i large enough, the constraint on $(\mathbf{x}_i, y_i)$ can always be met, and thus to avoid the trivial solution where all ξ_i take on large values, we need to penalize them in the objective function. To this end, the term $\sum_i \xi_i$ is included in 5.7, and the soft margin hyperplane is obtained by solving for some $C > 0$ the following optimization problem

$$\underset{\mathbf{w}\in\mathcal{H},\boldsymbol{\xi}\in\mathbb{R}^n}{\text{minimize}} \quad \tau(\mathbf{w},\boldsymbol{\xi}) = \frac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{n}\xi_i \tag{5.10}$$

subject to the constraints 5.8 and 5.9. Whenever the constraint 5.9 is met with $\xi_i = 0$, the corresponding point will not be a margin error. All nonzero slacks ξ, on the other hand, correspond to margin errors, and the fraction of them that is accepted increases with the second term of 5.10 and is managed through the positive constant C. As in the separable case, the optimization problem can be shown to have a dual formulation [98] that can be easily exploited to get the solution.

5.2.4 Nonlinear support vector classifiers and kernel trick

In the previous sections, we have only dealt with linear separators, and the case of more general nonlinear decision surfaces has not been treated yet. Here, we discuss the use of nonlinear support vector classifiers. In doing that, we observe that a set of training data that is not linearly separable, if projected onto a higher-dimensional space via some nonlinear transformation, has a high probability of being transformed into a linearly separable set [98]. We can, thus, construct a nonlinear map $\boldsymbol{\phi}(\mathbf{x}_i)$ to transform the input data $\mathbf{x}_i \in \mathcal{H}$ into vectors of a higher-dimensional feature space, and then separate the new data $\boldsymbol{\phi}(\mathbf{x}_i)$ using linear decision surfaces obtained by solving the optimization problem 5.10. Even if we can resort to problems of the same form as those presented in the previous sections, however, this new formulation of the optimization function requires the map $\boldsymbol{\phi}$ to be defined and computed on each $\mathbf{x}_i$, and suffers from the drawback that the calculation of dot products in high-dimensional spaces is a very expensive procedure from a computational viewpoint. To overcome these limitations, the so-called kernel trick to decrease the complexity of the optimization problem is proposed. A kernel can be thought of as a symmetric function $k : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ that is expressed as

$$k(\mathbf{x}_i, \mathbf{x}_j) = \langle \boldsymbol{\phi}(\mathbf{x}_i), \boldsymbol{\phi}(\mathbf{x}_j) \rangle \,, \tag{5.11}$$

and corresponds to the dot product in a high-dimensional feature space via the map $\boldsymbol{\phi}$. The most commonly exploited kernel functions are

$$\langle \mathbf{x}_i, \mathbf{x}_j \rangle \quad \text{(linear kernel)} \tag{5.12}$$

$$\langle \mathbf{x}_i, \mathbf{x}_j \rangle^d \quad \text{(polynomial kernel)} \tag{5.13}$$

$$e^{-\gamma ||\mathbf{x}_i - \mathbf{x}_j||^2} \quad \text{(radial basis function)} \tag{5.14}$$

but beyond their specific formulations, what makes them crucial in our treatment is the possibility of using the relation 5.11 to rewrite the dot products among high-dimensional vectors $\boldsymbol{\phi}(\mathbf{x}_i)$ and simplify the optimization problem we are dealing with.

5.3 Datasets

In Application 1 (classification of the quality of glycemic control), we used three datasets of CGM time-series collected from T1DM subjects. Two of the three are those already used in the parsimonious description of GV, and thus we refer the reader to Chapter 4 for details. The third dataset, made up of 22 profiles, was collected at the Institute for Clinical and Experimental Medicine (Prague, Czech Republic) within the same project and using the same protocol as the previous two. To perform the classification problem, an expert clinician divided the subjects into three classes based on the quality of their glycemic control as assessed from visual inspection of CGM profiles. Figure 5.3 shows a representative profile extracted from each class; comparing the top (good control), middle (medium control), and bottom (bad control) panels, a difference in the glycemic control of the three subjects can be already appreciated. To help in identifying hypo and hyperglycemic events, thresholds at 70 and 180 mg/dL are shown in the figure as dotted black lines; moreover, thresholds at 50 mg/dL represented as dotted red lines are reported to highlight the occurrence of SH. As can be seen, the subject with good control experiences few hypo events, no SH episodes, and some moderate hyper excursions. For subjects with medium and bad control, on the other hand, some SH events can be detectable from the CGM profiles, and, in the latter case, a sustained hyperglycemia can be seen as well.

In Application 2 (classification of the metabolic condition), 34 normal subjects at high risk of developing T2DM, 39 IGT, and 29 T2DM subjects monitored using the Medtronic® iPro® CGM System [85] were considered. As those used in the parsimonious description of GV in T2DM, these data have been collected within the ongoing EU FP7 project MOSAIC [86]. Examples of representative CGM recordings collected

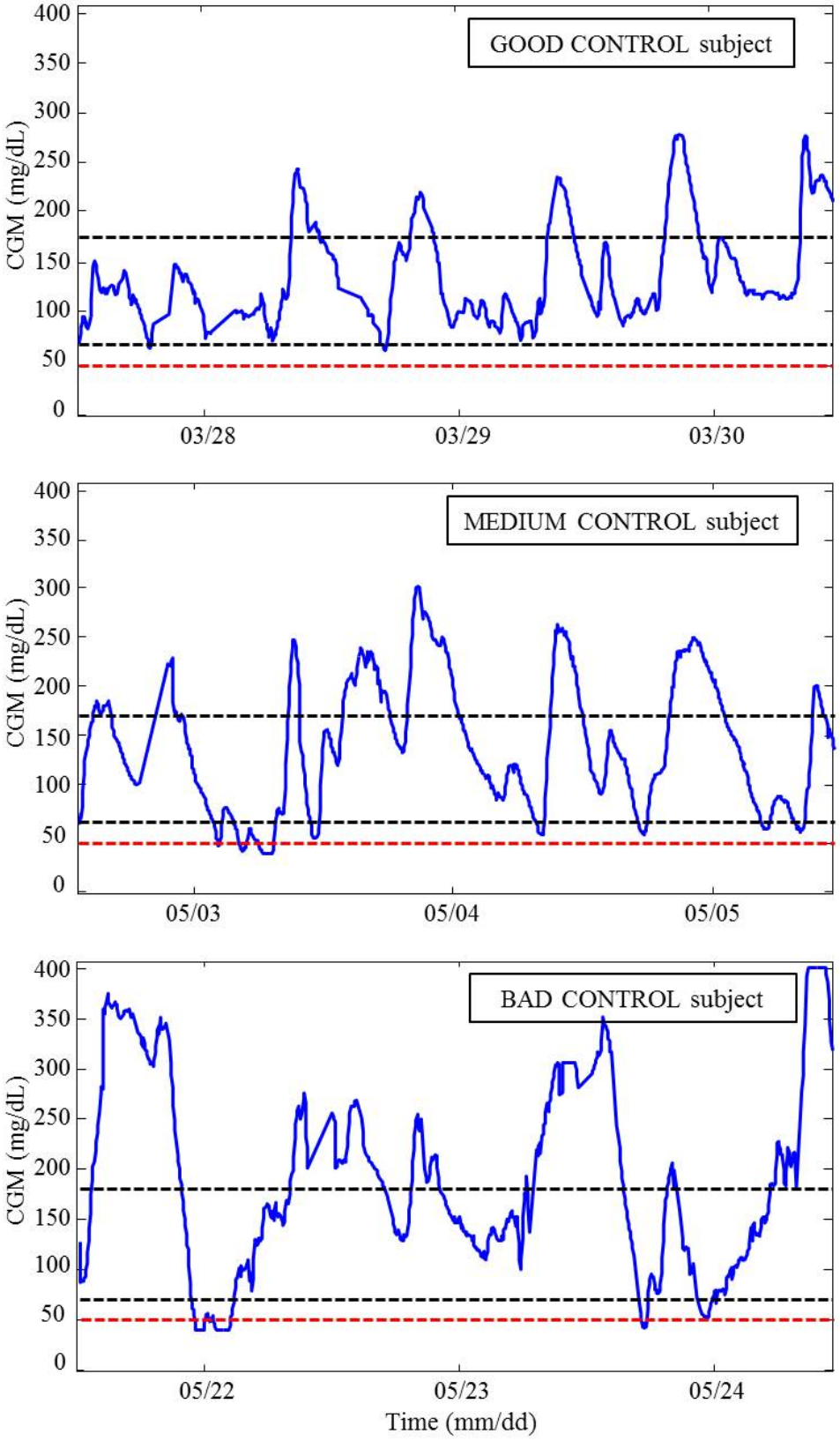

Figure 5.3 Representative CGM time-series extracted from each class of glycemic control according to the expert clinician classification; dotted black lines in the panels represent hypo and hyperglycemic thresholds at 70 and 180 mg/dL, respectively; dotted red lines represent the SH threshold at 50 mg/dL.

