Università degli Studi di Padova

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

UNIVERSITY OF PADUA

DEPARTMENT OF INFORMATION ENGINEERING

Master's Degree in Bioengineering

A new approach for automatic extraction of arterial function from DCE-MRI and estimation of cerebral blood flow

Supervisor

Prof.ssa Alessandra Bertoldo

Co-supervisor

Prof.ssa Letizia Squarcina

Graduand

Emma Cusin

Student ID

2120781

Academic Year 2025 - 2026

Graduation Date

July 6th, 2026

Abstract

Dynamic contrast-enhanced MRI (DCE-MRI) enables the quantitative characterization of tissue perfusion and microvascular permeability, but pharmacokinetic analysis requires an accurate estimation of the arterial input function (AIF). Manual AIF selection is operator-dependent and time-consuming, while population-averaged AIFs do not capture individual hemodynamics. This thesis develops and validates an automatic pipeline for AIF extraction from cerebral DCE-MRI, based on principal component analysis (PCA) of voxel-wise signal time-courses combined with physiological filtering and an explicit quality-control mechanism. The proposed approach evolves from a literature standard reference method through a set of structural adaptations to brain DCE-MRI and clinical safeguards: an extended candidate principal component search, plausibility filters with a fail-safe behavior that halts the extraction when no arterial component is reliably identifiable, and a composite ranking score that penalizes contaminated component morphologies. The pipeline is evaluated on synthetic phantoms with controlled Rician noise and partial volume contamination, characterizing its operating range and failure modes. It is then applied to a clinical cohort of brain DCE-MRI examinations, where the response of the pipeline across heterogeneous patient profiles is documented case by case, including instances where the fail- safe mechanism correctly refuses extraction. Cerebral blood flow, blood volume and mean transit time maps are subsequently estimated via SVD-regularized deconvolution and presented as a downstream, end-to-end check of the pipeline, with the limits of absolute quantitative calibration explicitly discussed.

Sommario

La risonanza magnetica dinamica con mezzo di contrasto (DCE-MRI) consente la caratterizzazione quantitativa della perfusione tissutale e della permeabilità microvascolare, ma l'analisi farmacocinetica richiede una stima accurata della funzione di input arteriosa (AIF). La selezione manuale dell'AIF è dipendente dall'operatore e dispendiosa in termini di tempo, mentre le AIF mediate sulla popolazione non catturano l'emodinamica individuale. Questa tesi sviluppa e valida una pipeline automatica per l'estrazione dell'AIF da DCE-MRI cerebrale, basata sull'analisi delle componenti principali (PCA) delle curve temporali di segnale voxel per voxel, combinata con un filtraggio fisiologico e un meccanismo esplicito di controllo di qualità. L'approccio proposto evolve da un metodo di riferimento standard della letteratura attraverso un insieme di adattamenti strutturali alla DCE-MRI cerebrale e di salvaguardie cliniche: una ricerca estesa delle componenti principali candidate, filtri di plausibilità con un comportamento fail-safe che interrompe l'estrazione quando nessuna componente arteriosa è identificabile in modo affidabile, e un punteggio composito di ranking che penalizza le morfologie di componente contaminate. La pipeline è valutata su fantocci sintetici con rumore Riciano controllato e contaminazione da volume parziale, caratterizzandone il campo operativo e le modalità di fallimento. Viene quindi applicata a una coorte clinica di esami DCE-MRI cerebrali, in cui la risposta della pipeline attraverso profili di pazienti eterogenei è documentata caso per caso, inclusi i casi in cui il meccanismo fail-safe rifiuta correttamente l'estrazione. Le mappe di flusso ematico cerebrale, volume ematico e tempo di transito medio sono successivamente stimate tramite deconvoluzione regolarizzata con SVD e presentate come verifica a valle dell'intera pipeline, discutendo esplicitamente i limiti della calibrazione quantitativa assoluta.

Contents

List of Figures

List of Tables

1

Introduction

1.1 Background and motivation

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a functional imaging technique that characterizes tissue microvasculature through the temporal evolution of a paramagnetic contrast agent. After the intravenous injection of a gadolinium-based bolus, the agent shortens the longitudinal relaxation time T_1 of the surrounding water protons. This produces a transient increase in the signal of T_1-weighted images which, under appropriate acquisition and modelling assumptions, can be related to the local contrast-agent concentration [1]. By acquiring a series of volumes before, during, and after the passage of the bolus, DCE-MRI provides, for each voxel, a dynamic enhancement curve that reflects the local delivery, extravasation, and wash-out of the agent.

These curves can be interpreted through kinetic models, which relate the measured tissue enhancement to physiological parameters of the microvasculature, ranging from descriptors of capillary permeability to measures of tissue perfusion [2, 3]. In the present work, the focus is on perfusion-related quantities derived from AIF-based deconvolution, namely cerebral blood flow, cerebral blood volume, and mean transit time. When the signal is converted to contrast-agent concentration, these quantities can provide quantitative, spatially resolved information on tissue microvasculature. In the present thesis, however, no T_1 mapping was used, and the resulting maps are therefore interpreted as semi-quantitative indices. In the brain, the sensitivity of DCE-MRI to vascular and blood–brain barrier alterations has motivated its application in oncology, cerebrovascular and neurodegenerative disease, and, more recently, psychiatric research [4, 5].

The quantitative interpretation of DCE-MRI relies on a quantity that is external to the tissue model itself: the arterial input function (AIF), defined as the concentration-time curve of the contrast agent in the arterial blood that supplies the tissue under examination.

In the convolution formulation common to tracer-kinetic and perfusion models, the AIF is the forcing function, and the tissue curve is obtained as its convolution with a tissue-specific response. The estimated physiological parameters are therefore determined by the AIF as much as by the tissue data. A systematic error in the AIF, whether in its amplitude, timing, or shape, propagates into the parameter estimates and cannot be compensated for simply by refining the tissue model [6]. An accurate and reproducible estimation of the AIF is therefore a central requirement for quantitative and semi-quantitative DCE-MRI analysis.

1.2 The challenge of arterial input function estimation

Despite its central role, the AIF remains one of the principal sources of uncertainty in quantitative DCE-MRI. The difficulty is both practical and physical, and the strategies proposed in the literature reflect a trade-off between fidelity to the individual subject and robustness of the estimate.

The most direct strategy is manual selection, in which an operator delineates a region of interest over a visible artery and averages the corresponding time courses. Although conceptually simple, this procedure depends on the experience of the operator, is subject to appreciable intra- and inter-observer variability, and does not scale to large cohorts. At the opposite end, the individual AIF can be replaced by a population-averaged function, obtained by fitting a fixed analytical form to the AIFs measured in a reference cohort [7]. This eliminates the operator dependence and yields a smooth, noise-free curve, but it imposes a single predetermined shape on every subject. It therefore cannot represent inter-individual differences in cardiac output, injection protocol, or vascular state [6].

Estimating the AIF from the individual acquisition preserves the subject specificity, but it must contend with several degrading factors. The caliber of the feeding arteries is frequently comparable to, or smaller than, the voxel size, so that a voxel positioned over a vessel mixes intravascular blood with the surrounding tissue. This partial volume effect attenuates and distorts the measured arterial curve. Voxels dominated by arterial signal are correspondingly scarce, and they are not straightforward to identify. The measurement is further degraded by acquisition noise and, in vivo, by motion and inflow effects. As a consequence, the automatic and reliable extraction of an individual AIF remains an open methodological problem.

A substantial body of work has addressed this problem with data-driven techniques, which aim to identify the arterial voxels automatically from the structure of the data rather than through manual intervention. Clustering methods partition the voxels according to the morphology of their time courses and designate the group whose representative

curve is most consistent with arterial behavior [8, 9]. Independent component analysis seeks to isolate the arterial contribution as a statistically independent source within the mixed signals [10]. Principal component analysis (PCA) decomposes the ensemble of time courses into orthogonal components ordered by explained variance, under the hypothesis that the arterial dynamics are concentrated in one of the leading components [11]. These methods are automatic and subject-specific, and under suitable conditions they have been reported to improve robustness and reproducibility with respect to manual or population-based approaches. They nonetheless share a common premise, namely that the arterial signal forms a distinct and identifiable population within the data. This premise holds when sufficiently pure arterial voxels are present, but it is precisely the condition that partial-volume mixing undermines.

1.3 Aim and contributions

The aim of this thesis is the development and validation of an automatic pipeline for the estimation of the AIF from cerebral DCE-MRI. The method extends a PCA-based reference approach from the literature [11], retaining its data-driven core while adapting it to the brain and equipping it with explicit physiological and quality-control mechanisms. To the principal component decomposition, the pipeline adds an extended search over candidate components, a set of plausibility constraints on the admissible arterial waveform, a composite ranking score that down-weights components with non-arterial morphology, and a fail-safe rule that suspends the extraction when no component satisfies the arterial criteria. The intended outcome is a procedure that operates without user intervention and that signals, rather than conceals, the cases in which a reliable arterial component cannot be identified.

The pipeline is evaluated in two complementary settings. It is first assessed on synthetic phantoms, for which a ground-truth AIF and a known arterial geometry are available. This setting allows the behavior of the method to be quantified under controlled Rician noise and controlled partial-volume contamination, and its operating range and failure modes to be delineated. The pipeline is then applied to a clinical cohort of brain DCE-MRI examinations, where its response is documented on a per-subject basis, including the case in which the fail-safe declines to return an estimate because no physiologically plausible arterial component is identified. Finally, as an end-to-end demonstration of the complete processing chain, cerebral blood flow, cerebral blood volume, and mean transit time are reconstructed from the extracted AIF by regularized deconvolution, and the conditions that limit their absolute quantification are examined.

Two results recur throughout the analysis and frame the discussion. The method

is comparatively robust to Rician noise, but markedly sensitive to the heterogeneity of partial-volume mixing, which emerges as the factor that most strongly governs its behavior on realistic data. In addition, a high correlation between the recovered AIF and a reference waveform is shown not to guarantee, on its own, that the underlying voxel selection is correct. This observation has direct implications for the validation of automatic AIF methods and constitutes one of the central methodological contributions of this work.

1.4 Outline of the thesis

Chapter 2 reviews the principles of DCE-MRI, the role of the AIF in pharmacokinetic and perfusion quantification, and the main families of AIF estimation methods. Chapter 3 describes the proposed pipeline in detail, together with its two versions, the synthetic and clinical datasets, and the deconvolution procedure used to obtain the perfusion maps. Chapter 4 reports the synthetic validation, examining the response of the method to noise and to partial volume. Chapter 5 reports the application of the pipeline to the clinical cohort and presents the resulting arterial input functions and perfusion maps. Chapter 6 discusses these results, relates them to previous work, and sets out the limitations of the approach. Chapter 7 draws the conclusions and outlines directions for future work.

2

Background: DCE-MRI and arterial input function

This chapter introduces the methodological background required to follow the content of the thesis. Section 2.1 describes the physical principles and the acquisition protocol adopted for dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), together with the relationship between the intensity of the MR signal and the concentration of contrast agents. Section 2.2 introduces the arterial input function (AIF), discusses its role in the quantitative analysis of DCE-MRI data, and reviews the practical difficulties found in its estimation. Section 2.3 provides an overview of the two quantitative frameworks used in the rest of this work, the Tofts pharmacokinetic model and deconvolution-based perfusion quantification. Finally, Section 2.4 surveys the main approaches proposed in the literature for AIF estimation, discusses their respective limitations, and identifies the open challenges that motivate the methodological work presented in Chapter 3.

2.1 DCE-MRI: acquisition principles

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a non-invasive imaging technique in which a time series of T_1-weighted volumes is acquired before, during, and after the intravenous administration of a paramagnetic contrast agent (CA). For every voxel, the resulting four-dimensional dataset provides a time course of the MR signal intensity. This reflects the local pharmacokinetics of the tracer as it enters, distributes within, and washes out of the tissue. Because the temporal evolution of the signal is governed by the microvascular properties of the tissue, DCE-MRI is used mainly to characterize vascular density, integrity and permeability, with applications in oncology,

stroke and neurodegenerative diseases [4, 1, 5].

2.1.1 Physical principles of T_1 and T_1-weighted imaging

MR images derive their contrast from the nuclear magnetic resonance properties of the hydrogen nuclei contained in tissue water. When placed in a strong static magnetic field B_0, these nuclei develop a net macroscopic magnetization M_0 aligned with the field. When a radio-frequency (RF) pulse is applied, the magnetization rotates away from its equilibrium state. After the pulse, the longitudinal component M_z recovers exponentially to M_0 with a tissue-specific time constant, the longitudinal relaxation time T_1:

$$M_z(t) \;=\; M_0 \left[\, 1 - e^{-t/T_1} \,\right]. \tag{2.1}$$

Equation (2.1) describes the ideal recovery of longitudinal magnetization after an excitation pulse and is introduced here to illustrate the physical meaning of T_1. However, in a real MR sequence, image contrast is determined not only by tissue relaxation times, but also by the sequence parameters. Two parameters are particularly important: the repetition time TR, defined as the time interval between successive excitation pulses, and the echo time TE, defined as the time between the excitation pulse and the signal readout.

In T_1-weighted imaging, TR is chosen sufficiently short that the longitudinal magnetization does not fully recover before the next RF pulse. Under these conditions, tissues with shorter T_1 recover a larger fraction of their longitudinal magnetization between successive excitations and therefore tend to produce a higher signal than tissues with longer T_1. In contrast, if TR were much longer than the T_1 values of the tissues of interest, the longitudinal magnetization would almost completely recover in all tissues, and the image contrast would depend mainly on proton density rather than on T_1.

A short TE is usually chosen together with a short TR to reduce the influence of transverse relaxation, especially T_2^* decay, on the measured signal. T_2^* is the effective transverse relaxation time and accounts for both intrinsic spin–spin relaxation and dephasing caused by local magnetic-field inhomogeneities. By minimizing T_2^*-related signal loss, the image contrast remains dominated by differences in T_1 [1, 3].

This principle is exploited in DCE-MRI using fast T_1-weighted spoiled gradient-echo sequences. When a gadolinium-based contrast agent arrives in a voxel, it shortens the local T_1, allowing faster recovery of the longitudinal magnetization between RF pulses and therefore increasing the measured signal intensity. Repeated acquisition of T_1-weighted volumes makes it possible to follow the arrival, distribution and wash-out of the tracer as a dynamic signal enhancement.

2.1.2 Gadolinium-based contrast agents and their effect on T_1

The CAs used in DCE-MRI are typically gadolinium-based chelates such as gadopentetate dimeglumine (Gd-DTPA) and gadoterate meglumine (Gd-DOTA) [2, 12]. They are paramagnetic compounds, meaning that the gadolinium ion has seven unpaired electrons whose magnetic moments generate local fluctuating fields, which efficiently shorten the longitudinal relaxation time of nearby water protons [1]. To a first approximation, the relationship between tissue CA concentration $C(t)$ and local T_1 is described by the relaxivity equation

$$\frac{1}{T_1(t)} = \frac{1}{T_{1,0}} + r_1 C(t), \tag{2.2}$$

where $T_{1,0}$ is the pre-contrast longitudinal relaxation time and r_1 is the longitudinal relaxivity of the CA, which depends on the specific agent and on the static magnetic field strength. The parameter $T_{1,0}$ reflects the intrinsic recovery rate of the tissue before any contrast agent is present [2, 3]. The consequence is that, as the CA arrives in a voxel, T_1 decreases and the T_1-weighted signal intensity increases; as the CA washes out, T_1 and signal intensity tend to return towards their baseline values, although persistent enhancement may be observed in tissues with substantial extravascular retention. The amplitude and the temporal profile of the signal carry information on the underlying microvascular dynamics.

After intravenous administration of the bolus, low-molecular weight gadolinium chelates rapidly distribute within the blood pool and, in tissues whose vessels are permeable to small molecules, leak into the extravascular extracellular space (EES) [2, 3]. In the healthy brain, the intact blood–brain barrier strongly limits the extravasation of gadolinium-based contrast agents, so that the agent remains predominantly intravascular. However, in pathological conditions in which the barrier is disrupted, the agent leaks into the EES, and the resulting signal–time curves carry quantitative information on barrier permeability. This can be used to characterize diseases such as cerebral small vessel disease, vascular dementia, and Alzheimer's disease [5].