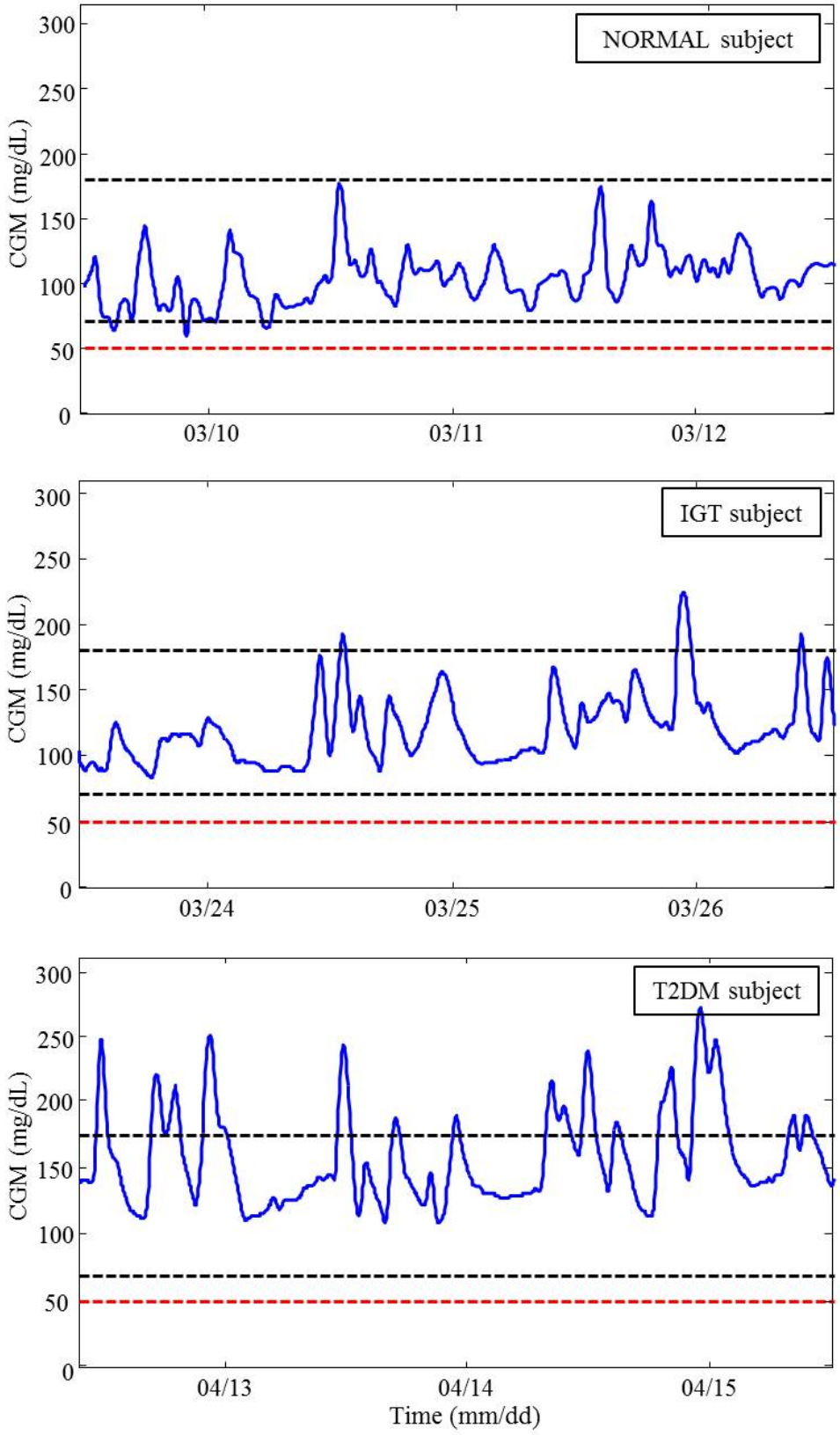

Figure 5.4 Representative CGM time-series extracted from normal, IGT, and T2DM classes; dotted black lines in the panels represent hypo and hyperglycemic thresholds at 70 and 180 mg/dL, respectively; dotted red lines represent the SH threshold at 50 mg/dL.

in a normal, IGT, and T2DM subject are shown in Figure 5.4, top, middle, and bottom panel, respectively. As for Application 1, the visual inspection of CGM profiles allows to capture significant differences in the GV of the three subjects, highlighting a sustained hyperglycemia experienced in T2DM, which decreases in the case of IGT, and is almost absent in the normal profile.

5.4 Design of the classification study

This section provides details about the features used to characterize the observations in our problems, the determination of training and test sets, and the estimation of some relevant coefficients in the design of linear and nonlinear classifiers.

5.4.1 Feature configurations

The starting point used to derive our observations was the application of SPCA to the dataset under analysis. This allowed us to manage two different feature configurations, where PCs and sparse PCs were subsequently tested as features to build the classifiers. Specifically, given that for all scenarios reported in Chapter 4 we saw that two PCs were enough to explain most of the variance of the original 25 GV metrics, the first two PCs provided by the application of PCA to the data matrix $\mathbf{X}$ conveniently built up were considered as first feature configuration. Then, to assess if the parsimonious combinations of indicators obtained from the application of the LASSO constraint to each selected PC could provide comparable classification performances, the two sparse PCs were considered as features too, defining the second analyzed configuration. Beyond the outputs obtained from SPCA, we decided to consider also each GV index as a single feature to perform the classification, to investigate if a combination of indicators was effectively necessary to comprehensively characterize GV and classify quality of glycemic control and metabolic condition, or if single indicators by themselves could already perform well in doing that.

5.4.2 Training and test sets

Once defined the feature configuration, we had to divide the whole observation set into training and test sets. To do that, we followed the strategy reported in Figure 5.5. We considered 80% of the available observations randomly extracted from the entire dataset to build the training set, and the remaining 20% to build the test set. To avoid

any bias due to the relatively small number of testing observations, the procedure was repeated 100 times, and the median classification performance was then computed.

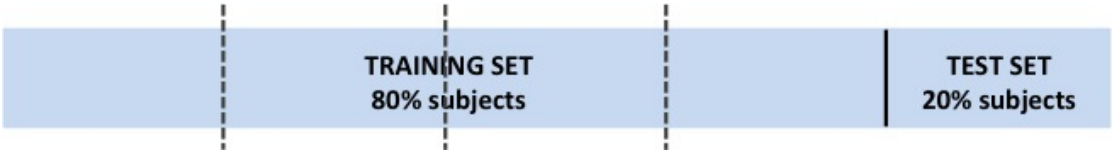

Figure 5.5 Definition of training and test sets in our SVM-based classification problems. Dotted lines that divide the training set into four subgroups indicate that a k-fold cross-validation was performed to determine the best SVM configuration.

Given that the implementation of SVM classifiers always depends on the regularization parameter C and on some other parameters that appear in certain kernel functions (e.g., d in the polynomial kernel or γ in the radial basis functions), an assessment of the more convenient values to use has to be made each time on the specific observation set. To this aim, in our implementation, we performed a k-fold cross-validation on the training set (see Figure 5.5) and used a grid search approach to determine the best values of our parameters. Based on this, we divided the training set into $k = 4$ subgroups of observations made up of approximately the same number of items, trained the classifier with a specific configuration of parameters on $k - 1$ groups, and tested it on the remaining one. By repeating the procedure for all the k subgroups, we could get a median classification performance for each specific combination of parameters. The best configuration was finally determined by keeping the parameterization that maximized the median classification accuracy in the cross-validation setting.