2.1.3 The DCE-MRI acquisition

A DCE-MRI acquisition consists of a dynamic series of three-dimensional T_1-weighted volumes covering the anatomical region of interest, acquired with a fast spoiled gradient-echo sequence characterized by short TR, short TE, and a moderate flip angle [1, 3]. After a few baseline volumes used to establish the pre-contrast signal level, the CA is administered as an intravenous bolus through a power injector, and image acquisition continues for several minutes so that both the wash-in and the (partial) wash-out of the tracer are captured. As a result, the temporal sequence of each voxel intensity exhibits three distinct

phases: a flat *baseline*, a rapid *wash-in* corresponding to the arrival of the bolus, and a slower *wash-out* (or a near-plateau in tissues with significant CA retention) [1].

The design of a DCE-MRI protocol involves a fundamental trade-off between spatial resolution, temporal resolution, and signal-to-noise ratio (SNR). A shorter TR improves temporal sampling, which is essential to capture the sharp and narrow arterial peak that follows bolus injection, but simultaneously reduces SNR and constrains achievable spatial resolution [3]. In neuroimaging, typical acquisition parameters consist of a static field of 3 T, millimetric spatial resolution, whole-brain or near whole-brain coverage, and temporal resolutions of the order of a few seconds per volume, sustained for a total acquisition time of approximately five to ten minutes.

2.1.4 From signal intensity to pseudo-concentration

Although the quantity of physiological interest is the local CA concentration $C(t)$, what is ultimately measured is the signal intensity $S(t)$, obtained from the scanner output through post-processing, and the conversion between the two is non-trivial. For a spoiled gradient-echo sequence in the steady state, the signal is a non-linear function of T_1. Combined with the relaxivity equation (2.2), this relationship allows the explicit estimation of $C(t)$ from $S(t)$, provided that the pre-contrast longitudinal relaxation time $T_{1,0}$ and the relaxivity r_1 are known [2, 3]. The most rigorous approach therefore requires an independent T_1 mapping step to be performed before dynamic acquisition.

An alternative, the one adopted throughout this work, is the so-called *semi-quantitative* analysis, in which a normalized signal-difference curve is used as a surrogate of concentration:

$$\Delta S(t) \;=\; \frac{S(t) - S_0}{S_0}, \tag{2.3}$$

where S_0 denotes the baseline signal averaged over the pre-contrast frames. Under the standard assumptions of small flip angle and small CA-induced changes in T_1, $\Delta S(t)$ is approximately proportional to $C(t)$ [2, 3]. This *pseudo-concentration* preserves the shape of the underlying concentration–time curve and is therefore adequate for shape-based analysis such as data-driven AIF extraction. However, since the proportionality constant between $\Delta S(t)$ and $C(t)$ is not recovered, parameters derived from these curves – in particular flow- and volume-related indices such as cerebral blood flow and cerebral blood volume – are obtained on an arbitrary scale and must be interpreted as semi-quantitative. When all subjects are acquired with the same protocol and processed with the same normalization procedure, pseudo-concentration curves can still be used for relative and shape-based comparisons across subjects. Nevertheless, they should not be interpreted as absolute contrast-agent concentrations. Accordingly, the curves used in the follow-

ing chapters should be interpreted as pseudo-concentration curves rather than absolute contrast-agent concentrations.

2.2 The arterial input function

The signal–time curve measured in a DCE-MRI experiment reflects the joint contribution of the tracer delivered to the tissue and of the tissue's intrinsic vascular and exchange properties. To recover the latter contribution, which represents the actual physiological quantity of interest, one needs to account for the arterial input of the contrast agent into the tissue, namely the arterial input function (AIF). The AIF is therefore fundamental for any quantitative analysis of DCE-MRI data, and its accurate estimation is the central methodological challenge addressed in this thesis.

2.2.1 Definition and physiological role

The arterial input function (AIF) describes the time course of the concentration of contrast agents in the arterial blood that supplies the tissue of interest [6]. Depending on the modeling context, this input function can be expressed either as the concentration in whole blood $C_b(t)$, or as the corresponding plasma concentration $C_p(t)$. This distinction is important because pharmacokinetic models are usually formulated in terms of plasma concentration. If the measured AIF represents the whole-blood concentration, it can be converted to plasma concentration as

$$C_p(t) = \frac{C_b(t)}{1 - \mathrm{Hct}}, \tag{2.4}$$

where Hct is the arterial hematocrit, which is the fraction of whole blood volume occupied by red blood cells. However, in the present work, the dynamic curves are derived from a relative signal enhancement rather than an absolute signal-to-concentration conversion. Therefore, the estimated AIF should be interpreted as an arterial pseudo-concentration curve.

In the indicator-dilution framework, the tissue concentration–time curve, or the corresponding pseudo-concentration curve in a semi-quantitative analysis, is related to the arterial input through a convolution:

$$C_t(t) \;=\; F \cdot \big[\, C_{\mathrm{AIF}}(t) \otimes R(t) \,\big], \tag{2.5}$$

where $C_{\mathrm{AIF}}(t)$ denotes the arterial input curve used in the analysis, F is the blood flow per unit mass of tissue, $\otimes$ denotes the temporal convolution, and $R(t)$ is the *impulse residue function*, that is the dimensionless curve that describes the fraction of tracer still present

in the voxel at time t after an instantaneous unit injection. The product $I(t) = F\,R(t)$, often called the *flow-scaled impulse residue function*, is the quantity directly recovered by deconvolution and is discussed further in Section 2.3.

Equation (2.5) makes the importance of the AIF clear. First, it is the only term that encodes the amount of tracer actually *delivered* to the tissue: any error in the AIF, whether it is a bias in its amplitude or a distortion of its shape, propagates directly into $I(t)$ and therefore into all derived parameters related to perfusion, such as CBF, CBV and MTT, as well as into pharmacokinetic parameters such as K^{trans}, v_e and v_p. Second, the AIF is in principle a *global* quantity shared by all voxels perfused by the same arterial territory, while $R(t)$ is voxel-specific and reflects the local microvascular architecture. Separating the tissue response from the arterial input is therefore essential to obtain perfusion estimates that reflect intrinsic tissue properties rather than the particular shape of the input.

2.2.2 Practical challenges in AIF estimation

Although the AIF is conceptually well defined, its accurate measurement *in vivo* is notoriously difficult, and the literature has consistently identified it as the largest source of variability in quantitative perfusion imaging [6]. The main practical difficulties relevant to brain DCE-MRI are summarized below.

Partial volume effect. The intracerebral arteries that can be visualized in a typical DCE-MRI acquisition have luminal diameters comparable to, or smaller than, the voxel size of the dynamic sequence. The signal of an "arterial" voxel is therefore inevitably a mixture of the intravascular blood signal and the contribution from the surrounding tissue. Since the normalized signal enhancement $\Delta S(t)$ of pure arterial blood is much higher than that of the parenchyma, partial-volume contamination systematically *underestimates* the peak amplitude of the AIF [6].

Inflow effects. In spoiled gradient-echo sequences, fresh inflowing blood, which has not been fully saturated by the imaging RF pulses, can produce an *inflow-enhanced* signal in proximal arterial voxels. This effect introduces a signal contribution that is not solely determined by local T_1 shortening. As a result, the arterial signal enhancement may be biased, and the measured AIF may no longer be proportional to the actual contrast agent concentration. The effect is particularly pronounced in 3D acquisitions with short TR and relatively thick slabs [6].

Signal non-linearity at the peak. At the bolus peak, the arterial concentration of the contrast agent can be high enough for the simple linear relationship between relative

signal enhancement and concentration to break down. In addition to the non-linear spoiled gradient-echo signal equation, susceptibility-related T_2^* effects and saturation phenomena may contribute to an underestimation or distortion of the arterial peak.

Inter-subject variability. Even when all the technical difficulties above are controlled, significant variability in the AIF arises from patient-specific factors: cardiac output, injection rate, vascular geometry, hematocrit. These factors influence both the amplitude and the temporal profile (bolus arrival, peak width) of the AIF, and motivate the use of *individual* AIFs whenever possible, versus population-averaged templates [6, 7].

Low signal-to-noise ratio. Arterial voxels are few in number relative to the total brain volume, and their pre-contrast signal can be low (especially in low-flip-angle 3D acquisitions). Estimating the AIF from such voxels leads to curves with a markedly lower SNR than the tissue curves themselves, particularly in the early baseline phase that is used for normalization.

Together, these difficulties motivate the development of *automated* and *data-driven* methods for AIF estimation, which attempt to identify the most arterial voxels available in a given dataset and combine them into a single representative curve. The main families of such methods are reviewed in Section 2.4.

2.3 Pharmacokinetic and perfusion quantification

Once an arterial input curve $C_{\mathrm{AIF}}(t)$ and the corresponding tissue dynamic curves are available, the quantitative or semi-quantitative analysis of DCE-MRI data can proceed within two complementary frameworks: a *model-based* pharmacokinetic analysis, in which a compartmental model of contrast agent exchange between blood and tissue is fitted to the data, and a deconvolution-based analysis, which does not impose a specific compartmental exchange model but still relies on the assumptions of indicator-dilution theory. The following sections describe both approaches in the form used in the rest of this work [3].

2.3.1 Pharmacokinetic modelling: the Tofts model

The Tofts model [2, 12] describes how a low-molecular-weight contrast agent distributes between the bloodstream and the tissue after injection. The model considers two compartments: the blood plasma and the extravascular extracellular space (EES), which is the fluid-filled space between cells outside the vessels. Plasma occupies a fraction v_p of total tissue volume and EES occupies a fraction v_e. Once the contrast agent reaches the

tissue through the bloodstream, it crosses the vessel wall into the EES at a rate governed by the *volume transfer constant* K^{trans}. Neglecting the intravascular plasma contribution to the measured tissue concentration, the standard Tofts model describes the tissue concentration as follows:

$$C_t(t) \;=\; K^{\text{trans}} \int_0^t C_p(\tau)\, \exp[-k_{ep}\,(t-\tau)]\; d\tau, \tag{2.6}$$

where $C_p(t)$ is the arterial plasma concentration of the contrast agent, and $k_{ep} = K^{\text{trans}}/v_e$ is the rate constant that describes the return of the tracer from the EES back to the plasma. If the measured input function is expressed as a whole-blood concentration $C_b(t)$, the plasma concentration is obtained through Equation (2.4).

The physiological interpretation of K^{trans} depends on local tissue conditions [12]. When the vessel wall is highly permeable and exchange is flow-limited, K^{trans} approximates the plasma flow per unit tissue volume. In contrast, when the barrier is mostly intact and only slightly permeable, as in the healthy brain, K^{trans} approximates the permeability–surface area product per unit of tissue volume and therefore reflects the degree of disruption of the blood–brain barrier.

When the intravascular plasma contribution to the measured tissue concentration cannot be neglected, the model is extended to include an explicit plasma-volume term:

$$C_t(t) \;=\; v_p\, C_p(t) \;+\; K^{\text{trans}} \int_0^t C_p(\tau)\, \exp[-k_{ep}\,(t-\tau)]\; d\tau. \tag{2.7}$$

This *extended Tofts model* [3] is particularly relevant in the setting of cerebral small-vessel-disease, where a further simplification (the Patlak model) is widely used to detect subtle leakage of the blood–brain barrier [5].

In this thesis, equations (2.6)–(2.7) are presented as the standard quantitative framework for the estimation of permeability with DCE-MRI. They are not the main subject of the methodological work presented in this thesis; rather, they illustrate why an accurate arterial input curve is essential: without a reliable estimate of $C_p(t)$, or of the corresponding pseudo-concentration input curve in a semi-quantitative setting, none of the above parameters can be robustly estimated.

2.3.2 Perfusion quantification: deconvolution-based approach

Cerebral perfusion refers to the delivery of blood to brain tissue per unit of time. It is commonly described by three haemodynamic parameters: cerebral blood flow (CBF), cerebral blood volume (CBV), and mean transit time (MTT). CBF is defined as the volume of blood passing through a unit mass of tissue per unit of time, typically expressed in $\text{mL}/(100\,\text{g}\cdot\text{min})$. CBV is the volume of blood contained within a unit mass of tissue,

typically expressed in mL/100 g. MTT is the average time required for blood to traverse the microvascular bed and is expressed in seconds. These parameters are related by the central volume theorem, MTT = CBV/CBF [13]. Spatial maps of CBF, CBV and MTT are commonly referred to as perfusion parametric maps and provide voxel-wise information on the haemodynamic status of the tissue.

Although DCE-MRI is most commonly associated with the estimation of permeability-related parameters, the first-pass portion of the dynamic curve can also be interpreted within the framework of indicator-dilution theory. In this setting, deconvolution-based analysis does not provide fully quantitative perfusion measurements unless an accurate conversion from signal intensity to contrast-agent concentration is available. However, when relative signal-enhancement curves are used as pseudo-concentration curves, the same framework can still provide semi-quantitative indices related to perfusion, provided that the limitations imposed by temporal resolution, signal non-linearity, partial-volume effects and AIF uncertainty are explicitly acknowledged.

Within this framework, the objective is to recover the flow-scaled residue function $I(t) = F\,R(t)$ by deconvolving the tissue pseudo-concentration curve with the estimated AIF [14, 15]. Once $I(t)$ is known, flow- and volume-related quantities follow from indicator-dilution theory:

$$F \;=\; \max_{t} I(t), \qquad V \;=\; \int_0^{\infty} I(t)\,dt, \qquad \mathrm{MTT} \;=\; \frac{V}{F}, \tag{2.8}$$

where F and V are proportional to blood flow and blood volume, respectively. When absolute concentration curves are available, these quantities can be converted, after appropriate corrections for tissue density and haematocrit, into CBF and CBV with conventional physiological units. However, in the semi-quantitative setting adopted in this thesis the same expressions yield relative indices of flow and volume, whose scale depends on the signal normalization and on the estimated AIF amplitude.

Recovering $I(t)$ from Equation (2.5) is a mathematically ill-posed problem. In practice, the convolution integral is discretized into a matrix equation, and this matrix is poorly conditioned: even small amounts of noise in the measured tissue dynamic curve, or pseudo-concentration curve in the semi-quantitative setting, can be amplified into large errors in the estimated $I(t)$. The standard solution proposed by Østergaard *et al.* [14] relies on singular value decomposition (SVD) of the convolution matrix, together with a regularization step that filters out the smallest singular values. In this work, we adopt the Tikhonov-regularized variant of this scheme [16]. This approach, like the pharmacokinetic one above, relies critically on the AIF: since it enters the deconvolution as the kernel, any inaccuracy in its shape or amplitude directly biases the resulting perfusion-related indices and, when absolute quantification is performed, the corresponding CBF, CBV and MTT maps.

2.4 Methods for AIF estimation: the state of the art

In view of the central role of the AIF and of the practical difficulties summarized in Section 2.2.2, many studies have been devoted to its estimation. The proposed approaches can be broadly grouped into four families, reviewed in the following subsections [6].

2.4.1 Population-averaged AIFs

One widely used strategy consists of replacing the individual AIF with a single parametric curve derived from measurements across a reference population. The reference model proposed by Parker *et al.* [7] describes the AIF as the sum of two Gaussian functions, representing the first-pass bolus and a recirculation peak, plus a modulated sigmoid that captures the slow wash-out (Figure 2.1).

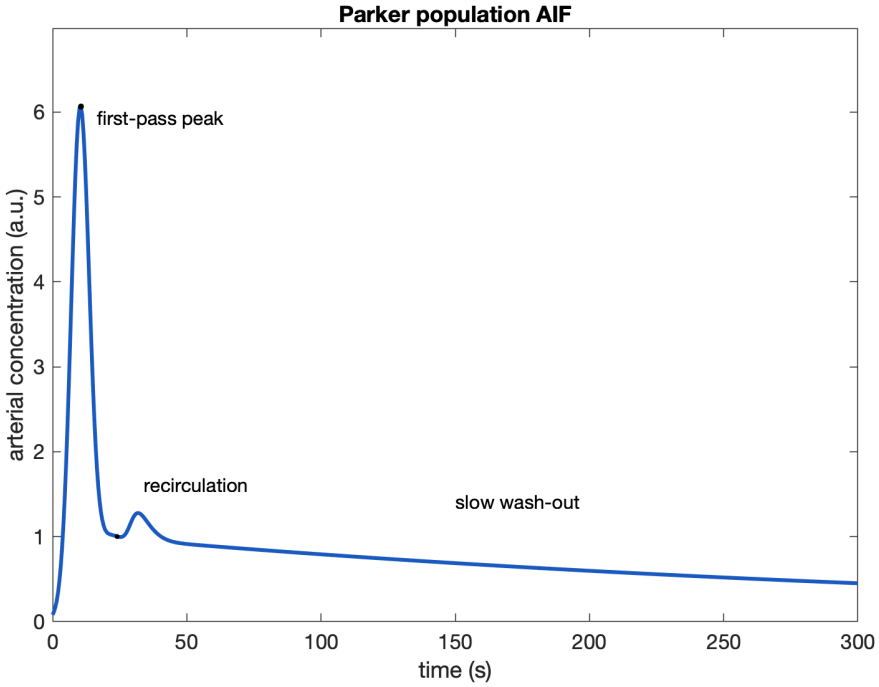


Figure 2.1: The population-averaged arterial input function of Parker *et al.* [7], used in this work as the reference arterial shape. It displays the characteristic features of an arterial input: a sharp, high first-pass peak, a lower recirculation peak, and a slow wash-out. The amplitude is in arbitrary units.