5.5 Application 1: quality of glycemic control

5.5.1 Statistical analysis

We start the discussion on the obtained results by reporting in Figure 5.6 a graphic representation of the set of observations (subjects) as depicted in the two two-dimensional feature spaces that we considered in our investigation. In the figure, we can see the placement of our samples in the PC (top panel) and sparse PC (bottom panel) feature space, where subjects are represented as blue circles if their glycemic control is good, as green triangles if their control is bad, and as red crosses if their control is in between. In both feature spaces, each vector is specified by a pair (x, y) of components that correspond to the first and second PC/sparse PC, respectively. From the inspection of the samples distribution, we can already observe that good and medium

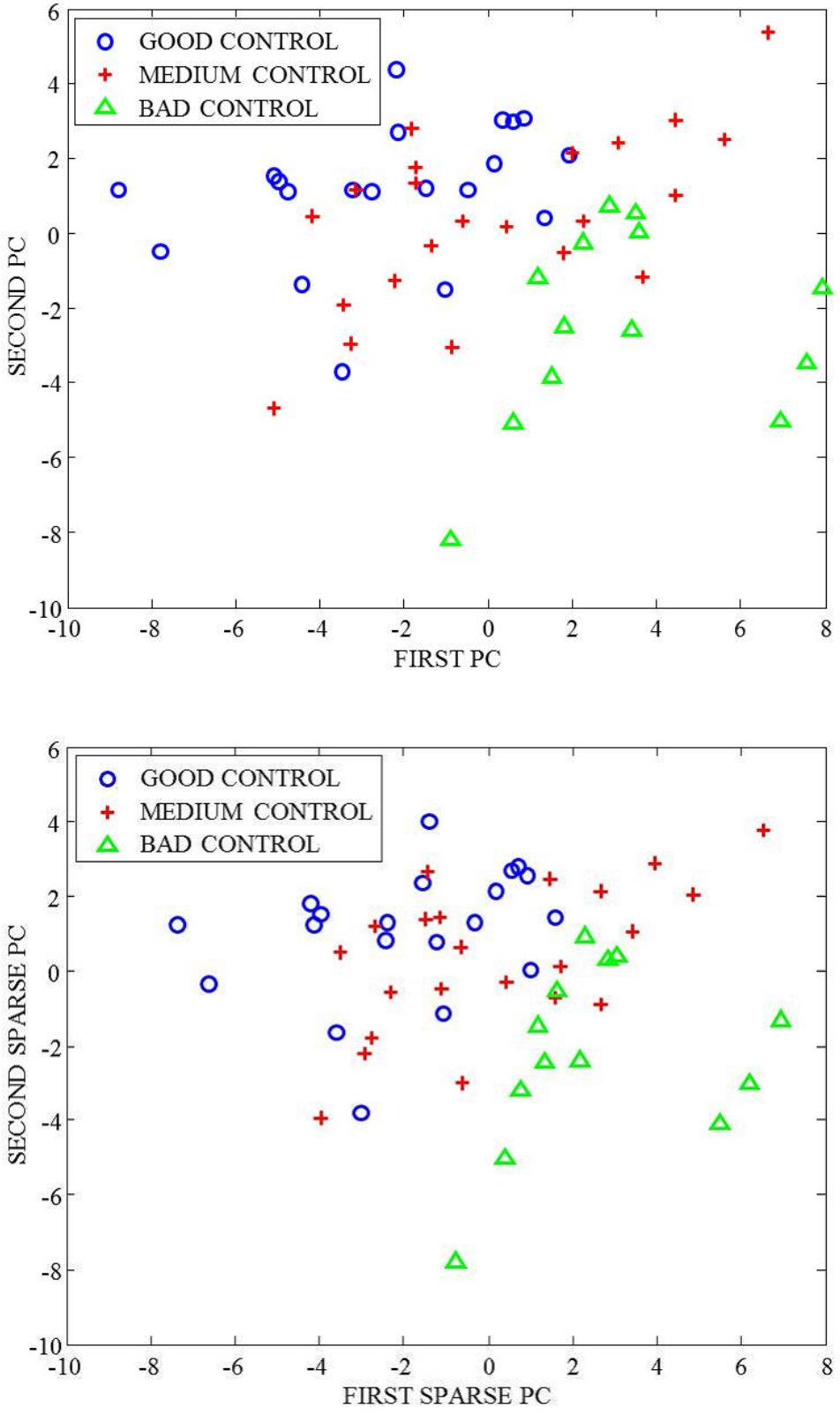

Figure 5.6 Representation of the observation set in the PC feature space (top panel) and sparse PC feature space (bottom panel).

control classes overlap significantly, and a separating surface which can generalize well will be probably difficult to find in this case. Good/medium vs bad control classes, on the other hand, seem to be separable quite enough. Moreover, at first sight, the distribution of the observations in the sparse PC feature space seems to be not so different from the distribution in the PC space, but to confirm the speculation that the two scenarios behave similarly, we need to quantitatively compare the classification performance achieved in the two cases.

Before entering the classification problem, we wanted to assess our choice of the feature settings and be sure that a statistically significant difference between the three classes was obtained for the considered features. To this aim, we computed a nonparametric Kruskal-Wallis test for PCs and sparse PCs between the three classes. This test was performed instead of one-way ANOVAs because one-sample Kolmogorov-Smirnov tests used to check the null hypothesis of normally distributed samples provided p-values < 0.05. Results form the Kruskal-Wallis tests are shown in Figure 5.7 (top panels referred to PCs and bottom panels referred to sparse PCs). In all cases, p-values obtained from the analyses resulted $\ll 0.005$, meaning that there is at least one class for each feature that behaves differently from the others. This suggested that our features and feature spaces had been conveniently chosen.

5.5.2 Classification performance

Once established that first and second PC/sparse PC were significantly different between well, medium, and badly-controlled subjects, we could move to the actual classification problem. In practice, we dealt with three two-class classification problems (i.e., good vs medium control, good vs bad, and medium vs bad), and for each of them we exploited linear, polynomial, and radial kernel functions.

To determine the optimal values of the SVM parameters, we performed the grid search described in Section 5.4. The best parameterization was chosen as the one that maximized the accuracy of classification in the cross-validation setting, and was then used to train the classifier on the whole training set and determine the optimal decision surface. This was made 100 times, with training and test sets randomly chosen, so that any dependence of the results on the small number of testing points was avoided. Classification performances obtained in the investigation are summarized in Table 5.1. Accuracy of classification achieved from the application of the optimal SVM to both training and test sets are listed, and values are reported as the median over the 100 repetitions performed with different training and test sets.

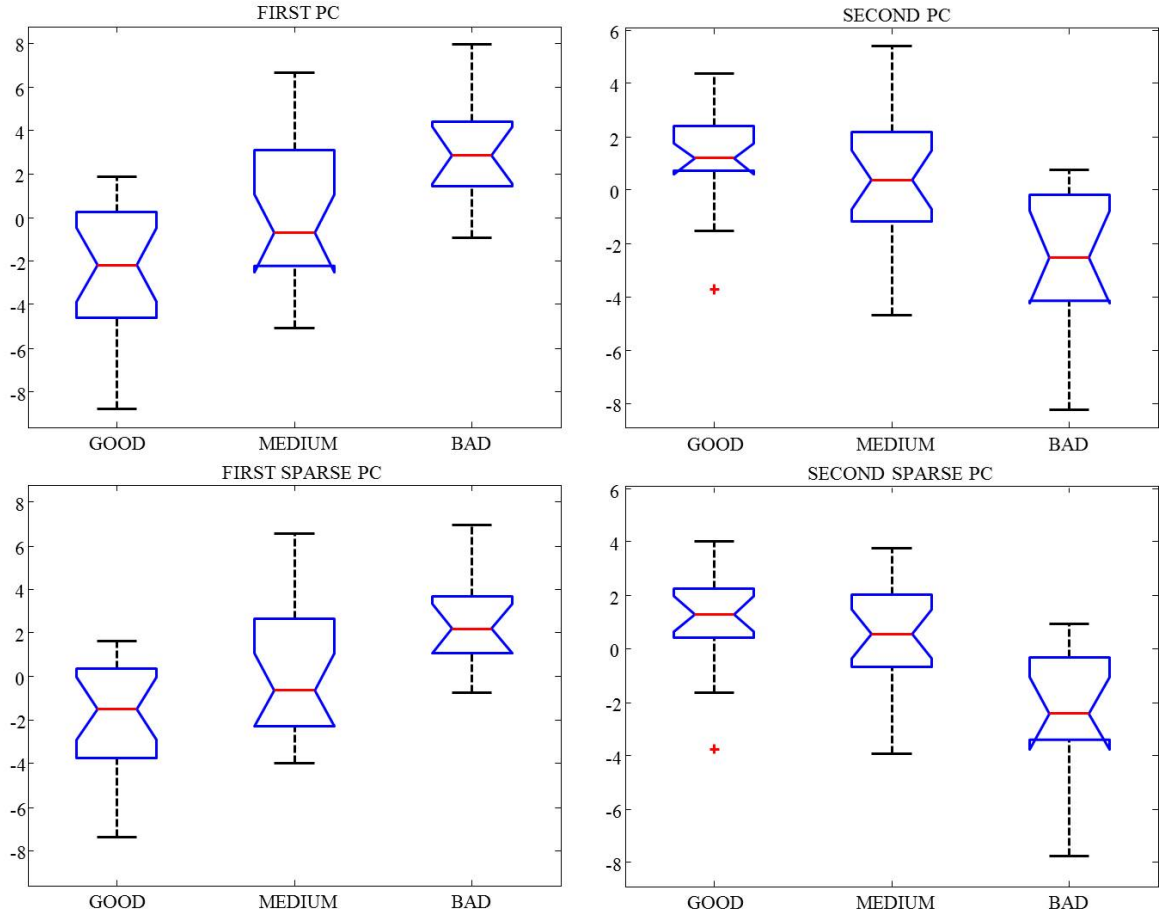

Figure 5.7 Application 1: quality of glycemic control. Boxplots of first and second PCs (top panels) and sparse PCs (bottom panels). On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme datapoints the algorithm considers to be not outliers, and the outliers are plotted individually.

Results obtained in this scenario seem satisfactory. In all cases, the performances we achieved are greater than 67%, and, in the good vs bad problem, we always got a 100% classification accuracy. Looking at the tables, we can observe that in almost all cases (except from the medium vs bad case with the radial kernel) linear kernels performed well enough, and thus we don't need to move to nonlinear configurations, and, above all, the use of sparse PCs instead of PCs led to highly comparable results. Specifically, in the good vs bad problem, performances of PCs and sparse PCs were exactly the same; in the good vs medium, PCs outperformed sparse PCs in two out of three cases; and in the medium vs bad, sparse PCs outperformed PCs in one out of three cases. Also, no GV index as considered alone was able to achieve a classification accuracy greater than 40% (results not documented for sake of space).

	GOOD VS MEDIUM CONTROL					
	PCs			**SPARSE PCs**		
	LIN	**POL**	**RAD**	**LIN**	**POL**	**RAD**
TRAINING	76%	76%	79%	76%	76%	76%
TEST	78%	78%	67%	72%	72%	77%
	GOOD VS BAD CONTROL					
	PCs			**SPARSE PCs**		
	LIN	**POL**	**RAD**	**LIN**	**POL**	**RAD**
TRAINING	96%	100%	96%	96%	96%	96%
TEST	100%	100%	100%	100%	100%	100%
	MEDIUM VS BAD CONTROL					
	PCs			**SPARSE PCs**		
	LIN	**POL**	**RAD**	**LIN**	**POL**	**RAD**
TRAINING	85%	85%	89%	85%	85%	85%
TEST	75%	75%	75%	75%	75%	75%

Table 5.1 Application 1: quality of glycemic control. Classification accuracy obtained in the PC and sparse PC feature space from the three two-class classification problems and the three exploited kernel functions (linear - LIN, polynomial - POL, and radial - RAD).

5.6 Application 2: metabolic condition

5.6.1 Statistical analysis

As done for the first investigation, we start the discussion by reporting a graphic representation of the set of observations as depicted in the PC (Figure 5.8 - top panel) and sparse PC (Figure 5.8 - bottom panel) feature space. In the figures, normal subjects are represented as blue circles, IGT as red crosses, and T2DM as green triangles, and each vector in the feature space is specified by a pair (x, y) of components that correspond to the first and second PC/sparse PC, respectively. Figure 5.8 shows that, even if our classes overlap significantly, there are several subjects that have clearly different positions in the feature spaces, and thus decision hyperplanes might perform well in separating our groups of observations. Also, we can notice again that, at first sight, what we can derive from the distribution in the sparse PC feature space is not so different from what we see in the PC space. But again this speculation needs to be

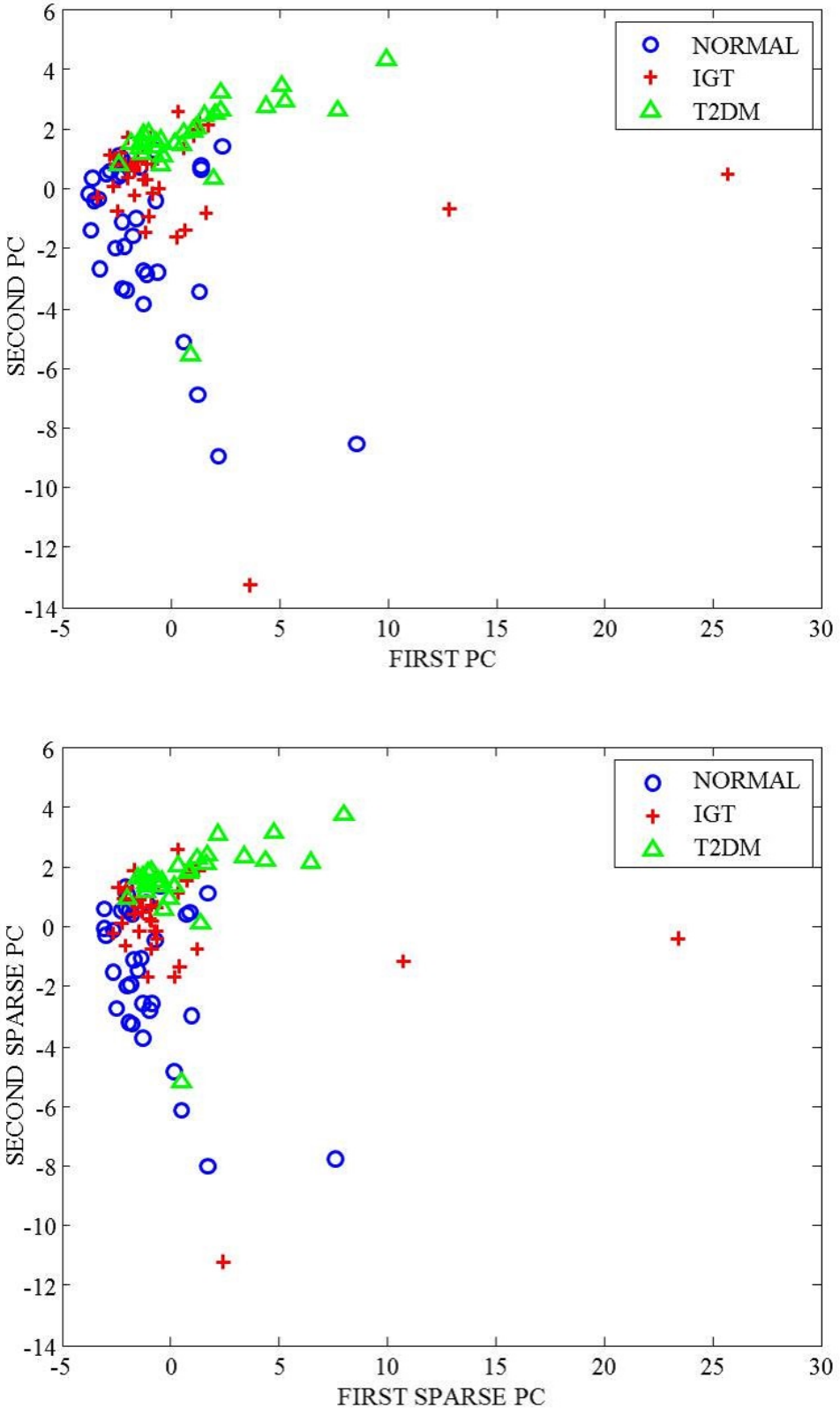

Figure 5.8 Representation of the observation set in the PC feature space (top panel) and sparse PC feature space (bottom panel).

confirmed by the final classification performances.

As for the Application 1, we computed a nonparametric Kruskal-Wallis test for PCs and sparse PCs between the three classes (one-sample Kolmogorov-Smirnov tests used to check the null hypothesis of normally distributed samples provided p-values < 0.05). Results from the analysis are shown in Figure 5.9 (top panels referred to PCs and bottom panels to sparse PCs). In all cases, p-values obtained from the analyses resulted $\ll 0.005$, and this confirmed again that our features had been conveniently chosen.