The main advantage of population AIFs is practical convenience: no arterial voxel needs to be identified in the image, the curve is smooth and noise-free by construction, and the approach is immediately reproducible across studies. The principal limitation is that a single average curve cannot capture the inter-subject variability arising from differences in cardiac output, body weight, injection rate, hematocrit and vascular geometry. When applied to an individual patient, a population AIF may be shifted in time, scaled in

amplitude or differently shaped relative to the true input, introducing systematic errors in all derived parameters [6]. For this reason, population AIFs are most useful as a reference template against which patient-specific AIFs are quality-checked, a role they also play in the validation framework of this thesis.

2.4.2 Manual and semi-automatic ROI-based AIFs

A second strategy is to place a manual region of interest (ROI) directly on a visible artery and to use the average signal-time curve of the selected voxels as the AIF. This approach is patient-specific and physiologically direct, and has therefore been widely used in clinical and research settings.

Its main limitation is operator dependency: the choice of artery, imaging slice and precise voxel set introduces inter- and intra-rater variability that is difficult to control. A further source of error is partial-volume contamination: cerebral arteries are small relative to the typical voxel size of a DCE-MRI acquisition, so the signal of an "arterial" voxel almost inevitably includes contributions from surrounding tissue, causing the AIF amplitude to be systematically underestimated. Finally, if the selected ROI includes venous voxels, which can be hard to distinguish from arterial ones on magnitude images, the estimated curve is delayed and widened relative to the true arterial input, biasing the subsequent deconvolution. Semi-automatic variants, in which the user places a single seed point and an algorithm selects the most arterial neighboring voxels, reduce operator workload, but do not resolve these fundamental limitations [6].

2.4.3 Criteria-based automatic voxel selection

A natural evolution of the manual ROI approach is to scan the entire image volume automatically and retain only those voxels whose signal-time curve satisfies a set of physiological criteria characteristic of arterial blood [6]. Commonly adopted criteria include: early bolus arrival time, since the contrast agent reaches the arteries before any tissue; high peak signal enhancement; rapid upslope during wash-in; narrow bolus width; and exclusion of curves that are too noisy or that arrive too late (which would suggest venous rather than arterial origin). The AIF is then estimated as the average of all voxels that pass the filters.

This approach removes the need for manual interaction and surveys the whole brain rather than a single pre-selected location. However, the selection thresholds are heuristic and typically dataset-specific: a threshold that works well for one acquisition protocol or patient population may fail on another. Furthermore, satisfying these criteria does not guarantee that a voxel is *purely* arterial; in fact, partial-volume mixing can still be

present. Despite these limitations, physiological criteria of this kind remain important in practice, and are incorporated as plausibility filters in the automated pipeline developed in this thesis, as described in Chapter 3.

2.4.4 Data-driven methods: clustering, ICA and PCA

Data-driven methods go one step further by exploiting the structure of the entire DCE-MRI dataset to separate arterial-like signals from tissue signals, reducing the dependence on manually defined ROIs and fixed selection thresholds. The three paradigms most relevant to this work are clustering, independent component analysis (ICA) and principal component analysis (PCA).

Clustering methods group all voxel time courses according to their mutual similarity and then identify which cluster best represents arterial dynamics. Both k-means clustering [8] and hierarchical clustering [9] have been applied to this problem. These approaches are conceptually simple and interpretable, but their result depends on the number of clusters, the choice of similarity features and the algorithm initialization. They also assume that arterial voxels form a well-separated group in the feature space, an assumption that may not hold when partial-volume mixing is severe.

Independent component analysis (ICA) decomposes the dataset into statistically independent temporal sources [10]. In principle, the arterial signal should emerge as one of these sources, since it is driven by a physiological process different from the tissue curves. ICA can be more flexible than clustering, but the decomposition is sensitive to initialization, the number of components usually has to be set in advance, and an additional selection step is still required to identify which component corresponds to the AIF.

Principal component analysis (PCA) [11] decomposes the data into orthogonal components ordered by the amount of variance they explain. Because the arterial bolus produces a sharp and high-amplitude temporal enhancement, it is expected to contribute substantially to the variance of the DCE-MRI data set and therefore may be represented within the first few principal components. The component most correlated with an expected arterial profile is then selected, and the voxels that project the most strongly onto it are used to estimate the AIF. PCA is computationally simple and does not require iterative optimization. Its main limitation is that the components maximize explained variance rather than physiological arterial specificity. Additional selection criteria are therefore needed both to choose the correct component and to identify the most representative voxels among those that contribute to it. The PCA-based approach constitutes the methodological foundation of this thesis and is described in detail in Chapter 3.

2.4.5 Open challenges and motivation of this work

Despite the variety of approaches surveyed above, several unresolved issues remain. The challenge is not simply to obtain *some* estimate of the AIF, but to do so in a way that simultaneously satisfies several requirements that are difficult to satisfy in combination. Specifically, AIF estimation should be *automatic*, without manual interaction; *robust to noise*, given the inherently limited SNR of arterial voxels; *robust to partial-volume effects*, since purely arterial voxels are rarely available at clinical resolution; *physiologically plausible*, so that the estimated curve can be verified against known arterial characteristics; and *validatable*, even in the absence of a clinical ground truth.

Existing methods typically involve trade-offs among these requirements. In particular, validation on data with known ground truth remains limited, because clinical DCE-MRI data do not provide direct access to the true arterial input function. The work described in the remainder of this thesis addresses these gaps by (i) implementing a PCA-based AIF extraction pipeline augmented with explicit physiological plausibility filters, (ii) validating the pipeline on synthetic digital reference objects with known ground truth before applying it to clinical data, and (iii) investigating the specific impact of noise and partial-volume effects on the quality of the estimated AIF.

3

Materials and methods

This chapter describes the data and the computational methods used in this work. Section 3.1 introduces the clinical DCE-MRI dataset, including the acquisition protocol and the patient population. Section 3.2 presents the PCA-based pipeline for the automatic extraction of the arterial input function. Section 3.3 describes the voxel-wise deconvolution procedure used to compute the cerebral perfusion maps.

3.1 Clinical dataset

The pipeline described in this chapter was applied to a clinical DCE-MRI dataset acquired in the context of a study on first-episode psychosis. The dataset is used here solely to evaluate the behavior of the proposed AIF-extraction method on real brain DCE-MRI data; no clinical or group-level inference about the underlying condition is drawn in this work.

3.1.1 Study population

The cohort consists of thirteen subjects, including both healthy controls and patients who experienced a first psychotic episode. The inclusion criterion was an age greater than 18 years. The exclusion criteria were the presence of intellectual disability or of neurological disease. The patients were recruited within six months of the onset of the first psychotic episode. Because the methodological focus of this thesis is the extraction of arterial input function rather than the characterization of the clinical condition, subjects are treated jointly as a single set of brain DCE-MRI acquisitions, without distinction between the two groups.

3.1.2 Image acquisition

Imaging was performed on a 3 T Philips Achieva scanner. The dynamic contrast-enhanced acquisition used a three-dimensional T_1-weighted spoiled gradient-echo sequence with an echo time $TE = 1.8$ ms and a flip angle of 15°. Each volume was acquired with a matrix of 112×112 over 44 slices, giving a reconstructed voxel size of $2.05 \times 2.05 \times 2.5$ mm^3, a field of view of approximately 230×230 mm^2 and nearly whole-brain coverage of about 110 mm.

The dynamic series comprised 70 consecutive volumes acquired with a temporal resolution of approximately 4.9 s per volume, corresponding to a total acquisition time of about 5.8 minutes (for one subject a shorter series of 23 volumes was available).

3.2 PCA-based AIF extraction pipeline

The core methodological contribution of this thesis is an automatic and data-driven pipeline for the extraction of the arterial input function from DCE-MRI data, based on principal component analysis (PCA). The pipeline takes as input a four-dimensional DCE-MRI volume and a binary brain mask, and returns both an estimated AIF and a mask of the voxels identified as arterial. Manual placement of regions of interest is not required, and no assumption is made about the anatomical location of the feeding arteries within the brain.

The method developed here builds on the PCA-based approach originally proposed by Sanz-Requena *et al.* [11], which is described in Section 3.2.1. Based on that reference framework, the present work introduces a number of modifications aimed at adapting the method to brain DCE-MRI, at improving its robustness to artifacts, and at making its output verifiable against known physiological characteristics of arterial curves. These modifications were developed incrementally and led to two successive versions of the pipeline, referred to as *v1* and *v2*, which are introduced in Section 3.2.2 and are described in the remainder of this section.

3.2.1 The reference method

The starting point of this work is the automatic AIF extraction method proposed by Sanz-Requena *et al.* [11], who applied PCA to volumetric DCE-MRI data of the prostate in order to identify, without user interaction, the voxels whose signal exhibits a purely arterial behavior.

In their approach, the dynamic data set is reorganized as a two-dimensional matrix $\mathbf{X}$ of size (voxels × time), in which each row is the time-course of the signal-intensity of one

voxel. PCA decomposes this matrix as $\mathbf{X} = \mathbf{T}\,\mathbf{P}^{\top} + \mathbf{E}$, where the loadings $\mathbf{P}$ are a set of orthogonal temporal patterns ordered by the fraction of variance they explain, the scores $\mathbf{T}$ describe how strongly each voxel projects onto each pattern, and $\mathbf{E}$ is a residual term. The authors observed that, when PCA is applied to a whole DCE-MRI volume, the largest sources of variance are typically associated with the arteries, with normally enhancing tissue, and with highly vascularized regions. The arterial pattern, although it explains a substantial fraction of the variance, does not occupy a fixed position in the ordering, but in their experience it always appeared within the first three principal components. To identify which component corresponds to the arterial dynamics, each loading is correlated with a reference arterial curve (in their case a population-averaged AIF), after aligning the two curves in time by matching their peaks; the component yielding the highest correlation is selected as the arterial one. The purest arterial voxels are then defined as those whose signal variance is explained for at least 95 % by the selected component, and the final AIF is obtained by averaging the time-courses of these voxels. Some practical choices of that work are worth noting, because they are revisited in the present implementation: the analysis is performed on signal-intensity curves converted to contrast-agent concentration through a dedicated T_1-mapping acquisition, and the two outermost slices of the volume are discarded, since in three-dimensional spoiled gradient-echo acquisitions they are the most affected by inflow and edge artifacts.

The method was validated on 65 prostate examinations and was shown to yield individual AIFs with lower variability than manual selection and with better discriminative power than a population-averaged curve, motivating its adoption as the methodological backbone of this thesis.

3.2.2 Overview of the proposed pipeline

The pipeline developed in this thesis retains the PCA backbone of the reference method but reorganizes it into two conceptually distinct macro-stages. In the first stage, PCA is used to identify the principal component that best represents the arterial dynamics and, through it, an initial set of candidate arterial voxels. In the second stage, this candidate set is refined by a series of voxel-level physiological criteria, so that the final AIF is computed only from voxels that are both statistically dominant in the arterial component and consistent with the expected shape of an arterial curve. This separation reflects one of the open challenges discussed in Chapter 2: PCA maximizes explained variance rather than arterial specificity, and therefore benefits from being combined with criteria-based filtering.

Compared to the reference method, the proposed pipeline introduces several adaptations, motivated both by the different anatomical setting (brain rather than prostate) and

by the absence of a T_1-mapping acquisition in the available dataset:

- the analysis is performed on *pseudo-concentration* curves $\Delta S/S_0$ rather than on absolute contrast-agent concentration, since no T_1 map is available (Section 3.2.3);
- *no slice exclusion* is applied, so that the whole brain volume contributes to the decomposition;
- the population AIF of Parker *et al.* [7] is used as the reference curve for component identification;
- the single 95 % variance criterion is replaced by a *two-stage percentile selection* that combines the arterial energy fraction with the spatial score (Section 3.2.6);
- an additional *voxel-level physiological refinement* is applied to the selected voxels, and the final AIF is estimated only from the voxels that survive this refinement (Sections 3.2.7 and 3.2.8).

The development of the pipeline proceeded incrementally. Version *v1* implements the adapted PCA-based framework and identifies the arterial component using the same basic criterion as the reference method, namely the highest correlation with the reference AIF among the first few principal components. During preliminary testing on clinical data, this criterion proved insufficient in some cases affected by strong artifacts, where a non-arterial component could obtain a high correlation and be incorrectly selected.

The second version, *v2*, was therefore developed to make the identification of the arterial component more robust. It searches over a larger number of components, applies a set of component-level physiological plausibility constraints, ranks the candidates through a composite score, and includes an explicit fail-safe mechanism for cases in which no plausible arterial component is found. The two versions differ only in the identification of the arterial component (Section 3.2.5); all subsequent steps, including voxel selection, physiological refinement and final AIF estimation, are identical. Version v2 is the one used for the clinical analysis presented in Chapter 5. The two versions are compared both on synthetic data, where they prove equivalent (Chapter 4), and on the clinical dataset, where the modifications introduced in v2 make a difference on specific patients (Chapter 5).

The individual steps of the pipeline are described in detail in the following subsections: signal pre-processing and construction of the PCA model (Sections 3.2.3–3.2.4), identification of the arterial component (Section 3.2.5), arterial voxel selection (Section 3.2.6), physiological refinement of the selected set (Section 3.2.7), and final AIF estimation (Section 3.2.8).

3.2.3 Signal pre-processing

As anticipated in Section 3.2.2, the analysis is not performed on the raw signal but on a pseudo-concentration time-course. One reason is that the reference method of Sanz-Requena *et al.* converts the signal intensity into an absolute contrast-agent concentration through a dedicated T_1-mapping acquisition. In its absence, the relative signal enhancement is used as a surrogate of the concentration, following the semi-quantitative approach introduced in Chapter 2. Consequently, the raw DCE-MRI signal $S(v,t)$ in voxel v and time frame t is converted to a pseudo-concentration time-course by normalizing it to the pre-contrast baseline signal:

$$C(v,t) \;=\; \frac{S(v,t)}{S_0(v)} - 1 \;=\; \frac{\Delta S(v,t)}{S_0(v)}, \tag{3.1}$$

where $S_0(v) = \frac{1}{N_{\text{bl}}} \sum_{t=1}^{N_{\text{bl}}} S(v,t)$ is the mean signal over the $N_{\text{bl}} = 10$ pre-contrast baseline frames. This normalization removes voxel-to-voxel differences in baseline signal intensity, which are not related to the contrast dynamics, and produces a dimensionless quantity that is approximately proportional to the local contrast-agent concentration under the linear approximation $\Delta S/S_0 \propto C(t)$. Since the subsequent AIF extraction is based on positive enhancement patterns, the residual negative values are set to zero. These values mainly reflect fluctuations around the baseline, noise or preprocessing imperfections. The implementation of this pre-processing step is reported in Listing A.1 of Appendix A.

The pre-processed time-courses of all the voxels within the mask are then stacked into a data matrix $\mathbf{X} \in \mathbb{R}^{N_v \times N_t}$, where N_v is the number of brain voxels and N_t is the number of time frames, so that each row of $\mathbf{X}$ is the pseudo-concentration curve of one voxel. This is the same arrangement adopted in the reference method, with the only difference that the entries are pseudo-concentration values rather than signal intensities, and that no slice is excluded from the volume. Before decomposition, the matrix is mean-centered column-wise:

$$\mathbf{X}_c \;=\; \mathbf{X} - \mathbf{1}\,\bar{\mathbf{x}}^{\top}, \tag{3.2}$$

where $\bar{\mathbf{x}} \in \mathbb{R}^{N_t}$ is the mean pseudo-concentration curve averaged over all voxels and $\mathbf{1}$ is a column vector of ones. Subtracting the mean curve leaves, for each voxel, only its deviation from the average dynamics. The decomposition therefore operates on the variance structure of the dataset rather than on the signal level common to all voxels. This step is important because arterial voxels are expected to differ from tissue voxels mainly in the temporal shape, amplitude and timing of their enhancement.

3.2.4 PCA decomposition

The purpose of the PCA decomposition is to re-express the N_v individual voxel curves, which are highly redundant because many voxels share similar dynamics, in terms of a small number of representative temporal patterns. Formally, the centered data matrix is written, as already introduced for the reference method in Section 3.2.1, as

$$\mathbf{X}_c \;=\; \mathbf{T}\,\mathbf{P}^{\top} + \mathbf{E}, \tag{3.3}$$

where the columns of $\mathbf{P}$ are the *loadings*, the columns of $\mathbf{T}$ are the *scores*, and $\mathbf{E}$ is the residual left unexplained when only a subset of components is retained. Each loading $\mathbf{p}_i \in \mathbb{R}^{N_t}$ is a temporal pattern, that is a characteristic time-course shared by a group of voxels, while the corresponding score $\mathbf{t}_i \in \mathbb{R}^{N_v}$ is a spatial map that quantifies how strongly, and with which sign, each voxel follows that pattern. A single voxel curve is thus approximated as a weighted sum of the loadings, the weights being the scores of that voxel.

In practice, the decomposition is computed through the singular value decomposition (SVD) of the centered matrix,

$$\mathbf{X}_c \;=\; \mathbf{U}\,\mathbf{S}\,\mathbf{P}^{\top}, \tag{3.4}$$

where $\mathbf{P}$ holds the loadings as its columns, $\mathbf{S}$ is the diagonal matrix of the singular values σ_i, and the scores are recovered as $\mathbf{T} = \mathbf{U}\,\mathbf{S}$. The SVD orders the components by decreasing singular value, and the fraction of the total variance explained by the i-th component is $\sigma_i^2 / \sum_j \sigma_j^2$. The leading components therefore describe the dominant and most coherent sources of temporal variation in the dataset, while the trailing ones account for progressively smaller and noisier contributions.

This ordering is what makes the arterial signal identifiable. The arrival of the contrast bolus produces a sharp, high-amplitude enhancement that is temporally coherent across all the arterial voxels (Figure 3.1), and therefore contributes a large amount of variance to the dataset; as a result, the arterial dynamics are expected to be represented by one of the first components.

However, as noted in the reference method, the exact position of the arterial component in the ordering is not fixed, because it depends on the relative amount of arterial, tissue and artefactual variance in each specific acquisition. The number of leading components retained as candidates for the subsequent identification step is the first point at which the two versions of the pipeline differ. Version v1 follows the reference method and considers only the first three components, whereas version v2 extends the search to the first $N_{\mathrm{PC}} = 8$ components. The motivation for this extension is that, while Sanz-Requena *et al.* reported that the arterial component always appeared within the first three components in their prostate data, brain acquisitions are more affected by motion

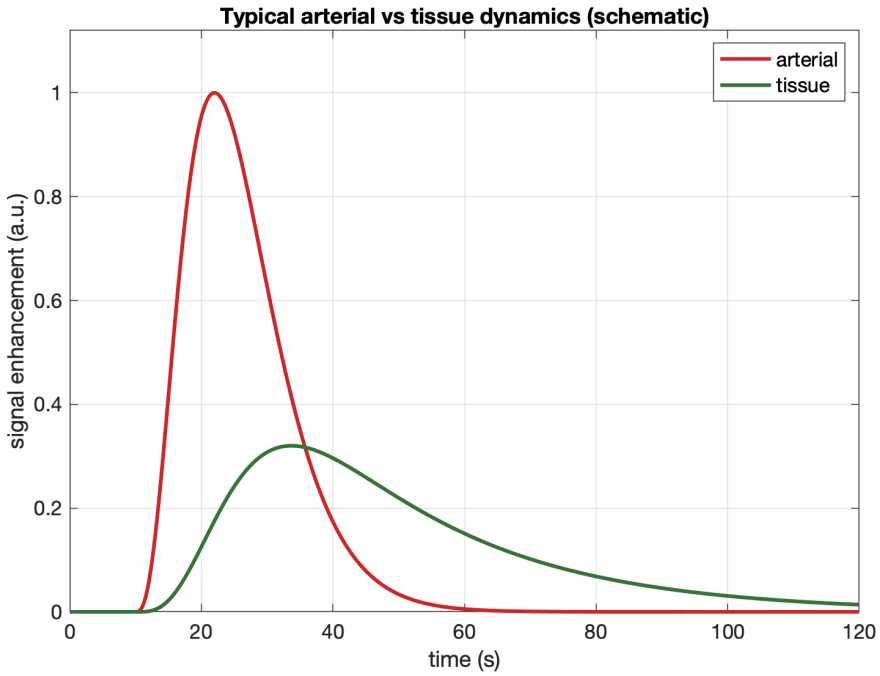


Figure 3.1: Schematic comparison of a typical arterial signal-time curve (sharp, early and high-amplitude) and a typical tissue curve (delayed, lower and broader). The strong and temporally coherent arterial enhancement contributes a large share of the dataset variance, which is what makes the arterial dynamics identifiable by PCA. The curves are illustrative and the amplitudes are in arbitrary units.

and inflow artifacts, which can introduce additional high-variance components and push the arterial one further down the ordering. Retaining eight candidates provides a safety margin against this effect at no practical cost, since the identification step described next discards the non-arterial candidates.

3.2.5 Identification of the arterial component

Among the candidate components retained by the PCA decomposition, the one that best represents the arterial dynamics is identified by comparing each loading $\mathbf{p}_i$ with a population reference AIF. As in the reference method, the comparison uses a curve of known arterial shape as *a priori* knowledge; here this role is played by the population-averaged AIF proposed by Parker *et al.* [7] (Figure 2.1), evaluated on the acquisition time grid with an assumed bolus arrival time of 10 s. The analytical form used to evaluate this reference curve is reported in Listing A.2 of Appendix A. Both the candidate loading and the reference AIF are z-score normalized before the comparison, so that the criterion measures the similarity of the *shape* of the two curves and is insensitive to their amplitude.

The comparison must resolve two ambiguities that are intrinsic to PCA. First, the sign of a principal component is arbitrary, because $\mathbf{p}_i$ and $-\mathbf{p}_i$ describe the same axis

of variance. Therefore, each component is tested at both polarities, and the polarity matching the positive enhancement of the reference is retained. Second, the bolus arrival time of the data differs substantially from the nominal timing of the reference curve: the Parker template has its first-pass peak at about 10 s, whereas in the clinical acquisitions the contrast agent is injected only after the pre-contrast baseline (the first $N_{\mathrm{bl}} = 10$ frames, of the order of 50 s), so that the measured arterial enhancement peaks several tens of seconds later than the template. Therefore, each candidate loading is shifted in time by the lag k^* that aligns its peak with that of the reference. The search for k^* is restricted to a window of $\pm 15\%$ of N_t (of the order of ± 50 s for the present acquisitions), wide enough to bridge this baseline-induced offset while still excluding non-physical alignments to the late wash-out. After sign selection and peak alignment, the Pearson correlation coefficient r_i between the shifted loading and the reference is computed and used as the primary measure of arterial similarity. This framework is shared by both versions of the pipeline.

In version v1, the identification ends at this point: the candidate with the highest correlation r_i is selected as the arterial component, exactly as in the reference method. This criterion is simple and performs well on good-quality data. However, because it relies on correlation alone, a non-arterial component can occasionally obtain a high correlation and therefore be incorrectly selected by v1; an example of this failure on clinical data is shown in Chapter 5 (Section 5.1.3).

Version v2 was introduced specifically to reduce this risk by augmenting the component-selection step with a measure of physiological specificity, a set of component-level plausibility constraints, and a fail-safe mechanism. First, a *tail penalty* is computed on the sign-corrected and peak-normalized candidate loading. This quantity is designed to penalize components that remain elevated in the late part of the acquisition, as expected for tissue-like or artifactual curves, rather than returning towards baseline after the first pass:

$$\tau_i \;=\; \frac{\max\left(0,\; \dfrac{1}{N_{\mathrm{tail}}} \displaystyle\sum_{t \,\geq\, 0.7\,N_t} \tilde{p}_i(t)\right)}{\max_t \, \tilde{p}_i(t)}, \tag{3.5}$$

where $\tilde{p}_i(t)$ denotes the sign-corrected candidate loading after peak normalization, and N_{tail} is the number of frames in the late window, taken here as the last 30 % of the acquisition. A component with a high and sustained plateau has τ_i close to one, whereas a component that washes out has τ_i close to zero. The correlation and the tail penalty are then combined into a single *composite score*,

$$s_i \;=\; r_i\,(1 - \tau_i), \tag{3.6}$$

which rewards components that are simultaneously well correlated with the reference and characterized by a clean wash-out. In v2, this composite score replaces the simple

correlation as the ranking criterion.

Second, a candidate is admitted as a plausible arterial component only if it simultaneously satisfies three physiological constraints:

1. $r_i \geq R_{\min} = 0.6$, to discard components that are too weakly correlated with the reference shape;

2. $k^* > 0$, so that the components that peak before the reference curve are discarded. This constraint was introduced as a conservative rule to avoid selecting early artifactual components whose peak precedes the expected arterial enhancement.

3. the peak of $\mathbf{p}_i$ must fall within the interval [10%, 70%] of N_t, so that components peaking in the very first or in the last frames of the acquisition are rejected.

Rather than being adopted from a specific published method, these constraints translate to the component level the physiological principles common to the criteria-based AIF-selection family reviewed in Section 2.4.3: plausible bolus timing, similarity to a known arterial shape, and a clean first pass [6]. The threshold values themselves ($R_{\min} = 0.6$, the sign of the lag, and the $[10\%, 70\%]$ peak window) were not taken from the literature but were set conservatively on the present data. Therefore, a candidate component is not selected solely because it explains the variance or correlates with the reference AIF, but also because its timing and shape are compatible with an arterial first-pass curve.

Finally, among the components that satisfy all three constraints, the one with the highest composite score s_i is selected as the arterial component. If no candidate is plausible, v2 does not force a choice: it flags the dataset as a *fail-safe* case and returns no AIF estimate. This behavior, which has no counterpart in v1, prevents the pipeline from producing a spurious AIF on datasets in which none of the components genuinely corresponds to arterial dynamics, typically due to severe artifacts.

3.2.6 Arterial voxel selection

Once the arterial component has been identified, the voxels that best represent the arterial dynamics are localized. This step follows the same idea as the reference method, in which the purest arterial voxels are those whose signal variance is largely explained by the selected component, but it replaces the single fixed threshold of that work with a two-stage percentile selection. The procedure is identical in both versions of the pipeline.

Let $\mathbf{p}^*$ denote the loading of the selected arterial component and $\mathbf{t}^*$ its spatial score map (the corresponding column of $\mathbf{T}$). For each voxel v, the part of its time-course attributable to the arterial component is the rank-one reconstruction

$$\hat{\mathbf{x}}_k(v) \;=\; t^*_v \cdot (\mathbf{p}^*)^\top, \tag{3.7}$$

and the *arterial energy fraction* is the proportion of the total variance of the voxel that this term explains,

$$f_k(v) \;=\; \frac{\|\hat{\mathbf{x}}_k(v)\|^2}{\|\mathbf{x}_c(v)\|^2}, \tag{3.8}$$

where $\mathbf{x}_c(v)$ is the mean-centered time-course of voxel v. The quantity $f_k(v)$ is the same measure used by the reference method: it equals one for a voxel whose dynamics are entirely described by the arterial component, and approaches zero for a voxel dominated by other sources of variance.

Whereas Sanz-Requena *et al.* retain every voxel whose energy fraction exceeds an absolute value of 95 %, the present pipeline uses a *two-stage relative* selection that combines a relative and an absolute criterion:

1. **Stage 1 — energy-fraction threshold.** Only the voxels with $f_k(v) \geq \pi_{95}(f_k)$ are retained, where π_{95} is the 95th percentile of f_k over all brain voxels; this keeps the top 5 % of voxels by arterial energy dominance. This stage identifies voxels whose dynamics are *predominantly* arterial, but f_k is a purely relative measure: it is insensitive both to the polarity of the projection and to the absolute strength of the arterial signal.

2. **Stage 2 — spatial-score threshold.** The score t_v^* of the selected component (the corresponding entry of $\mathbf{T}$, introduced in Section 3.2.4) measures how strongly and with which sign voxel v expresses the arterial temporal pattern $\mathbf{p}^*$. Among the stage-1 voxels, only those with a positive score ($t_v^* > 0$) — i.e. voxels that enhance in phase with the arterial pattern rather than against it — are kept, and among these only the 10 % with the highest score ($t_v^* \geq \pi_{90}(t^*)$) are retained.

The choice of percentiles rather than an absolute cut-off is dictated by the semi-quantitative setting: in brain DCE-MRI without T_1 mapping, the energy fraction of the candidate voxels rarely reaches the 95 % level reported for prostate arteries, so a relative threshold is needed to adapt to the actual distribution of each dataset. The second stage, acting on the signed spatial score, therefore selects — among the predominantly-arterial voxels of stage 1 — those that express the arterial dynamics most strongly and with the correct polarity, removing voxels that match the arterial *shape* but with the opposite polarity or a weak amplitude, and thus further increasing the purity of the retained set.

The voxels surviving both stages form the *candidate* arterial mask. From this set, the PCA stage also produces a first estimate of the AIF, computed as the weighted mean of the candidate time-courses with the energy fractions $f_k(v)$ used as weights, so that the most arterial voxels contribute the most. The complete implementation of this first stage — PCA decomposition, identification of the arterial component and two-stage voxel

selection — is reported in Listing A.3 of Appendix A. This weighted estimate is the AIF used directly when the pipeline is run without the physiological refinement described below, as is the case in the synthetic validation of Chapter 4. In the clinical pipeline, instead, the candidate mask is passed to the refinement stage of Section 3.2.7, and the final AIF is estimated from the refined voxels (Section 3.2.8).

3.2.7 Physiological refinement of the selected voxels

In the clinical pipeline, the candidate arterial mask produced by the PCA stage is passed to a voxel-level refinement stage based on explicit physiological criteria. The rationale, anticipated in Section 3.2.2, is that belonging to the arterial principal component guarantees temporal *similarity* to an arterial curve, but does not guarantee by itself that the individual voxel curve has the characteristic shape of an arterial first pass. This voxel-level refinement stage is common to both versions of the pipeline and is therefore distinct from the component-level plausibility constraints introduced in v2.

For each candidate voxel, the baseline-subtracted enhancement $\Delta S(v,t) = S(v,t) - S_0(v)$ is computed, and four descriptors of its first-pass shape are extracted:

- the *peak amplitude* and the *time-to-peak* (TTP), the frame at which the enhancement is maximal;
- the *arrival time*, defined as the first frame in which ΔS exceeds 5 % of its peak value;
- the *upslope*, the mean rate of increase of the enhancement between arrival and peak;
- the *tail ratio*, the ratio between the area under the curve after the peak and the area under the curve before the peak, both evaluated over a first-pass window; a low value denotes a curve that washes out quickly rather than remaining elevated.

A candidate voxel is retained as arterial only if it satisfies, at the same time, all the following conditions: a positive peak amplitude; a time-to-peak no later than the median TTP of the candidates plus three frames, since the bolus reaches all arterial voxels almost simultaneously; an upslope of at least 0.1 enhancement units per frame, denoting a steep wash-in; and a tail ratio below 0.25, denoting a clean wash-out. These criteria translate, at the individual voxel level, the same arterial features used as criteria-based heuristics in the literature reviewed in Chapter 2.

A fall-back rule completes the procedure: if fewer than three voxels satisfy all the criteria, the refinement instead retains the five candidates with the highest peak amplitude. The implementation of the physiological refinement and of the final AIF estimation is

reported in Listing A.4 of Appendix A. In the clinical data analyzed in this work, this fall-back rule was triggered for almost all patients: the candidate sets produced by the PCA stage rarely contained three voxels satisfying all physiological criteria simultaneously. This behavior is expected rather than anomalous and is a direct consequence of the partial-volume problem discussed in Chapters 2 and 4: at clinical spatial resolution, the brain contains few purely arterial voxels. The refinement therefore reduces, in practice, to selecting the few voxels with the strongest enhancement among the PCA candidates. The consequences of this observation on the quality of the estimated AIF are discussed together with the clinical results in Chapter 5.

3.2.8 Final AIF estimation

The first AIF representation is computed as the simple average of the baseline-subtracted enhancement curves of the voxels retained by the physiological refinement. This curve is referred to as the *raw AIF estimate*:

$$\hat{C}_{\mathrm{AIF}}^{\mathrm{raw}}(t) \;=\; \frac{1}{|\mathcal{V}|}\sum_{v\in\mathcal{V}} \Delta S(v,t), \tag{3.9}$$

where $\mathcal{V}$ is the set of refined arterial voxels. Averaging over a small set of strongly enhancing voxels suppresses the voxel-wise noise while preserving the shape of the arterial curve.