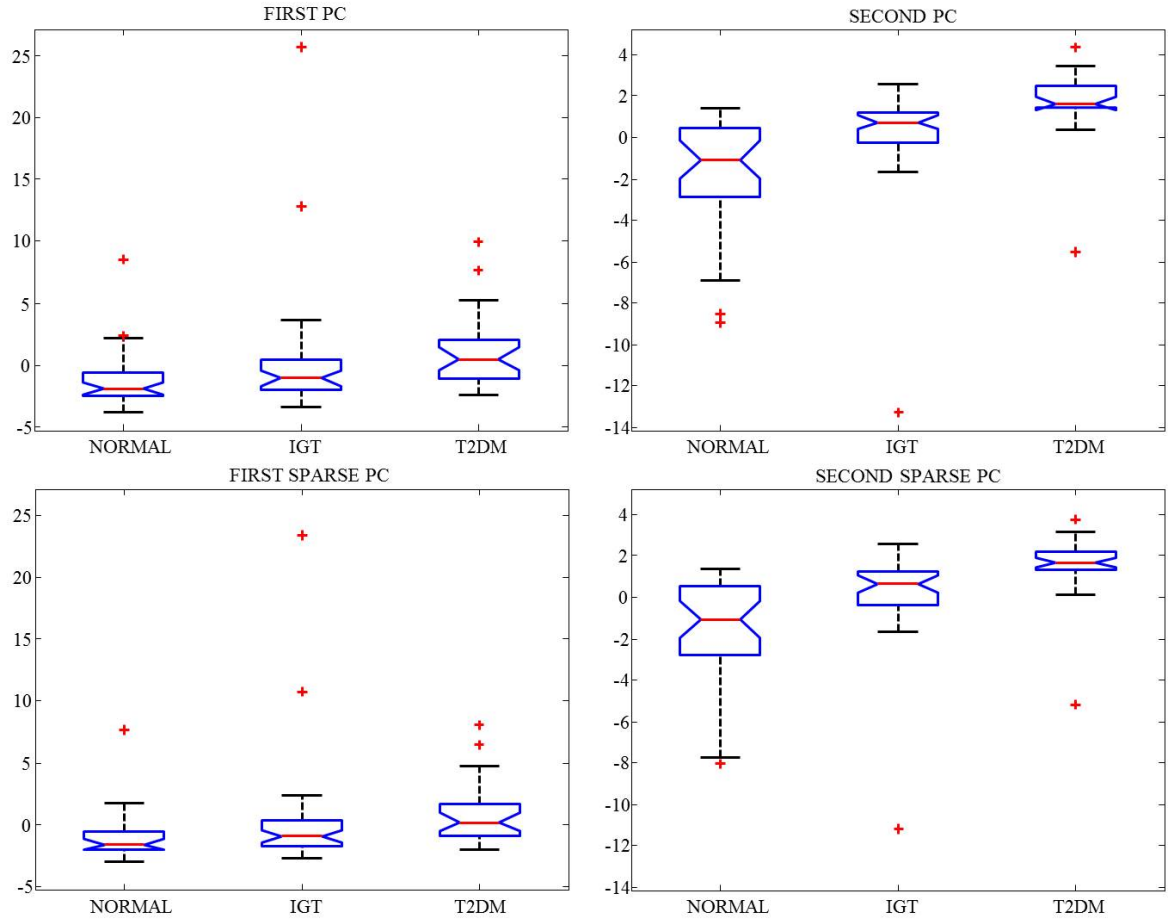

Figure 5.9 Application 2: metabolic condition. Boxplots of first and second PCs (top panels) and sparse PCs (bottom panels). On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme datapoints the algorithm considers to be not outliers, and the outliers are plotted individually.

5.6.2 Classification performance

Since PCs and sparse PCs were significantly different between normal, IGT, and T2DM subjects, we could move to the actual classification problem. In practice, we dealt

with three two-class classification problems (i.e., normal vs IGT, normal vs T2DM, and IGT vs T2DM), and for each of them we exploited linear, polynomial, and radial kernel functions.

Again, we performed the grid search described in Section 5.4, selected the best combination of parameters, and determined the optimal decision surface 100 times to avoid any bias in the obtained results.

Classification performances obtained in the investigation are summarized in Table 5.2. Accuracy of classification obtained on both training and test sets are listed, and values are reported as the median over the 100 repetitions performed with different training and test sets. We can observe that in general we were able to achieve a satisfactory accuracy of classification. The are no performances lower than 71%, and, in some cases, values greater than 90% were obtained. We can also notice that, again, sparse PCs behave similarly to PCs, allowing to achieve the same classification accuracy on the test set in all cases, and that GV index considered alone could not perform comparably to PCs/sparse PCs (results not shown).

	NORMAL VS IGT					
	PCs			**SPARSE PCs**		
	LIN	**POL**	**RAD**	**LIN**	**POL**	**RAD**
TRAINING	76%	78%	81%	74%	76%	79%
TEST	73%	73%	73%	73%	73%	73%
	NORMAL VS T2DM					
	PCs			**SPARSE PCs**		
	LIN	**POL**	**RAD**	**LIN**	**POL**	**RAD**
TRAINING	88%	88%	90%	88%	88%	90%
TEST	92%	85%	85%	85%	92%	85%
	IGT VS T2DM					
	PCs			**SPARSE PCs**		
	LIN	**POL**	**RAD**	**LIN**	**POL**	**RAD**
TRAINING	76%	83%	83%	76%	83%	83%
TEST	71%	71%	71%	71%	71%	71%

Table 5.2 Application 2: metabolic condition. Classification accuracy obtained in the PC and sparse PC feature space from the three two-class classification problems and the three exploited kernel functions (linear - LIN, polynomial - POL, and radial - RAD).

5.7 Conclusions

In this chapter, we exploited the use of the outputs obtained from SPCA (i.e., PCs and sparse PCs) as tools for the classification of the quality of glycemic control (Application 1) and of the metabolic condition (Application 2) of disordered subjects. Specifically, we developed SVM-based linear and nonlinear classifiers, and attempted to solve the two classification problems using the first two PCs and their sparse representations as features to build the classifiers.

Results obtained from this investigation were satisfactory. With the relevant percentages of classification accuracy achieved, we can speculate that PCs and sparse PCs could be exploited, e.g., within glucose sensors, to characterize the glycemic condition of disordered subjects, providing feedback about the ongoing therapy and alerts about the possible transition to more dangerous metabolic diseases (e.g., from a condition at risk of T2DM to T2DM). Moreover, given that equal performances were achieved with PCs and sparse PCs, a confirmation of the real capability of SPCA of parsimoniously describing GV in diabetes was provided.

It is also worth observing that PCs and sparse PCs outperformed any GV index as considered alone in solving our classification problems, and this suggests that a combination of metrics is effectively necessary to well characterize GV. Nonetheless, it could be pointed out that combinations of indicators with a number of predictors smaller than five could work as well. This kind of investigation could be matter of future works, but it is important to stress that the approach to get the combinations of metrics will be exactly the same, exploiting the SPCA-based strategy presented here.

A limitation of the study is related to the relatively small number of subjects that could be used in the test sets. This condition led us to repeat the analysis 100 times, without being able to provide a final formulation of the SVM classifier. With larger datasets, it could be possible to isolate a certain amount of subjects within a test set, and identifying a definitive version of the SVM that could be formalized and exploited in practice to solve classification problems like those proposed here. This issue will be addressed in future works.

From an algorithmic viewpoint, developments of this investigation shall assess the implementation of classification strategies different from SVMs and move from two-class classification problems to multi-class ones, exploiting the common one-vs-one and one-vs-rest SVM approaches or other techniques.

Finally, to strengthen the results obtained in Application 1 and make them more consistent, a larger number of expert clinicians should be asked to perform the initial

classification of the quality of glycemic control. This would render the study less dependent on potential subjective evaluations.

Chapter 6

Conclusions

6.1 Summary of the results

GV is widely considered as a risk factor for the development of long-term complications from diabetes, and a number of indicators have been developed to quantitatively measure it from either SMBG or CGM profiles.

The first problem considered in this dissertation was extending the use of some GV indices designed and validated only from SMBG data, to CGM recordings. In particular, we performed this kind of analysis on the well-known LBGI and HBGI, indicators that, as computed from SMBG time-series, have been shown to be predictive of significant glycemic events and allow to classify the risk of hypo and hyperglycemia of diabetic subjects. The proper use of LBGI and HBGI on CGM signals has been here enabled by developing suitable transformation functions that adapt the indicators to the characteristics of CGMs. The study was designed based on a dataset of 28 T1DM subjects monitored for up to four weeks with both SMBG and CGM systems. The alternate versions of LBGI and HBGI proposed allow references to previous works and clinically relevant cutoff values, and extend the possibility of exploiting these indicators to classify the overall quality of glucose control from CGM time-series.

The second problem we dealt with in this thesis concerned the fact that several indicators to quantify GV are available and the information conveyed by them is usually redundant. Therefore, we proposed the use of SPCA to provide a parsimonious description of GV in both T1DM and T2DM. Specifically, the approach was applied to a pool of 25 literature GV indices evaluated on 33 T1DM and 13 T2DM CGM time-series, and allowed to select a reduced subset of up to 10 indicators that could save more than 60% of the original variance in both scenarios. Of note is that a small group

of metrics was shared by the parsimonious sets of indicators in T1DM and T2DM, suggesting that some GV indices could be particularly representative of GV in diabetes regardless of the specific type of disease.
The third investigation developed in the study was aimed at managing the outputs of SPCA (i.e., PCs and sparse PCs) to develop a GV-based SVM classifier of the metabolic state of normal and diabetic subjects. Specifically, we attempted to classify the quality of glucose control of 55 T1DM subjects, and the actual metabolic condition of 34 subjects at high risk of developing T2DM, 39 IGT subjects, and 29 subjects with T2DM diagnosed. Results obtained from the analysis showed high percentages of classification accuracy in both scenarios, with 100% accuracy obtained in the classification of good vs bad glycemic control subjects with both PC and sparse PC feature configurations, and 92% accuracy in the classification of normal vs T2DM subjects with the linear kernel in the PC feature space and with the polynomial in the sparse PC one. The similar classification performance obtained with PCs and sparse PCs strengthened our speculation that SPCA can be used as a technique to obtain parsimonious descriptions of GV in diabetes.