Two additional representations of the AIF are derived from this raw estimate. The first is a lightly smoothed curve, obtained with a three-point moving average, which removes the residual high-frequency noise without altering the position or the width of the first-pass peak. The second is a gamma-variate model fitted to the first pass,

$$g(t) \;=\; A\,(t-t_0)^{\alpha}\,\exp\!\left(-\frac{t-t_0}{\beta}\right), \qquad t > t_0, \tag{3.10}$$

where A is an amplitude factor, t_0 the bolus arrival time, and α and β are shape parameters governing the rise and the decay of the curve; $g(t) = 0$ for $t \le t_0$. These four parameters are estimated by non-linear least squares on the first-pass portion of the raw AIF, and the smoothed curve is used as a fall-back when the fit does not converge. The gamma-variate provides a noise-free, analytically smooth description of the first pass, but by construction it decays to zero and therefore does not reproduce the recirculation and the slow wash-out present in the measured curve.

For the computation of the perfusion maps described in Section 3.3, the smoothed AIF is adopted, as it offers the best compromise between noise reduction and fidelity to the measured dynamics: unlike the raw curve, it is not affected by single-frame noise spikes, and unlike the gamma-variate, it retains the recirculation tail, whose omission would otherwise introduce a systematic mismatch with the tissue curves during the deconvolution.

3.2.9 Summary of the two pipeline versions

As discussed throughout this section, the two versions of the pipeline share the entire processing chain and differ only in the way the arterial principal component is identified (Section 3.2.5). The purpose of v2 is therefore not to modify the downstream AIF estimation strategy, but to make the upstream identification of the arterial component more conservative and less dependent on correlation alone. Table 3.1 collects these differences in a single view. All the remaining steps, namely the signal pre-processing, the PCA decomposition, the two-stage voxel selection, the physiological refinement and the final AIF estimation, are identical in the two versions.

Table 3.1: Differences between the two versions of the AIF extraction pipeline. All steps not listed here are identical.

Aspect	v1	v2
Candidate components	first 3	first 8
Ranking criterion	correlation r_i	composite score $r_i(1-\tau_i)$
Component-level constraints	none	$r_i \geq 0.6$, $k^* > 0$, peak in $[10\%, 70\%]$
Fail-safe	absent	present

Thus, the comparison between v1 and v2 isolates the effect of the arterial-component identification strategy, while keeping the subsequent voxel selection and AIF estimation unchanged.

3.3 Computation of cerebral perfusion maps

Cerebral perfusion maps are computed from the estimated AIF and the voxel-wise tissue pseudo-concentration time-courses using the indicator-dilution framework of Meier and Zierler [13], as described in Section 2.3. The specific implementation follows the SVD-based deconvolution approach of Ostergaard *et al.* [14], with Tikhonov regularization of Calamante *et al.* [16] and a first-pass integration window for the CBV estimate. The implementation of the perfusion-map computation is reported in Listing A.5 of Appendix A.

3.3.1 AIF pre-processing

The AIF produced by the extraction pipeline is expressed in units of raw signal enhancement $\Delta S = S(t) - S_0$, because the physiological refinement of Section 3.2.7 operates on baseline-subtracted curves rather than on the normalized pseudo-concentration used inside the PCA stage. In contrast, tissue curves are expressed as pseudo-concentration

$\Delta S/S_0$ (Section 3.2.3). To bring the arterial input and the tissue curves onto the same scale, which is a prerequisite for an internally consistent deconvolution, the AIF is divided by the mean pre-contrast baseline signal of the arterial mask voxels:

$$\hat{C}^{\mathrm{norm}}_{\mathrm{AIF}}(t) \;=\; \frac{\hat{C}^{\mathrm{raw}}_{\mathrm{AIF}}(t)}{\overline{S}_{0,\mathrm{art}}}, \tag{3.11}$$

where $\overline{S}_{0,\mathrm{art}}$ is the mean of $S_0(v)$ over the voxels of the arterial mask. After this step, the AIF is expressed in the same $\Delta S/S_0$ units as the tissue pseudo-concentration curves.

Because partial-volume contamination causes the pipeline to underestimate the AIF peak [6] (even the most arterial voxels contain some tissue contribution, as confirmed by the synthetic validation of Chapter 4), the normalized AIF is additionally rescaled so that its peak equals a target value:

$$\hat{C}^{\mathrm{sc}}_{\mathrm{AIF}}(t) \;=\; \hat{C}^{\mathrm{norm}}_{\mathrm{AIF}}(t) \;\times\; \frac{C^{\mathrm{target}}_{\mathrm{peak}}}{\max_t \hat{C}^{\mathrm{norm}}_{\mathrm{AIF}}(t)}, \tag{3.12}$$

with $C^{\mathrm{target}}_{\mathrm{peak}} = 5.0$ in $\Delta S/S_0$ units. Correcting the AIF for partial-volume-induced amplitude distortion is an established step in quantitative perfusion MRI [17]: the most common variants rescale the measured AIF against a partial-volume-free reference, typically a venous output function from the sagittal sinus matched in area under the curve, or a reference arterial curve [18, 19]. However, in the present semi-quantitative setting, neither an absolute concentration calibration (no T_1 mapping) nor a venous reference is available, so the AIF amplitude is pragmatically fixed to a single empirical target peak, chosen to bring the extracted amplitude into the range expected for strong arterial enhancement. This is a simplified, target-based form of the peak-scaling approach [19], and only partially compensates for the systematic amplitude underestimation; its consequence for the absolute scale of the perfusion maps is addressed in Section 3.3.4.

3.3.2 SVD deconvolution with Tikhonov regularization

Under the linear indicator-dilution model, the tissue pseudo-concentration in voxel v is related to the AIF through a convolution:

$$C_t(v,t) \;=\; \int_0^t \hat{C}^{\mathrm{sc}}_{\mathrm{AIF}}(\tau)\, I(v, t-\tau)\, \mathrm{d}\tau, \tag{3.13}$$

where $I(v,t) = F(v)\, R(v,t)$ is the flow-scaled residue function, $F(v)$ is the blood flow and $R(v,t)$ is the dimensionless impulse residue function. The tissue density and the haematocrit correction are applied later, at the unit-conversion step (Section 3.3.4). In discretized form, this becomes a matrix equation:

$$\mathbf{c}_t(v) \;=\; \mathbf{H}\,\mathbf{i}(v), \tag{3.14}$$

where $\mathbf{i}(v)$ is the discretized flow-scaled residue function and $\mathbf{H}$ is the lower triangular Toeplitz convolution matrix:

$$H_{ij} \;=\; \begin{cases} \hat{C}^{\mathrm{sc}}_{\mathrm{AIF}}(i-j+1)\cdot\Delta t & \text{if } i \geq j, \\ 0 & \text{otherwise.} \end{cases} \tag{3.15}$$

Recovering the flow-scaled residue function by directly inverting $\mathbf{H}$ is an ill-posed problem: the convolution matrix is severely ill-conditioned, so the noise present in the measured tissue curve $\mathbf{c}_t(v)$ would be strongly amplified by a naive inversion, producing oscillatory and non-physical residue functions. For this reason, $\mathbf{H}$ is inverted using a Tikhonov-regularized SVD pseudo-inverse: the SVD deconvolution framework was introduced for perfusion by Østergaard *et al.* [14], and the Tikhonov-regularized variant adopted here follows Calamante *et al.* [16]. Decomposing $\mathbf{H} = \mathbf{U}_H\,\mathbf{S}_H\,\mathbf{V}_H^{\top}$, the regularized pseudo-inverse is:

$$\mathbf{H}^{+} \;=\; \mathbf{V}_H\,\mathrm{diag}\!\left(\frac{s_j}{s_j^2+\lambda^2}\right)\mathbf{U}_H^{\top}, \tag{3.16}$$

where s_j are the singular values of $\mathbf{H}$ and $\lambda = \lambda_{\mathrm{rel}}\,s_{\max}$ is the regularization threshold, with $\lambda_{\mathrm{rel}} = 0.05$. This corresponds to a light regularization (5% of the largest singular value), empirically set as a fixed value that suppressed the noise-driven oscillations of the estimated residue functions while preserving the first-pass shape, rather than selected by a data-adaptive criterion such as the L-curve or generalized cross-validation [16]. A more principled, voxel-wise choice of the regularization strength is left to future work. This form of Tikhonov regularization damps the contribution of small singular values, which would otherwise amplify noise in the deconvolution. The pseudo-inverse $\mathbf{H}^{+}$ is computed once and applied in parallel to all brain voxels.

3.3.3 Derivation of CBF, CBV, and MTT

The discretized flow-scaled residue function at each voxel is obtained as:

$$\hat{\mathbf{i}}(v) = \mathbf{H}^{+}\mathbf{c}_t(v). \tag{3.17}$$

A non-negativity constraint is applied by setting any negative values to zero: $\hat{i}(v,t) \leftarrow \max(\hat{i}(v,t),\,0)$.

The three perfusion parameters are then extracted according to the Meier-Zierler indicator-dilution relations [13, 14]:

- **Cerebral blood flow (CBF)** is proportional to the peak of the flow-scaled residue function:

$$F(v) \;=\; \max_t \hat{i}(v,t). \tag{3.18}$$

- **Cerebral blood volume (CBV)** is proportional to the integral of the flow-scaled residue function over an early first-pass interval, here defined from the beginning of the curve to $T_{\mathrm{fp}} = 30$ s after the residue-function peak:

$$V(v) \;=\; \int_0^{t_{\mathrm{peak}}(v)+T_{\mathrm{fp}}} \hat{i}(v,\tau)\,\mathrm{d}\tau \;\approx\; \Delta t \sum_{t=1}^{t_{\mathrm{peak}}(v)+N_{\mathrm{fp}}} \hat{i}(v,t), \tag{3.19}$$

  where $N_{\mathrm{fp}} = \lfloor T_{\mathrm{fp}}/\Delta t \rfloor$ and $t_{\mathrm{peak}}(v)$ is the frame index of the peak of $\hat{i}(v,t)$. The first-pass window suppresses the contribution of the regularization tail, which would otherwise inflate the CBV estimate.

- **Mean transit time (MTT)** follows from the central volume theorem:

$$\mathrm{MTT}(v) \;=\; \frac{V(v)}{F(v)}. \tag{3.20}$$

3.3.4 Conversion to physiological units

The quantities $F(v)$ and $V(v)$ derived from the deconvolution are converted to conventional perfusion-map units for visualization and comparison. However, because both the tissue curves and the AIF are pseudo-concentration curves, this conversion should be interpreted as a scaling convention rather than as a fully quantitative calibration. Following the convention adopted in perfusion MRI [15], a scaling factor was included to account for brain tissue density and the plasma fraction of blood:

$$\kappa \;=\; \frac{1 - H_{\mathrm{ct}}}{\rho_{\mathrm{brain}}}, \tag{3.21}$$

where $H_{\mathrm{ct}} = 0.45$ is the large-vessel haematocrit and $\rho_{\mathrm{brain}} = 1.04\ \mathrm{g\,mL^{-1}}$ is the brain tissue density. The resulting maps are reported in conventional CBF and CBV units:

$$\mathrm{CBF} \;=\; \frac{F}{\kappa} \times 60 \times 100 \quad [\mathrm{mL}/(100\,\mathrm{g}\cdot\mathrm{min})], \tag{3.22}$$

$$\mathrm{CBV} \;=\; \frac{V}{\kappa} \times 100 \quad [\mathrm{mL}/100\,\mathrm{g}], \tag{3.23}$$

$$\mathrm{MTT} \;=\; \frac{\mathrm{CBV}}{\mathrm{CBF}/60} \quad [\mathrm{s}]. \tag{3.24}$$

Voxels with non-physical values (negative or non-finite CBF, CBV, or MTT) are excluded and set to NaN in the output maps. Although these expressions report the maps in conventional perfusion units, the resulting values should be interpreted as *semi-quantitative*. Because no T_1 mapping is available, the linear approximation $\Delta S/S_0 \propto C(t)$ replaces the true contrast-agent concentration both in the AIF and in the tissue curves. Moreover, the empirical peak rescaling of Equation (3.12) fixes the AIF amplitude only up to a target

value chosen *a priori*. As a consequence, the absolute scale of CBF, CBV and MTT may be systematically biased, and the maps should be interpreted as relative, intra-subject indices rather than as absolute physiological measurements. This limitation is reflected in the clinical perfusion maps of Chapter 5 and its implications are discussed in Chapter 6.

4

Synthetic validation

This chapter reports on the validation of the PCA-based AIF extraction pipeline on synthetic data. The purpose of this validation is not to reproduce the full complexity of clinical DCE-MRI acquisitions, but to evaluate the behavior of the pipeline under controlled conditions in which the true arterial input and the true arterial voxel locations are known. Section 4.1 describes the digital reference object (DRO) used as a controlled test environment, together with its ground-truth arterial input function. Section 4.2 defines the performance metrics adopted to quantify the fidelity of the AIF shape and the accuracy of the selection of voxels. Section 4.3 presents the noise robustness study, in which Rician noise of increasing intensity is added to the clean DRO to assess how the pipeline behaves under degraded signal quality. Section 4.4 presents the partial-volume (PV) study. In this experiment, arterial voxel curves are progressively mixed with a representative tissue curve by assigning a different arterial fraction to each voxel. This allows the simulation to reproduce the heterogeneous artery–tissue signal mixing that is expected in clinical acquisitions. Section 4.5 discusses the results of both studies and draws the main conclusions about the strengths and limitations of the proposed approach.

4.1 The QIBA digital reference object

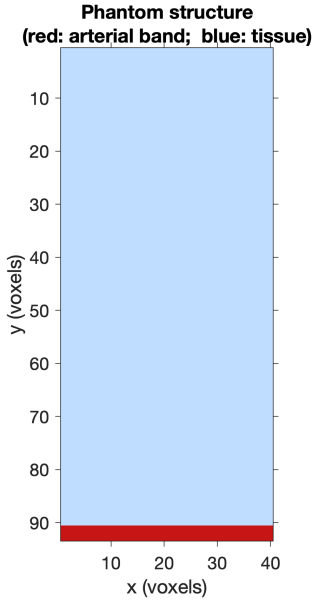
A digital reference object (DRO) is a synthetic dataset for which the ground truth is known exactly by construction. In the context of DCE-MRI, a DRO provides ideal conditions for the validation of the method: noise can be controlled, the arterial input function is known, and the spatial location of every tissue type is explicit. In this work, we used the open-source DCE-MRI DRO developed by Gill *et al.* [20], which implements a configurable DCE-MRI simulator in MATLAB and is distributed together with the arterial input

function defined by the Quantitative Imaging Biomarkers Alliance (QIBA).

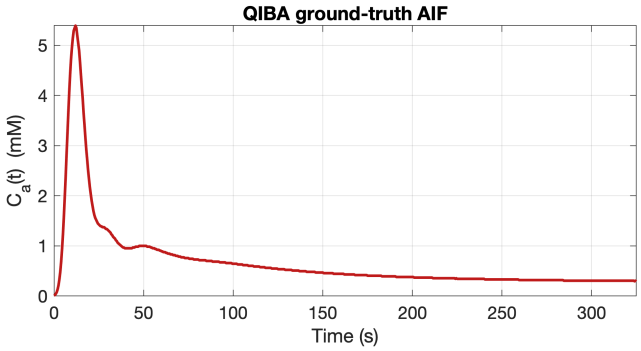
4.1.1 Phantom structure

The phantom consists of a three-dimensional volume divided into three tissue compartments and a dedicated arterial band. The tissue compartments differ in their pharmacokinetic properties, defined by the plasma volume fraction v_p, the extravascular extracellular volume fraction v_e, and the volume transfer constant K^{trans}. The arterial band occupies the outer rows of the image and represents a region of pure intravascular contrast: its signal time-course corresponds to the uncontaminated arterial input function. Therefore, it provides the ground-truth arterial signal used to validate the AIF extracted by the pipeline.

Figure 4.1 summarizes the two elements of the DRO that are used as ground truth in this validation: the spatial arterial mask and the corresponding QIBA AIF.



(a) Spatial layout



(b) Ground-truth AIF

Figure 4.1: Ground-truth information provided by the QIBA DRO. **(a)** Spatial layout of the phantom on a representative slice. The tissue compartments are shown in blue, while the red strip corresponds to the arterial band used as the ground-truth arterial mask. **(b)** Time course of the QIBA ground-truth AIF, characterized by a sharp first-pass peak followed by a recirculation shoulder and a slow washout.