6.2 What's next?

Further developments of the work presented in this dissertation will concern, first of all, the enlargement of the considered datasets. With a larger number of observations, the number of rows in the data matrix given as input to PCA and the ratio between number of observations and number of variables would increase, leading to a more robust estimation of the covariance matrix of the GV metrics and, thus, of the final SPCA results. Given this observation, having additional CGM recordings is also a condition necessary for extending the initial pool of indicators given as input to SPCA and providing an established linear combination of GV indices to be considered as a gold-standard metric in the characterization of GV. With a larger dataset, then, we would be able to avoid the necessity of 100 repetitions of the classification strategy proposed in Chapter 5 and we could provide a final description of the best SVM formulation to classify quality of glycemic control and metabolic condition, allowing a possible practical implementation and use of the strategy, e.g., within glucose sensors. From a methodological viewpoint, multi-class problems could be managed instead of two-class ones and, beside classifiers developed using SVM, also other less sophisticated approaches, such as the k-nearest neighbors, could be investigated; simultaneously, approaches different from the LASSO constraint (e.g., elastic net regression) could be

assessed in performing the variable selection task after the PCA step within SPCA. Results obtained from different techniques in both cases could be compared, suggesting different interpretations of the problem or strengthening the results obtained so far. Concerning the correction of LBGI and HBGI, further developments of the study will be focused on the assessment of the transformation functions with other CGM systems, to allow the application of the indicators in several scenarios involving the exploitation of different CGM devices. Also, investigating how much the transformation functions depend on the different sampling frequency of SMBG and CGM, and how much on possible sensor inaccuracy could be matter of future studies.

Bibliography

[1] American Diabetes Association, "Diagnosis and classification of diabetes mellitus," *Diabetes Care*, vol. 27, suppl. 1, pp. S5–S10, 2004.

[2] A. R. Saltiel and C. R. Kahn, "Insulin signalling and the regulation of glucose and lipid metabolism," *Nature*, vol. 414, no. 6865, pp. 799–806, 2001.

[3] P. Cryer, "Hypoglycemia: the limiting factor in the glycaemic management of type I and type II diabetes," *Diabetologia*, vol. 45, no. 7, pp. 937–948, 2002.

[4] http://www.who.int/.

[5] http://www.diabetes.org/.

[6] http://www.idf.org/.

[7] http://www.diabetes.co.uk/.

[8] M. J. Fowler, "Microvascular and macrovascular complications of diabetes," *Clin Diabetes*, vol. 26, no. 2, pp. 77–82, 2008.

[9] D. M. Nathan, "Long-term complications of diabetes mellitus," *N Engl J Med*, vol. 328, no. 23, pp. 1676–1685, 1993.

[10] D. Giugliano, A. Ceriello, and G. Paolisso, "Oxidative stress and diabetic vascular complications," *Diabetes Care*, vol. 19, no. 3, pp. 257–267, 1996.

[11] W. T. Cade, "Diabetes-related microvascular and macrovascular diseases in the physical therapy setting," *Phys Ther*, vol. 88, no. 11, pp. 1322–1335, 2008.

[12] P. H. Sonksen, S. L. Judd, and C. Lowy, "Home monitoring of blood-glucose: method for improving diabetic control," *Lancet*, vol. 311, no. 8067, pp. 729–732, 1978.

[13] E. M. Benjamin, "Self-monitoring of blood glucose: the basics," *Clin Diabetes*, vol. 20, no. 1, pp. 45–47, 2002.

[14] http://www.onetouch.com/.

[15] J. M. Evans, R. W. Newton, D. A. Ruta, T. M. MacDonald, R. J. Stevenson, and A. D. Morris, "Frequency of blood glucose monitoring in relation to glycaemic control: observational study with diabetes database," *BMJ*, vol. 319, no. 7202, pp. 83–86, 1999.

[16] M. Franciosi, F. Pellegrini, G. De Berardis, M. Belfiglio, D. Cavaliere, B. Di Nardo, S. Greenfield, S. H. Kaplan, M. Sacco, G. Tognoni, M. Valentini, and A. Nicolucci, "The impact of blood glucose self-monitoring on metabolic control and quality of life in type 2 diabetic patients: an urgent need for better educational strategies," *Diabetes Care*, vol. 24, no. 11, pp. 1870–1877, 2001.

[17] M. Cox, "An overview of continuous glucose monitoring systems," *J Pediatr Health Care*, vol. 23, no. 5, pp. 344–347, 2009.

[18] G. McGarraugh, "The chemistry of commercial continuous glucose monitors," *Diabetes Technol Ther*, vol. 11, suppl. 1, pp. S17–S24, 2009.

[19] M. Joubert and Y. Reznik, "Personal continuous glucose monitoring (CGM) in diabetes management: review of the literature and implementation for practical use," *Diabetes Res Clin Pract*, vol. 96, no. 3, pp. 294–305, 2012.

[20] J. E. Lane, J. P. Shivers, and H. Zisser, "Continuous glucose monitors: current status and future developments," *Curr Opin Endocrinol Diabetes Obes*, vol. 20, no. 2, pp. 106–111, 2013.

[21] A. Facchinetti, S. Del Favero, G. Sparacino, J. R. Castle, W. K. Ward, and C. Cobelli, "Modeling the glucose sensor error," *IEEE Trans Biomed Eng*, vol. 61, no. 3, pp. 620–629, 2014.

[22] http://www.dexcom.com/.

[23] C. Choleau, J. C. Klein, G. Reach, B. Aussedat, V. Demaria-Pesce, G. Wilson, R. Gifford, and W. K. Ward, "Calibration of a subcutaneous amperometric glucose sensor implanted for 7 days in diabetic patients. part 2. superiority of the one-point calibration method." *Biosens Bioelectron*, vol. 17, no. 8, pp. 647–654, 2002.

[24] C. Cobelli, E. Renard, and B. P. Kovatchev, "Artificial pancreas: past, present, future," *Diabetes*, vol. 60, no. 11, pp. 2672–2682, 2011.

[25] K. G. Ruedy and W. V. Tamborlane, "The landmark JDRF continuous glucose monitoring randomized trials: a look back at the accumulated evidence," *J Cardiovasc Transl Res*, vol. 5, no. 4, pp. 380–387, 2012.

[26] J. C. Pickup, S. C. Freeman, and A. J. Sutton, "Glycaemic control in type 1 diabetes during real time continuous glucose monitoring compared with self monitoring of blood glucose: meta-analysis of randomised controlled trials using individual patient data," *BMJ*, pp. 343–d3805, 2011.

[27] D. C. Klonoff, B. Buckingham, J. S. Christiansen, V. M. Montori, W. V. Tamborlane, R. A. Vigersky, and H. Wolpert, "Continuous glucose monitoring: an Endocrine Society Clinical Practice Guideline," *J Clin Endocrinol Metab*, vol. 96, no. 10, pp. 2968–2979, 2011.

[28] R. A. Vigersky, F. S.J., M. Chellappa, M. S. Walker, and N. M. Ehrhardt, "Short- and long-term effects of real-time continuous glucose monitoring in patients with type 2 diabetes," *Diabetes Care*, vol. 35, no. 1, pp. 32–38, 2012.

[29] R. M. Bergenstal, W. V. Tamborlane, A. Ahmann, J. B. Buse, G. Dailey, S. N. Davis, C. Joyce, T. Peoples, B. A. Perkins, J. B. Welsh, S. M. Willi, and M. A. Wood, "Effectiveness of sensor-augmented insulin-pump therapy in type 1 diabetes," *N Engl J Med*, vol. 363, no. 4, pp. 311–320, 2010.

[30] The Diabetes Control and Complications Trial Research Group, "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus," *N Engl J Med*, vol. 329, no. 14, pp. 977–986, 1993.

[31] The Diabetes Control and Complications Trial Research Group, "The relationship of glycemic exposure (HbA1c) to the risk of development and progression of retinopathy in the Diabetes Control and Complications Trial," *Diabetes*, vol. 44, no. 8, pp. 968–983, 1995.

[32] B. P. Kovatchev, D. J. Cox, A. Kumar, L. Gonder-Frederick, and C. W. L., "Algorithmic evaluation of metabolic control and risk of severe hypoglycemia in type 1 and type 2 diabetes using self-monitoring blood glucose data," *Diabetes Technol Ther*, vol. 5, no. 5, pp. 817–828, 2003.