After conversion from the original DICOM series to the NIfTI format used by the pipeline, the phantom has spatial dimensions $40 \times 93 \times 6$ voxels with an in-plane resolution of 1×1 mm^2 and a slice thickness of 10 mm. The time series comprises $N_t = 650$ frames acquired at a temporal resolution of $\Delta t = 0.5$ s, for a total acquisition time of 325 s. The

arterial band spans the last three rows of the image ($y = 91 \ldots 93$), giving a ground-truth arterial mask of exactly $40 \times 3 \times 6 = 720$ voxels.

4.1.2 Ground-truth arterial input function

The reference AIF used in the simulator is the QIBA AIF, distributed as a tabulated curve with 0.5 s temporal resolution. It is characterized by a sharp first-pass peak followed by a lower and broader recirculation peak and a slow washout, a morphology broadly consistent with the population AIF proposed by Parker *et al.* [7], although the timing is adapted to the simulator timeline. In this work, the QIBA AIF is used as ground-truth against which the shape of the PCA-estimated AIF is compared.

4.2 Performance metrics

Two complementary aspects of performance are evaluated for each experiment: (i) the accuracy of the extracted AIF shape, and (ii) the spatial accuracy of the arterial voxel selection.

4.2.1 AIF shape fidelity

The estimated AIF and the QIBA reference AIF are both transformed to z-scores prior to comparison, so that the metric captures shape rather than amplitude. The primary reference is the Pearson correlation coefficient computed after temporal peak alignment:

$$r_{\text{aligned}} \;=\; \operatorname{corr}\big(\hat{c}_{\text{AIF}}^{(k-k^*)},\; c_{\text{GT}}\big), \tag{4.1}$$

where $\hat{c}_{\text{AIF}}^{(k-k^*)}$ denotes the estimated AIF shifted by k^* frames so that its peak aligns with the peak of the ground-truth curve c_{GT}. Peak alignment is necessary because the absolute timing of the bolus arrival differs between the QIBA AIF (constructed on the simulator timeline) and the Parker reference AIF used internally by the pipeline to rank the principal components. Because peak alignment removes the temporal delay between the two curves, r_{aligned} should be interpreted as a measure of shape similarity rather than as a complete measure of AIF accuracy. For this reason, the temporal lag required for alignment is reported separately.

4.2.2 Voxel-selection accuracy

The binary mask composed of arterial voxels returned by the pipeline is compared with the ground-truth arterial mask using the standard confusion-matrix quantities: true positives (TP), false positives (FP), and false negatives (FN). From these, the following

metrics are derived:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \qquad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{4.2}$$

$$\text{Dice} = \frac{2\,\text{TP}}{2\,\text{TP} + \text{FP} + \text{FN}}. \tag{4.3}$$

Precision measures the purity of selection, i.e. the fraction of selected voxels that are genuinely arterial, while recall measures completeness, i.e. the fraction of the true arterial region that is recovered. The Dice coefficient provides a single summary measure that combines both precision and recall.

4.3 Noise robustness study

4.3.1 Experimental setup

Additive zero-mean Rician noise was added independently to each voxel and each time frame of the clean DRO to produce three noisy datasets of increasing intensity. The noise standard deviation σ was set according to the signal-to-noise ratio (SNR) defined with respect to the mean baseline signal of the phantom:

$$\sigma = \frac{\overline{S}_{\text{baseline}}}{\text{SNR}}, \tag{4.4}$$

where $\overline{S}_{\text{baseline}}$ is the mean signal intensity over the first ten frames (pre-bolus) averaged across all phantom voxels. This convention provides an interpretable noise level referenced to the pre-contrast signal and independent of the contrast-enhancement magnitude. Three SNR levels were tested: SNR $\in \{50, 20, 10\}$, corresponding, respectively, to mild, moderate, and severe noise. The original noise-free dataset (SNR $\to \infty$) served as baseline.

Both pipeline variants considered in this validation, denoted as v1 and v2, were evaluated using the same synthetic datasets. Their algorithmic differences are described in detail in Chapter 3.

For each noise level, the entire PCA pipeline was executed with the parameters frozen at the values used for the clinical dataset (Chapter 3): $N_{\text{PC}} = 8$, $R_{\text{min}} = 0.6$, spatial-score percentile = 90, energy-fraction percentile = 95. Because the DRO has shorter acquisition than the clinical scans, the peak-location acceptance window for component selection was extended to [2%, 95%] of the total number of frames to avoid spurious rejections due to the different bolus timing.

4.3.2 Results

Table 4.1 summarizes the performance of the pipeline in the four noise levels.

Table 4.1: Noise robustness study results. The ground-truth arterial mask contains 720 voxels. SNR is defined relative to the mean pre-bolus signal (Eq. 4.4).

Level	SNR	r_{raw}	r_{aligned}	Lag (s)	TP	FP	FN	Precision	Recall	Dice
0	∞	0.746	0.977	2.5	720	0	0	1.000	1.000	1.000
1	50	0.746	0.977	2.5	72	0	648	1.000	0.100	0.182
2	20	0.746	0.977	2.5	72	0	648	1.000	0.100	0.182
3	10	0.746	0.977	2.5	72	0	648	1.000	0.100	0.182

The most striking observation from Table 4.1 is the near-perfect stability of r_{aligned} across all noise levels: the correlation between the extracted AIF and the QIBA reference AIF remains 0.977 even at SNR = 10, where the noise standard deviation equals one tenth of the mean pre-bolus signal used for normalization. This indicates that, in this controlled DRO setting, the shape of the estimated AIF is highly robust to the tested levels of additive Rician noise.

In contrast, the voxel classification metrics change markedly after the introduction of noise. With SNR = ∞ the method selects all 720 ground-truth arterial voxels with no false positives, obtaining Dice = 1. From SNR = 50 onward, only 72 out of 720 arterial voxels are retained (Dice = 0.18, Recall = 0.10), and no further change occurs as the noise increases to SNR = 10. Essentially, precision remains 1 at all noise levels: every selected voxel is genuinely arterial.

Figure 4.2 shows the two key metrics as a function of the noise level. The left panel confirms the negligible variation in AIF correlation; the right panel illustrates the sharp drop in Dice that occurs between the noise-free and the mildest noisy condition.

4.4 Partial volume effect study

4.4.1 Partial volume mixing model

In real DCE-MRI acquisitions, the voxels at the boundary of a blood vessel contain a mixture of arterial blood and surrounding tissue. The measured signal is a weighted average of the two contributions, governed by the arterial volume fraction $f_a \in [0, 1]$:

$$S_{\text{pv}}(v,\, t) \;=\; f_a(v)\, S_{\text{art}}(v,\, t) \;+\; \big(1 - f_a(v)\big)\, S_{\text{tissue}}(t), \tag{4.5}$$

where $S_{\text{art}}(v, t)$ is the pure arterial signal in voxel v and time t, and $S_{\text{tissue}}(t)$ is a tissue reference time-course. This linear mixing model is a simplified approximation of the

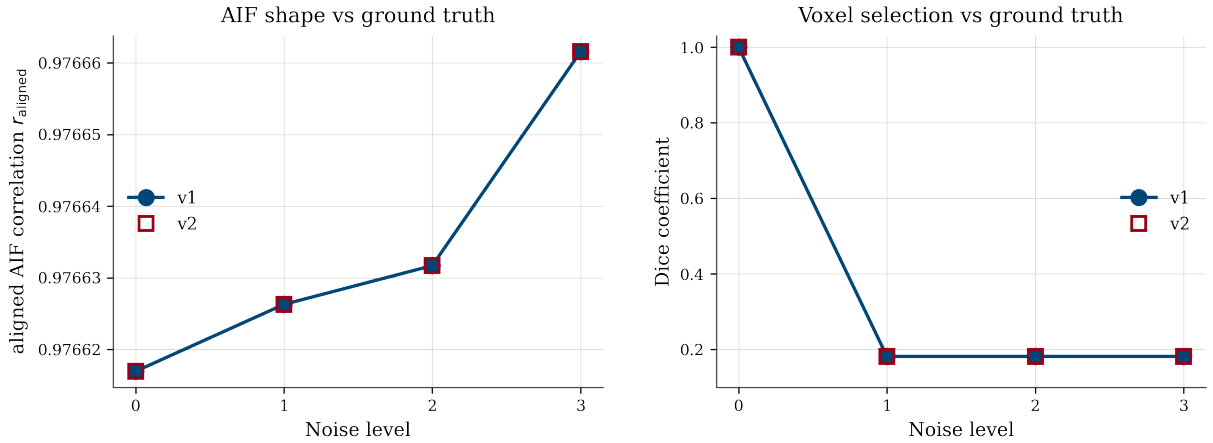


Figure 4.2: Noise robustness study. *Left:* aligned correlation r_{aligned} between the extracted AIF and the QIBA reference as a function of noise level. *Right:* Dice coefficient of the arterial voxel mask relative to the ground truth. For both metrics, the results of the two pipeline variants (v1, blue circles; v2, red squares) overlap exactly at all noise levels.

partial volume effect at the signal-intensity level. It does not explicitly model within-voxel magnetization physics, but provides a controlled way to attenuate the arterial contribution while preserving a known ground-truth arterial region. The tissue reference curve $S_{\text{tissue}}(t)$ was taken as the mean raw signal across all non-arterial phantom voxels, computed from the noise-free DRO and applied consistently in units of raw signal intensity.

A key feature of clinical partial volume contamination is that it is *heterogeneous*: within the same image, voxels close to a vessel carry different proportions of arterial and tissue signal depending on their position relative to the vessel lumen. To reproduce this, and unlike a single spatially uniform contamination level, each arterial voxel was assigned an *independent* arterial fraction $f_a(v)$ drawn at random, so that within an image the arterial band spans a continuous range of contamination from heavily mixed to nearly pure. The mixing was applied only within the arterial band: no arterial or partially arterial signal was introduced into the tissue voxels, so that the spatial location of the ground-truth arterial region was preserved exactly. The relevant question then becomes not whether the pipeline recovers the *entire* arterial band – which would be undesirable, since the heavily contaminated voxels should *not* be labeled as arterial – but whether it selects the *least contaminated* voxels (highest f_a) and rejects the most contaminated ones.

Two contamination regimes were constructed, both with $f_a \in [0.2, 1.0]$, but differing in the shape of the per-voxel distribution:

- **High-heterogeneity regime:** f_a drawn from a uniform distribution, $f_a(v) \sim \mathcal{U}(0.2,\, 1.0)$. Arterial voxels are spread evenly across the entire contamination range, with as many heavily mixed voxels as nearly pure ones (a stress test of the selection).
- **Low-heterogeneity regime:** f_a skewed toward purity, $f_a(v) = f_{a,\max} - (f_{a,\max} -$

$f_{a,\min})\, u^p$ with $u \sim \mathcal{U}(0,1)$ and $p = 2.5$. Most arterial voxels are nearly pure, with only a minority strongly contaminated: a coherent cluster of near-arterial voxels accompanied by a tail of mixed ones.

The two regimes are not claimed to bracket the true clinical contamination, which is unknown; they probe the two qualitatively different situations the selection can face — a population dominated by a coherent pure cluster, and one in which contamination is so dispersed that no such cluster stands out.

4.4.2 Experimental setup

For each regime, a single mixed-PV dataset was generated (fixed random seed for reproducibility) and processed with the same frozen PCA pipeline parameters used in the noise robustness study, so that any performance change could be attributed to the imposed contamination.

Because the goal is no longer to reproduce the entire arterial band, the evaluation is based on quantities that capture the *purity* of the selection rather than its completeness: the number of selected voxels N_{sel}; how many of them are truly arterial (TP) or tissue (FP); the mean arterial fraction $\bar{f}_a^{\text{sel}}$ of the selected arterial voxels; and the aligned AIF correlation r_{aligned}. The Dice and recall metrics used in the noise study are not reported for the partial-volume experiment. In this setting, recovering voxels with a low arterial fraction would increase Dice and recall, but would not represent a desirable outcome for the AIF estimation, since such voxels are strongly contaminated by tissue signal.

Finally, to characterize the realistic clinical condition in which noise and partial volume act together, noise was added on top of each PV dataset using the same noise model and levels as the noise robustness study (Eq. 4.4; SNR $\in \{50, 20, 10\}$, plus the noise-free case SNR $\to \infty$), for both regimes. This combined experiment is the synthetic counterpart of the conditions met in the clinical data of Chapter 5.

4.4.3 Results

Table 4.2 reports the combined noise-and-PV results for both regimes; the noise-free rows (SNR $= \infty$) correspond to the pure partial-volume case.

Partial volume alone (SNR $= \infty$). In the **low-heterogeneity** regime, the pipeline behaves exactly as intended. Of 720 arterial voxels spanning $f_a \in [0.2, 1.0]$, 52 are selected, *all* of them are arterial (FP $= 0$) and all essentially pure: the selected voxels have a mean arterial fraction of 0.999 (minimum 0.998). Figure 4.3(a) shows that the selected voxels fall entirely into the purest bin of the arterial-fraction distribution. Therefore, the

Table 4.2: Combined partial-volume and noise study; results are identical for the two pipeline variants v1 and v2. The ground-truth arterial band contains 720 voxels, and the noise-free rows (SNR = ∞) correspond to the pure partial-volume case. PC: index of the selected principal component; N_{sel}: number of selected voxels; TP/FP: selected voxels that are arterial / tissue; $\bar{f}_a^{\text{sel}}$: mean arterial fraction of the selected arterial voxels (undefined when TP = 0); r_{aligned}: peak-aligned correlation between the extracted AIF and the QIBA reference.

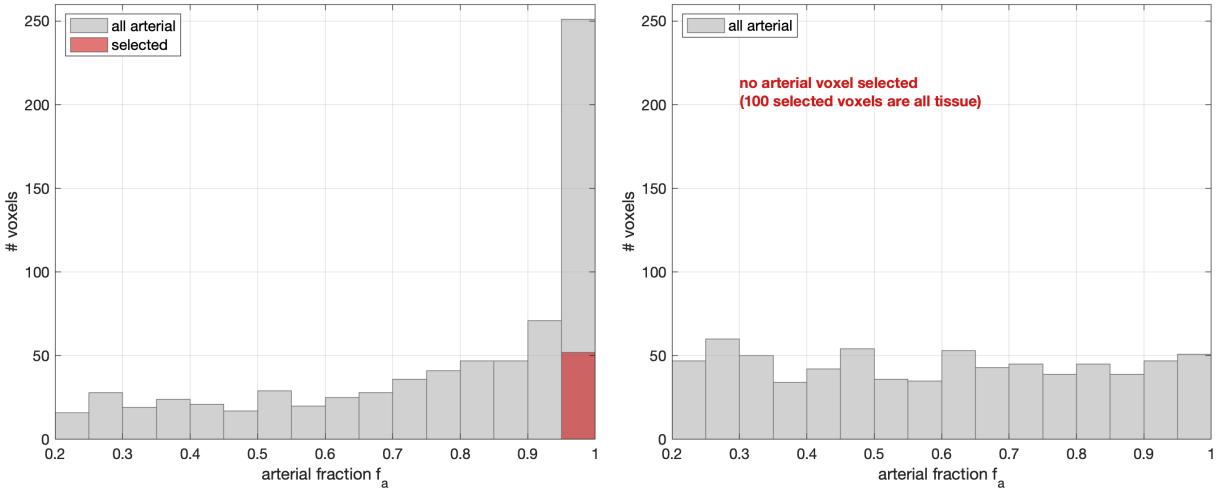
Regime	SNR	PC	N_{sel}	TP	FP	$\bar{f}_a^{\text{sel}}$	r_{aligned}
High heterogeneity	∞	2	100	0	100	—	0.865
	50	2	17	0	17	—	0.869
	20	2	14	0	14	—	0.935
	10	2	10	0	10	—	0.934
Low heterogeneity	∞	1	52	52	0	0.999	0.971
	50	1	50	50	0	0.998	0.971
	20	1	58	58	0	0.993	0.971
	10	1	68	68	0	0.991	0.971

algorithm isolates the least contaminated voxels and discards the rest, and the resulting AIF correlates well with the QIBA reference ($r_{\text{aligned}} = 0.971$).

In the **high-heterogeneity** regime, the outcome is the opposite: none of the arterial voxels is selected (TP = 0) and the 100 selected voxels are all tissue (FP = 100; Figure 4.3(b)). This failure is not caused by a lack of pure voxels — 98 arterial voxels still have $f_a > 0.9$ — but by the way the relative energy-fraction metric behaves when the arterial population is dispersed, as analyzed in Section 4.5.2.

Combined noise and partial volume. Noise addition does not change the qualitative picture (Table 4.2, Figure 4.4). In the **low-heterogeneity** regime, the selection remains fully arterial (FP = 0) for every SNR, with selected voxels remaining nearly pure ($\bar{f}_a^{\text{sel}}$ from 0.999 at SNR = ∞ to 0.991 at SNR = 10) and the AIF correlation unchanged at $r_{\text{aligned}} = 0.971$. The method is thus robust to the simultaneous presence of noise and partial volume, provided that a coherent cluster of near-pure voxels exists.

In the **high-heterogeneity** regime, the selection remains entirely wrong (TP = 0) at every SNR: noise does not rescue a case that already fails without it. A counter-intuitive observation is that the AIF correlation is nonetheless moderate and even *increases* with noise (r_{aligned} from 0.865 at SNR = ∞ to 0.93 at SNR = 10), despite the selection being composed of 100% false positives. This point is discussed in Section 4.5.2, as it carries a methodological warning about validating AIF extraction by shape correlation alone.



(a) Low heterogeneity (b) High heterogeneity

Figure 4.3: Distribution of the arterial fraction f_a over all arterial voxels (gray) and over the voxels selected by the pipeline (red), for the two contamination regimes (pure partial volume, SNR $= \infty$). **(a)** Low heterogeneity: the selection falls entirely in the purest bin ($f_a \approx 1$). **(b)** High heterogeneity: no arterial voxel is selected — the 100 selected voxels are all tissue.

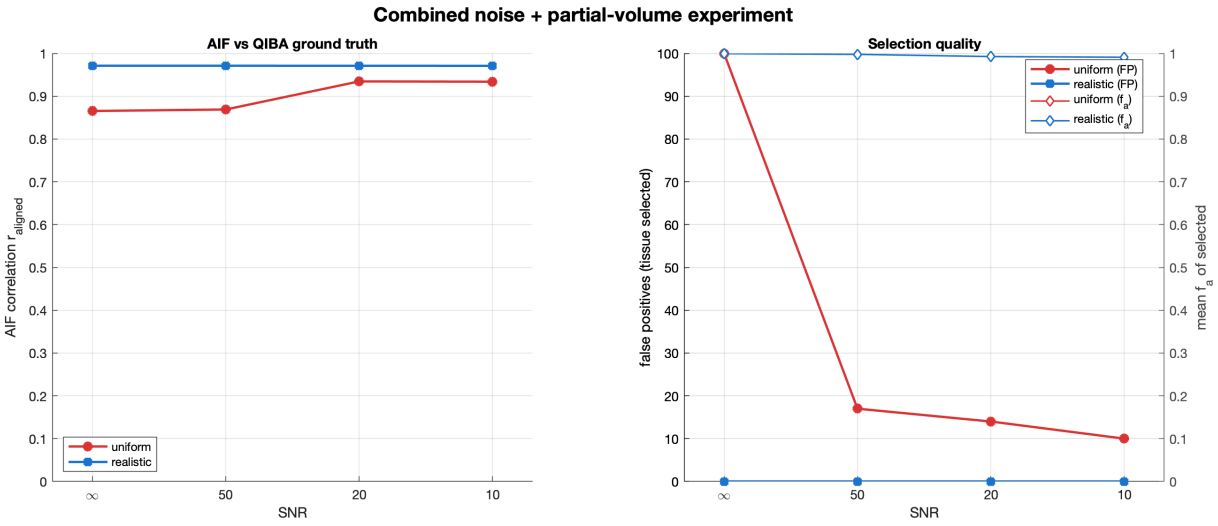


Figure 4.4: Combined noise and partial-volume study, for both contamination regimes (red: high heterogeneity; blue: low heterogeneity) as a function of SNR (∞ = pure partial volume). *Left:* aligned AIF correlation with the QIBA reference. *Right:* number of false positives (solid markers, left axis) and mean arterial fraction of the selected voxels (open markers, right axis). In the low-heterogeneity regime the selection stays purely arterial at all SNR; in the high-heterogeneity regime it is entirely tissue, yet the AIF correlation remains misleadingly high.

4.5 Discussion

4.5.1 Noise robustness

The noise study demonstrates that the shape of the AIF estimated by PCA is remarkably stable under Rician noise perturbations. The aligned correlation r_{aligned} remains essentially unchanged in all SNR levels tested, a result that can be understood from the structure of the PCA decomposition. The arterial band contributes 720 voxels to the data matrix, all of which share the same underlying arterial time-course. In this idealized setting, the principal component that captures this common variance behaves similarly to an average over many realizations of the arterial signal. Under the simplifying assumption of independent zero-mean noise, the noise contribution would be expected to decrease approximately as $1/\sqrt{N_{\text{art}}}$. With $N_{\text{art}} = 720$, this averaging effect helps explain why even the SNR = 10 dataset shows negligible degradation in AIF shape correlation.

The reduction in Dice from 1.00 to 0.18 that occurs between the noise-free and the first noisy condition (SNR = 50) has a different explanation. In the noise-free DRO, all 720 arterial voxels have nearly identical energy-fraction values because they share the same arterial time-course. As a result, the percentile threshold is not selective and the entire arterial band is retained. Once noise is added, the energy-fraction distribution becomes heterogeneous: small voxel-wise differences are introduced in the estimated energy-fraction values, causing only the upper tail of the arterial distribution to survive the 95th-percentile threshold. The result is a conservative selection of the 72 most unambiguously arterial voxels, all genuine (precision = 1). The invariance of both the selected voxel count and the AIF correlation across noise levels 1–3 indicates that the thresholding behavior stabilizes after the transition from the noise-free to the noisy case: within the tested range, additional Rician noise does not further alter which voxels are selected.

4.5.2 Partial volume effect

The partial-volume study indicates that the outcome of the voxel selection is determined by the structure of the contamination: the pipeline correctly retains the least contaminated voxels when a coherent cluster of near-pure voxels is present and fails to identify any arterial voxel when the contamination is heterogeneously distributed. Retaining only the least contaminated voxels and discarding the heavily mixed ones is in fact the desired behavior, since highly contaminated voxels should not be labeled arterial. What the two regimes reveal is the condition under which this behavior holds. That condition, and the difference between the two regimes, is explained by the relative energy-fraction quantity $f_k(v)$ used by the double-threshold voxel selection.

The quantity $f_k(v)$ is the fraction of a voxel's centered temporal variance that is captured by the selected principal component; it is therefore a *relative* measure, normalized to each voxel's own energy, and is maximized when a voxel's time-course is well aligned with that single component.

In the **low-heterogeneity** regime most arterial voxels are nearly pure and thus share essentially the same arterial waveform. They collapse into a single dominant principal component, in which each of them has a high f_k. They consequently dominate the energy-fraction ranking and are the ones selected, while the minority of contaminated voxels, whose mixed waveforms are mutually dissimilar, scatter across components and fall below the threshold. This is why the selection rate is non-zero only for $f_a \in [0.9, 1.0]$ (Figure 4.3(a)): the method isolates the coherent cluster of pure voxels.

In the **high-heterogeneity** regime no such cluster exists. Because f_a is distributed throughout the range, in fact every arterial voxel is a different mixture, and the common arterial dynamics is split into several components — the first two principal components carry 80.3% and 13.4% of the variance, respectively. The component eventually selected (the second, $r = 0.857$ with the Parker reference) accounts for only about 9.5% of each arterial voxel's own energy, so the arterial voxels all have a low and nearly identical $f_k \approx 0.095$. Meanwhile the untouched tissue voxels carry far less temporal energy (their median centered energy is roughly two orders of magnitude smaller than that of the arterial voxels): a low-energy voxel whose small fluctuation happens to align with the selected component reaches f_k values up to 0.94. Since f_k is relative, these degenerate tissue voxels – not the arterial ones – occupy the top of the ranking and are selected.

This also explains why the algorithm still returns 100 voxels instead of declaring failure. The double-threshold selection is purely *percentile-based* (top 5% of f_k, then top 10% of the spatial score): it has no absolute cutoff and therefore always returns a non-empty set, regardless of whether any selected voxel is genuinely arterial. The pipeline's only fail-safe acts earlier, at the component-identification stage — if none of the candidate components satisfies the plausibility constraints, the method returns no AIF — but here a plausible arterial-like component does exist, so the fail-safe is not triggered and the relative threshold proceeds to select the highest f_k voxels, which in this regime are tissue. The absence of an absolute criterion that can reject an implausible *voxel* selection is thus a limitation of the current approach.

A particularly instructive observation concerns the AIF correlation in the high-heterogeneity regime. Although the selection is entirely made up of tissue voxels (TP = 0), the extracted AIF still correlates moderately with the QIBA reference, and the correlation even *increases* with noise (from 0.865 to 0.93). The reason is that the AIF is a weighted average over the selected voxels, which reproduces the shape of the selected component irrespective of which voxels carry it. Adding noise makes the relative threshold even more selective,

retaining the few voxels most tightly aligned with that component and pushing the averaged waveform closer to the arterial shape. This residual — and partly spurious — correlation must not be read as evidence of correct arterial identification. The lesson is methodological: *a good AIF shape correlation does not guarantee a correct voxel selection*, and an automated AIF method cannot be validated by waveform similarity alone; spatial selection metrics and physiological plausibility checks are indispensable.

4.5.3 Implications for clinical application

The synthetic validation defines a clear performance envelope for the proposed pipeline. The noise robustness study demonstrates that, within the tested SNR range, additive Rician noise has a negligible impact on AIF shape accuracy, confirming that PCA-based averaging is an effective denoising strategy. The partial-volume study, in turn, shows that contamination is not in itself a source of error: the pipeline correctly selects the least contaminated voxels and discards the rest. What it requires in order to succeed is the *coexistence*, in the image, of a coherent cluster of near-pure voxels; when such a cluster is present the pipeline isolates it even under noise (low-heterogeneity regime, and the combined experiment of Table 4.2), whereas when contamination is so dispersed that no cluster stands out, the relative energy-fraction metric is captured by spurious low-energy voxels (high-heterogeneity regime). The decisive factor is therefore not the average level of partial volume, but the *heterogeneity* of the contamination and the existence of a pure sub-population.

Finally, the two pipeline variants v1 and v2 produced identical voxel selections and AIF correlations under every synthetic condition tested, all noise levels, both contamination regimes, and their combination. The synthetic validation therefore cannot discriminate between the two variants: the behavior reported in this chapter is a property of the shared PCA-based selection strategy, rather than of a specific implementation. The plausibility filters and fail-safe introduced in v2 remain inactive here, because the selected component in the DRO is already plausible; their effect emerges only on clinical data, where the two variants are compared directly in Chapter 5.

In the conditions tested, the pipeline reliably retains only voxels with $f_a \gtrsim 0.9$ and rejects everything below, which is an appropriate operating point: a voxel that is 90% arterial is a reasonable practical definition of a usable arterial voxel. The open question for real data is whether voxels of such purity exist at all at clinical resolution.

This is precisely where the synthetic envelope meets the clinical reality. In clinical DCE-MRI data, purely arterial voxels are expected to be rare, especially at typical clinical spatial resolutions: even when a major artery is visible, its apparent signal is affected by vessel diameter, voxel size, slice thickness, vessel orientation and motion, leading to

variable degrees of partial-volume contamination. Whether a coherent cluster of near-pure voxels survives these effects — the condition that the present study identifies as necessary for success — is examined directly on the patient data in Chapter 5.

It should be noted that the DRO remains an idealized scenario in several respects. The arterial band has a regular rectangular geometry, and the tissue partner curve is a single averaged time-course rather than the spatially varying parenchymal dynamics of real brain imaging. The heterogeneity of the contamination, although now reproduced within a single image, is modeled by an independent random arterial fraction per voxel rather than by realistic vascular geometry.

A further limitation concerns the evaluation metric itself. Because the AIF curves were compared after z-score normalization and, for r_{aligned}, after peak alignment, the correlation metric primarily assesses waveform similarity. It does not quantify absolute amplitude errors and only indirectly reflects temporal delay, which is reported separately through the alignment lag. As the partial-volume study made explicit, this shape-only correlation can remain high even when the underlying voxel selection is entirely wrong. Future work should therefore complement correlation-based metrics with amplitude- and timing-sensitive measures, such as peak-amplitude error, time-to-peak error, area-under-the-curve error or normalized root mean square error.

5

Clinical results

This chapter presents the application of the PCA-based pipeline to the clinical DCE-MRI dataset described in Chapter 3. It is organized into two parts. Section 5.1 compares the two versions of the pipeline, v1 and v2, directly on the clinical data: the synthetic validation of Chapter 4 showed that the two versions are equivalent in the absence of artifacts, and the purpose here is to examine where and why they differ in real subjects, thereby assessing whether the additional mechanisms introduced in v2 are useful in the clinical setting for which they were designed. Section 5.2 then presents the clinical results obtained with version v2, which is the configuration adopted for the analysis: the extracted arterial input functions, the spatial selection of the arterial voxels, and the cerebral perfusion maps. The limitations of the approach are discussed in Chapter 6.

5.1 Comparison of the two pipeline versions

5.1.1 Component selection across the cohort

Both versions of the pipeline were applied to the thirteen subjects of the dataset. The two versions share the entire processing chain and differ only in the way the arterial principal component is selected (Section 3.2.5): version v1 examines the first three components and selects the one with the highest correlation with the Parker reference AIF, without further constraint, while version v2 extends the search to the first eight components, ranks them through the composite score $r_i(1 - \tau_i)$ subject to a set of physiological plausibility constraints, and declares a fail-safe case when none of the components is plausible. Because the component selection of v1 can be reconstructed exactly from the per-component quantities computed by v2, the two versions can be compared without re-running the pipeline.

Table 5.1 reports, for each subject, the component selected by the two versions and the corresponding correlation with the reference AIF.

Table 5.1: Component selected by the two versions of the pipeline for each subject, with the correlation r with the Parker reference AIF. "Identical" denotes subjects for which the two versions select the same component; for P006, version v2 returns a fail-safe and produces no AIF.

Subject	v1: PC (r)	v2: PC (r)	Outcome
P001	PC2 (0.759)	PC2 (0.759)	identical
P002	PC2 (0.804)	PC2 (0.804)	identical
P003	PC2 (0.810)	PC2 (0.810)	identical
P004	PC2 (0.791)	PC2 (0.791)	identical
P005	PC3 (0.632)	PC3 (0.632)	identical
P006	PC1 (0.339)	—	fail-safe
P007	PC1 (0.502)	PC4 (0.749)	v2 differs
P008	PC2 (0.764)	PC2 (0.764)	identical
P009	PC2 (0.768)	PC7 (0.794)	v2 differs
P010	PC2 (0.838)	PC2 (0.838)	identical
P011	PC1 (0.827)	PC2 (0.736)	v2 differs
P012	PC2 (0.819)	PC2 (0.819)	identical
P013	PC2 (0.804)	PC2 (0.804)	identical

The comparison shows that the two versions select the same component in nine of the twelve participants for which v2 produces an AIF. In these cases, the arterial component is already the most correlated among the first three, and the additional machinery of v2 leaves the result unchanged. The two versions diverge in three subjects (P007, P009 and P011), in addition to the fail-safe case P006. However, in P009, the difference has limited practical impact: both versions select an arterial-like component (PC2 for v1 and PC7 for v2), with similar correlation values and no evident failure of the AIF extraction. For this reason, the following subsections focus on P007, P011 and P006, which each illustrate a distinct mechanism introduced in v2.

5.1.2 A case of agreement

In the majority of subjects the two versions coincide. Subject P010 is a representative example: both versions select the second principal component as the arterial one, with the same correlation ($r = 0.838$, the highest in the cohort), and therefore produce the same AIF. Figure 5.1 shows the selected component against Parker reference AIF: the component reproduces the characteristic arterial profile, with a steep first-pass peak, a lower recirculation peak and a slow wash-out toward the baseline. For subjects of this

kind, in which the arterial dynamics dominate the initial principal components, the extension of the search and the plausibility constraints of v2 are not necessary, and the two versions are interchangeable, consistently with the equivalence observed on synthetic data in Chapter 4.