[33] American Diabetes Association, "Standards of medical care in diabetes," *Diabetes Care*, vol. 329, suppl. 1, pp. S15–S35, 2004.

[34] I. B. Hirsch, "Glycemic variability: it's not just about A1C anymore," *Diabetes Technol Ther*, vol. 7, no. 5, pp. 780–783, 2005.

[35] I. B. Hirsch and M. Brownlee, "Should minimal blood glucose variability become the gold standard of glycemic control?" *J Diabetes Complications*, vol. 19, no. 3, pp. 178–181, 2005.

[36] E. S. Kilpatrick, A. S. Rigby, and S. L. Atkin, "The effect of glucose variability on the risk of microvascular complications in type 1 diabetes," *Diabetes Care*, vol. 29, no. 7, pp. 1486–1490, 2006.

[37] I. M. E. Wentholt, W. Kulik, R. P. J. Michels, J. B. L. Hoekstra, and J. H. DeVries, "Glucose fluctuations and activation of oxidative stress in patients with type 1 diabetes," *Diabetologia*, vol. 51, no. 1, pp. 183–190, 2008.

[38] C. Weber and O. Schnell, "The assessment of glycemic variability and its impact on diabetes-related complications: an overview," *Diabetes Technol Ther*, vol. 11, no. 10, pp. 623–633, 2009.

[39] M. Brownlee and I. B. Hirsch, "Glycemic variability: a hemoglobin A_{1c}-independent risk factor for diabetes complications," *JAMA*, vol. 295, no. 14, pp. 1707–1708, 2006.

[40] I. B. Hirsch and M. Brownlee, "Beyond hemoglobin A_{1c} - Need for additional markers of risk for diabetic microvascular complications," *JAMA*, vol. 303, no. 22, pp. 2291–2291, 2010.

[41] A. Ceriello, S. Kumar, L. Piconi, K. Esposito, and D. Giugliano, "Simultaneous control of hyperglycemia and oxidative stress normalizes endothelial function in type 1 diabetes," *Diabetes Care*, vol. 30, no. 3, pp. 649–654, 2007.

[42] A. Ceriello, K. Esposito, L. Piconi, M. A. Ihnat, J. E. Thorpe, R. Testa, M. Boemi, and D. Giugliano, "Oscillating glucose is more deleterious to endothelial function and oxidative stress than mean glucose in normal and type 2 diabetic patients," *Diabetes*, vol. 57, no. 5, pp. 1349–1354, 2008.

[43] D. Rodbard, "Optimizing display, analysis, interpretation and utility of self-monitoring of blood glucose (SMBG) data for management of patients with diabetes," *J Diabetes Sci Technol*, vol. 1, no. 1, pp. 62–71, 2007.

[44] D. Rodbard, "New and improved methods to characterize glycemic variability using continuous glucose monitoring," *Diabetes Technol Ther*, vol. 11, no. 9, pp. 551–565, 2009.

[45] D. Rodbard, "Interpretation of continuous glucose monitoring data: glycemic variability and quality of glycemic control," *Diabetes Technol Ther*, vol. 11, suppl. 1, pp. S55–S67, 2009.

[46] W. Clarke and B. P. Kovatchev, "Statistical tools to analyze continuous glucose monitor data," *Diabetes Technol Ther*, vol. 11, suppl. 1, pp. S45–S54, 2009.

[47] D. Rodbard, T. Bailey, L. Jovanovic, H. Zisser, R. Kaplan, and S. K. Garg, "Improved quality of glycemic control and reduced glycemic variability with use of continuous glucose monitoring," *Diabetes Technol Ther*, vol. 11, no. 11, pp. 717–723, 2009.

[48] D. Rodbard, "Glycemic variability: measurement and utility in clinical medicine and research - one viewpoint," *Diabetes Technol Ther*, vol. 13, no. 11, pp. 1077–1080, 2011.

[49] F. J. Service, "Glucose variability," *Diabetes*, vol. 62, no. 5, pp. 1398–1404, 2013.

[50] J. H. DeVries, "Glucose variability: where it is important and how to measure it," *Diabetes*, vol. 62, no. 5, pp. 1405–1408, 2013.

[51] J. C. Kuenen, R. Borg, H. Zheng, D. Schoenfeld, R. J. Heine, and D. M. Nathan, "Does glucose varibility influence the relationship between average glucose and HbA1c levels?" *Diabetes*, vol. 57, suppl. 1, p. A245, 2008.

[52] D. M. Nathan, J. C. Kuenen, R. Borg, H. Zheng, D. Schoenfeld, and R. J. Heine, "Translating the a1c assay into estimated average glucose values," *Diabetes Care*, vol. 31, no. 8, pp. 1473–1478, 2008.

[53] N. L. Pernick and D. Rodbard, "Personal computer programs to assist with self-monitoring of blood glucose and self-adjustment of insulin dosage," *Diabetes Care*, vol. 9, no. 1, pp. 61–69, 1986.

[54] D. Rodbard, "Potential role of computers in clinical investigation and management of diabetes mellitus," *Diabetes Care*, vol. 11, suppl. 1, pp. 54–61, 1988.

[55] R. S. Mazze, D. Lucido, O. Langer, K. Hartmann, and D. Rodbard, "Ambulatory glucose profile: representation of verified self-monitored blood glucose data," *Diabetes Care*, vol. 10, no. 1, pp. 111–117, 1987.

[56] R. S. Mazze, E. Strock, D. Wesley, S. Borgman, B. Morgan, R. Bergenstal, and R. Cuddihy, "Characterizing glucose exposure for individuals with normal glucose tolerance using continuous glucose monitoring and ambulatory glucose profile analysis," *Diabetes Technol Ther*, vol. 10, no. 3, pp. 149–159, 2008.

[57] S. Garg, H. Zisser, S. Schwartz, T. Bailey, R. Kaplan, S. Ellis, and L. Jovanovic, "Improvement in glycemic excursions with a transcutaneous, real-time continuous glucose sensor: a randomized controlled trial," *Diabetes Care*, vol. 29, no. 1, pp. 44–50, 2006.

[58] S. Garg and L. Jovanovic, "Relationship of fasting and hourly blood glucose levels to HbA1c values: safety, accuracy, and improvements in glucose profiles obtained using a 7-day continuous glucose sensor," *Diabetes Care*, vol. 29, no. 12, pp. 2644–2649, 2006.

[59] T. Bailey, H. Zisser, and S. Garg, "Reduction in hemoglobin A1C with real-time continuous glucose monitoring: results from a 12-week observational study," *Diabetes Techol Ther*, vol. 9, no. 3, pp. 203–210, 2007.

[60] J. M. Wojcicki, "J-index. A new proposition of the assessment of current glucose control in diabetic patients," *Horm Metab Res*, vol. 27, no. 1, pp. 41–42, 1995.

[61] F. J. Service, G. D. Molnar, J. W. Rosevear, E. Ackerman, L. C. Gatewood, and W. F. Taylor, "Mean amplitude of glycemic excursions, a measure of diabetic instability," *Diabetes*, vol. 19, no. 9, pp. 2291–2291, 1970.

[62] P. A. Baghurst, "Calculating the mean amplitude of glycemic excursions from continuous glucose monitoring data: an automated algorithm," *Diabetes technol Ther*, vol. 13, no. 3, pp. 296–302, 2011.

[63] J. Schlichtkrull, O. Munck, and M. Jersild, "The M-value, an index of blood-sugar control in diabetics," *Acta Med Scand*, vol. 177, no. 1, pp. 95–102, 1965.

[64] J. M. Wojcicki, "Mathematical descriptions of the glucose control in diabetes therapy. Analysis of the Schlichtkrull M-value," *Horm Metab Res*, vol. 27, no. 1, pp. 1–5, 1995.

[65] B. P. Kovatchev, D. J. Cox, L. A. Gonder-Frederick, and W. Clarke, "Symmetrizatioit of the blood glucose measurement scale and its applications," *Diabetes Care*, vol. 20, no. 11, pp. 1655–1658, 1997.

[66] B. P. Kovatchev, D. J. Cox, L. A. Gonder-Frederick, D. Young-Hyman, D. Schlundt, and W. Clarke, "Assessment of risk for severe hypoglycemia among adults with IDDM. Validation of the low blood glucose index," *Diabetes Care*, vol. 21, no. 11, pp. 1870–1875, 1998.

[67] B. P. Kovatchev, M. Straume, D. J. Cox, and L. S. Farhy, "Risk analysis of blood glucose data: a quantitative approach to optimizing the control of insulin dependent diabetes," *J Theor Med*, vol. 3, no. 1, pp. 1–10, 2000.

[68] B. P. Kovatchev, E. Otto, D. J. Cox, L. A. Gonder-Frederick, and W. Clarke, "Evaluation of a new measure of blood glucose variability in diabetes," *Diabetes Care*, vol. 20, no. 11, pp. 2433–2438, 2006.

[69] N. R. Hill, P. C. Hindmarsh, R. J. Stevens, I. M. Stratton, J. C. Levy, and D. R. Matthews, "A method for assessing quality of control from glucose profiles," *Diabet Med*, vol. 24, no. 7, pp. 753–758, 2007.

[70] N. R. Hill, B. Thompson, J. Bruce, D. R. Matthews, and P. C. Hindmarsh, "Glycaemic risk assessment in children and young people with type 1 diabetes mellitus," *Diabet Med*, vol. 26, no. 7, pp. 740–743, 2009.

[71] A. L. McCall and B. P. Kovatchev, "The median is not the only message: a clinician's perspective on mathematical analysis of glycemic variability and modeling in diabetes mellitus," *J Diabetes Sci Technol*, vol. 3, no. 1, pp. 3–11, 2009.

[72] B. P. Kovatchev, D. J. Cox, L. Gonder-Frederick, and W. Clarke, "Methods for quantifying self-monitoring blood glucose profiles exemplified by an examination of blood glucose patterns in patients with type 1 and type 2 diabetes," *Diabetes Technol Ther*, vol. 4, no. 3, pp. 295–303, 2002.

[73] B. P. Kovatchev, M. D. Breton, and W. Clarke, "Analytical methods for the retrieval and interpretation of continuous glucose monitoring data in diabetes," *Methods Enzymol*, vol. 454, no. 1, pp. 69–86, 2009.

[74] B. P. Kovatchev, W. Clarke, M. D. Breton, K. Brayaman, and M. A. L., "Quantifying temporal glucose variability in diabetes via continuous glucose monitoring:

mathematical methods and clinical application," *Diabetes Technol Ther*, vol. 7, no. 6, pp. 849–862, 2005.

[75] B. P. Kovatchev, W. Clarke, M. D. Breton, K. Brayaman, and M. A. L., "Quantifying temporal glucose variability in diabetes via continuous glucose monitoring: mathematical methods and clinical application," *Diabetes Technol Ther*, vol. 7, no. 6, pp. 849–862, 2005.

[76] D. J. Cox, L. Gonder-Frederick, L. Ritterband, W. Clarke, and B. P. Kovatchev, "Prediction of severe hypoglycemia," *Diabetes Care*, vol. 30, no. 6, pp. 1370–1373, 2007.

[77] V. Perea, A. Amor, M. Giménez, J. Blanco, and I. Conget, "Glycemic variability measures in a group of subjects with type 1 diabetes and repeated severe and non-severe hypoglycemia," *J Diabetes Sci Technol*, vol. 7, no. 1, pp. 289–290, 2013.

[78] Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group, "Factors predictive of severe hypoglycemia in type 1 diabetes analysis from the juvenile diabetes research foundation continuous glucose monitoring randomized control trial dataset," *Diabetes Care*, vol. 34, no. 3, pp. 586–590, 2011.

[79] C. Fabris, S. D. Patek, and M. D. Breton, "Are risk indices derived from CGM interchangeable with SMBG-based indices?" Submitted to *J Diabetes Sci Technol*.

[80] H. Zou, T. Hastie, and R. Tibshirani, "Sparse principal component analysis," *J Comput Graph Stat*, vol. 15, no. 2, pp. 265–286, 2006.

[81] C. Fabris, A. Facchinetti, G. Sparacino, S. Guerra, A. Maran, and C. Cobelli, "Glucose variability indices in type 1 diabetes: parsimonious set of indices revealed by sparse principal component analysis," *Diabetes Technol Ther*, vol. 16, no. 10, pp. 644–652, 2014.

[82] C. Fabris, A. Facchinetti, G. Sparacino, and C. Cobelli, "Parsimonious description of glucose variability in type 2 diabetes by sparse principal component analysis," Submitted to *J Diabetes Sci Technol*.

[83] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 1986.

[84] http://www.diadvisor.eu/.

[85] http://www.medtronicdiabetes.net/.

[86] http://www.mosaicproject.eu/.

[87] D. Rodbard, "Clinical interpretation of indices of quality of glycemic control and glycemic variability," *Postgrad Med*, vol. 123, no. 4, pp. 107–118, 2011.

[88] C. R. Marling, N. W. Struble, R. C. Bunescu, J. H. Shubrook, and F. L. Schwartz, "A consensus perceived glycemic variability metric," *J Diabetes Sci Technol*, vol. 7, no. 4, pp. 871–879, 2013.

[89] S. Jung and J. S. Marron, "PCA consistency in High Dimension, Low Sample Size context," *Ann Stat*, vol. 37, no. 6B, pp. 4104–4130, 2009.

[90] S. Jung, A. Sen, and J. S. Marron, "Boundary behavior in High Dimension, Low Sample Size asymptotics of PCA," *J Multivar Anal*, vol. 109, no. 1, pp. 190–203, 2012.

[91] D. Shen, H. Shen, and J. S. Marron, "Consistency of sparse PCA in High Dimension, Low Sample Size contexts," *J Multivar Anal*, vol. 115, no. 1, pp. 317–333, 2013.

[92] F. J. Service and R. L. Nelson, "Characteristics of glycemic stability," *Diabetes Care*, vol. 3, no. 1, pp. 58–62, 1980.

[93] C. M. McDonnell, S. M. Donath, S. I. Vidmar, W. G. A., and F. J. Cameron, "A novel approach to continuous glucose analysis utilizing glycemic variation," *Diabetes Technol Ther*, vol. 7, no. 2, pp. 253–263, 2005.

[94] S. Guerra, G. Sparacino, A. Facchinetti, M. Schiavon, C. Dalla Man, and C. Cobelli, "A dynamic risk measure from continuous glucose monitoring data," *Diabetes Technol Ther*, vol. 13, no. 8, pp. 843–852, 2011.

[95] L. Crenier, "Poincaré plot quantification for assessing glucose variability from continuous glucose monitoring systems and a new risk marker for hypoglycemia: application to type 1 diabetes patients switching to continuous subcutaneous insulin infusion," *Diabetes Technol Ther*, vol. 16, no. 4, pp. 247–254, 2014.

[96] L. Vigil, E. Condes, M. Varela, C. Rodriguez, A. Colas, B. Vargas, M. Lopez, and E. Cirugeda, "Glucose series complexity in hypertensive patients," *J Am Soc Hypertens*, vol. 8, no. 9, pp. 630–636, 2014.

[97] C. Fabris, G. Sparacino, A. S. Sejling, A. Goljahani, J. Duun-Henriksen, L. S. Remvig, C. B. Juhl, and C. Cobelli, "Hypoglycemia-related electroencephalogram changes assessed by multiscale entropy," *Diabetes Technol Ther*, vol. 16, no. 10, pp. 688–694, 2014.

[98] B. Scholkopf and A. J. Smola, *Learning with Kernels - Support Vector Machines, Regularization, Optimization, and Beyond.* Cambridge, MA: The MIT Press, 2002.

[99] C. M. Bishop, *Pattern recognition and machine learning.* New York: Springer, 2006.

[100] M. H. Jensen, T. F. Christensen, L. Tarnow, M. D. Johansen, and O. K. Hejlesen, "An information and communication technology system to detect hypoglycemia in people with type 1 diabetes," *Stud Health Technol Inform*, vol. 192, no. 1, pp. 38–41, 2013.

[101] M. H. Jensen, T. F. Christensen, L. Tarnow, Z. Mahmoudi, M. D. Johansen, and O. K. Hejlesen, "Professional continuous glucose monitoring in subjects with type 1 diabetes: retrospective hypoglycemia detection," *J Diabetes Sci Technol*, vol. 7, no. 1, pp. 38–41, 2013.

[102] M. H. Jensen, T. F. Christensen, L. Tarnow, E. Seto, M. D. Johansen, and O. K. Hejlesen, "Real-time hypoglycemia detection from continuous glucose monitoring data of subjects with type 1 diabetes," *Diabetes Technol Ther*, vol. 15, no. 7, pp. 538–543, 2013.

[103] J. Bondia, C. Tarin, W. Garcia-Gabin, E. Esteve, J. M. Fernández-Real, W. Ricart, and J. Vehi, "Using support vector machines to detect therapeutically incorrect measurements by the MiniMed CGMS," *J Diabetes Sci Technol*, vol. 2, no. 4, pp. 622–629, 2008.

[104] Y. Leal, L. Gonzalez-Abril, C. Lorencio, J. Bondia, and J. Vehi, "Detection of correct and incorrect measurements in real-time continuous glucose monitoring systems by applying a postprocessing support vector machine," *IEEE Trans Biomed Eng*, vol. 60, no. 7, pp. 1891–1899, 2013.