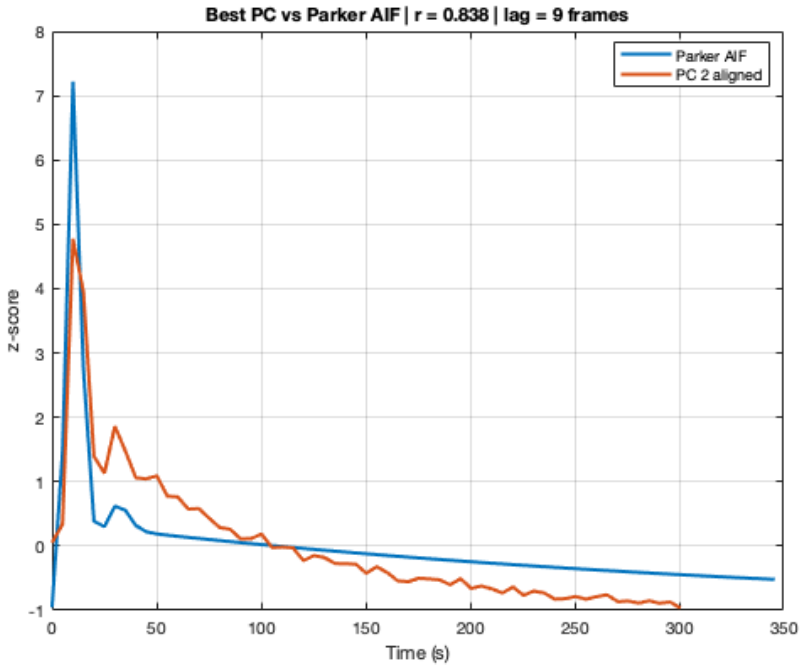


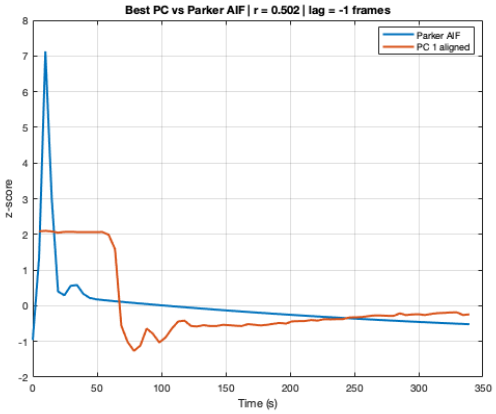
Figure 5.1: Case of agreement (P010): the principal component selected by both versions (PC2, $r = 0.838$) compared with the Parker reference AIF, after sign correction and peak alignment. Both curves are shown as z-scores.

5.1.3 Cases where version v2 changes the outcome

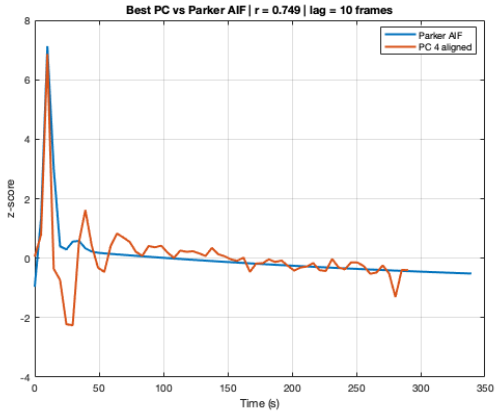
Among the cases of disagreement, P007, P011 and P006 are the most informative because they illustrate three distinct mechanisms introduced in v2.

Extended search and plausibility filtering (P007). In P007 the arterial dynamics are captured by the fourth principal component, beyond the first three examined by v1. Among the first three components, the most correlated is the first, but it is not arterial: its correlation is low ($r = 0.502$) and its peak precedes the reference (a negative alignment lag), so it represents a slow baseline trend rather than a bolus passage. Version v1, which simply selects the most correlated of the first three components, would therefore choose this non-arterial component (Figure 5.2, left). Version v2 instead rejects the first three components, because none of them satisfies the plausibility constraints, and selects the fourth component, which has a clear arterial profile ($r = 0.749$, Figure 5.2, right).

This case shows the combined effect of the extended search and of the component-level plausibility constraints introduced in v2.



(a) v1: selects PC1

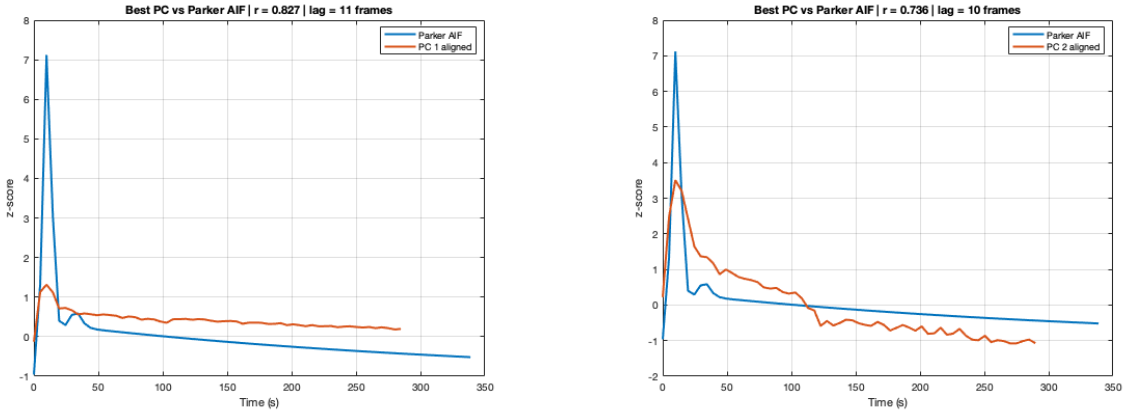


(b) v2: selects PC4

Figure 5.2: Subject P007. **(a)** Version v1 selects the first component ($r = 0.502$, negative lag), a non-arterial slow trend. **(b)** Version v2 rejects the first three components and selects the fourth ($r = 0.749$), which has a clear arterial first-pass profile. Curves are z-scores, aligned to the Parker reference.

Composite score and tail penalty (P011). In P011 both versions find an arterial-like component among the first three, but select different ones. The first component is the most correlated ($r = 0.827$), and version v1 selects it; however, its curve does not return to the baseline after the first pass but maintains an elevated tail, quantified by a tail penalty $\tau = 0.20$ (Figure 5.3, left). Version v2 ranks the components by the composite score $r(1 - \tau)$, which penalizes this sustained tail, and therefore selects the second component instead: although slightly less correlated ($r = 0.736$), its curve washes out completely ($\tau = 0.00$, Figure 5.3, right). This case isolates the effect of the composite score: v2 prefers a component with a cleaner wash-out over a more correlated one with a venous-like tail.

Fail-safe (P006). Subject P006 is the only case in which version v2 does not produce an AIF. None of the first eight components satisfies the three plausibility constraints simultaneously, and the fail-safe mechanism flags the dataset without forcing a selection. Version v1, which does not have such a mechanism, would instead select the most correlated of the first three components (Figure 5.4): a component with a very low correlation ($r = 0.339$) and a plateau shape that does not correspond to an arterial passage. This case illustrates the value of the fail-safe: where v1 would silently return an implausible AIF, v2 reports the failure explicitly. Subject P006 is therefore excluded from the analysis of Section 5.2.



(a) v1: selects PC1 ($\tau = 0.20$)

(b) v2: selects PC2 ($\tau = 0.00$)

Figure 5.3: Subject P011. **(a)** Version v1 selects the most correlated component (PC1, $r = 0.827$), whose curve keeps an elevated tail. **(b)** Version v2, ranking by the composite score $r(1 - \tau)$, selects the second component ($r = 0.736$), which washes out completely. Curves are z-scores, aligned to the Parker reference.

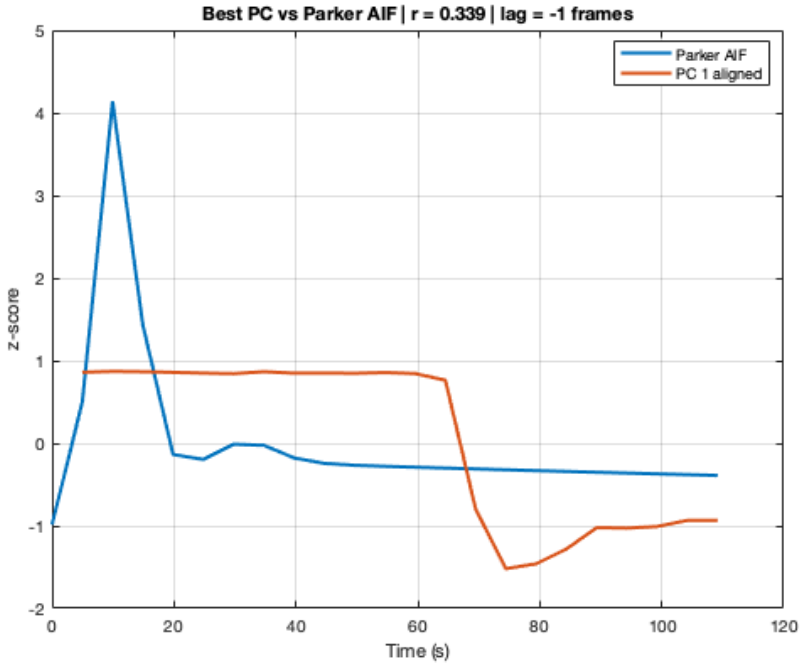


Figure 5.4: Fail-safe case (P006): the component that version v1 would select (PC1, $r = 0.339$), a plateau-shaped, non-arterial pattern. Version v2 rejects all components and returns no AIF.

5.1.4 Summary of the comparison

In clinical data, the two versions of the pipeline agree in nine of the twelve successful participants and differ in three. In one of these cases (P009), the difference has limited practical impact, since both versions select an arterial-like component. The remaining disagreements, together with the fail-safe case P006, highlight the role of the additional mechanisms introduced in v2.

The extended search and the component-level plausibility constraints recover the arterial component when it lies beyond the first three positions or when the early components are not physiologically valid (P007); the composite score discards a highly correlated but tail-heavy component in favour of one with a cleaner wash-out (P011); and the fail-safe mechanism prevents the selection of an implausible component when none is reliable (P006). These observations confirm, on the real data, the behaviour anticipated in the description of the method (Section 3.2.5) and complement the synthetic comparison of Chapter 4, in which the two versions were indistinguishable.

5.2 Clinical results of the proposed pipeline

This section presents the results obtained with version v2 in the twelve subjects for which an AIF was extracted.

5.2.1 Arterial input function extraction

For all twelve subjects the selected component reproduces the arterial profile, with correlations with the Parker reference AIF ranging from 0.632 to 0.838 (Table 5.1). The lowest value, found in P005, lies just above the acceptance threshold $R_{\min} = 0.6$ and corresponds to a component that is only weakly arterial-like; all other participants exceed 0.73. The number of candidate voxels returned by the two-stage percentile selection (Section 3.2.6) varies substantially throughout the cohort, from 188 in P012 to 907 in P007, and is not associated with a higher correlation: the subject with the most candidates does not have the highest r.

The final AIF of the representative subject P010 is shown in Figure 5.5, together with its smoothed and gamma-variate representations (Section 3.2.8). The figure illustrates a point already anticipated in the description of the method: the gamma-variate fit reproduces the first pass accurately but, decaying to zero by construction, does not capture the recirculation and the slow wash-out that are clearly visible in the measured curve. For this reason, the smoothed curve, which preserves these features while suppressing the

frame-to-frame noise, is the representation used for the subsequent computation of the perfusion maps.

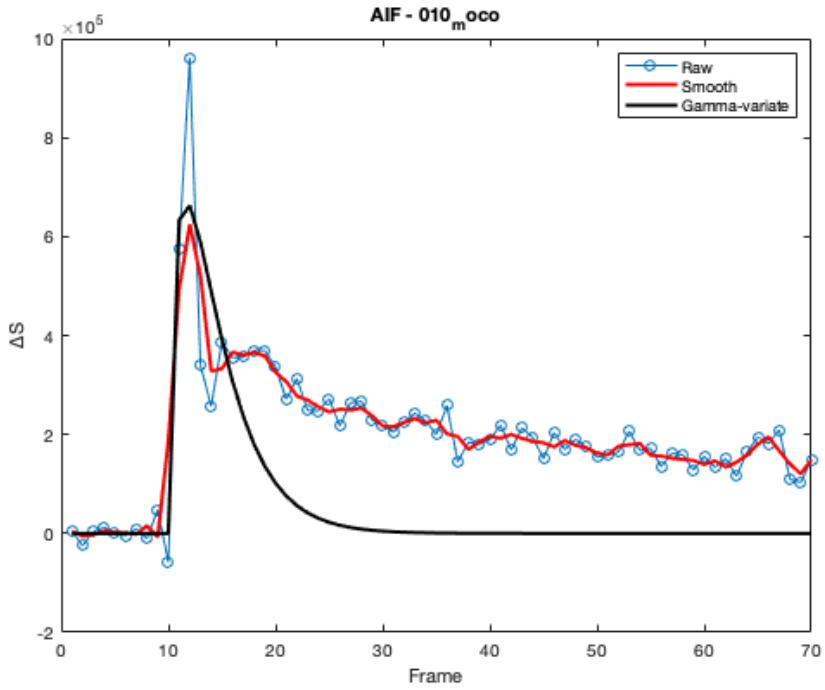
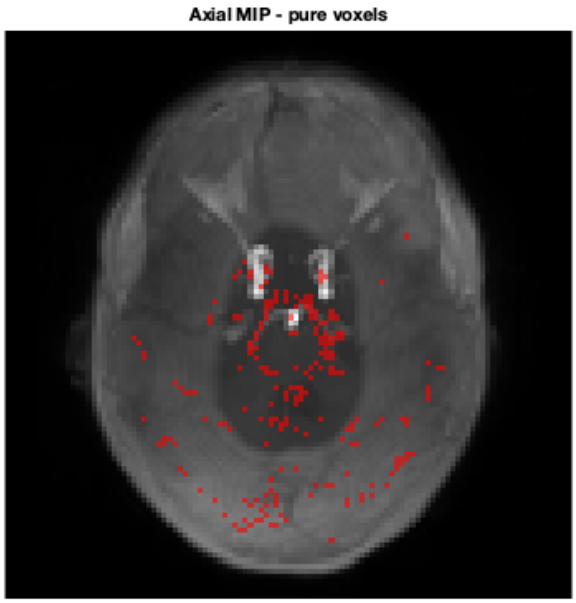


Figure 5.5: Representative case (P010): final extracted AIF. The raw curve (blue) is the average of the enhancement time-courses of the selected arterial voxels; the smoothed curve (red) is a three-point moving average; the gamma-variate (black) is the model fitted to the first pass. The gamma-variate decays to zero and does not reproduce the recirculation and wash-out retained by the smoothed curve.

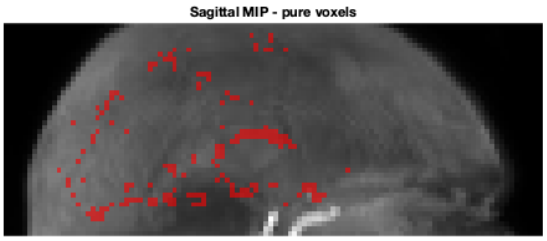
5.2.2 Spatial localization and the scarcity of identifiable pure arterial voxels

The spatial distribution of the selected voxels is visualized through maximum intensity projections (MIPs). A MIP is a two-dimensional image obtained by projecting a three-dimensional volume onto a plane and assigning to each pixel the maximum value encountered along the corresponding projection ray; it makes it possible to display, in a single image, all the selected voxels contained in the volume, regardless of the slice in which they lie. Figure 5.6 shows the axial, the coronal and the sagittal MIPs of the candidate arterial voxels for the representative subject, overlaid in red on the anatomical image.

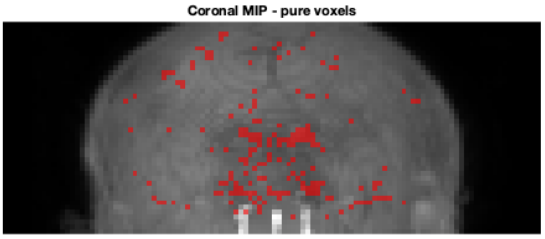
The candidate voxels concentrate around the major intracranial arteries, but they are accompanied by a substantial number of isolated voxels scattered throughout the parenchyma and near the skull, which do not correspond to arterial structures. This spatial heterogeneity is reflected in the voxel-level physiological refinement applied after



(a) Axial



(b) Sagittal



(c) Coronal

Figure 5.6: Representative case (P010): maximum intensity projections of the candidate arterial voxels, overlaid in red on the first dynamic frame. The voxels concentrate around the central arteries but also include scattered peripheral candidates.

the PCA selection. In almost every subject, fewer than three voxels satisfy all the arterial criteria, so that the fall-back rule is triggered. In these cases, the final AIF is computed from the five candidate voxels with the strongest enhancement. This indicates that the method does not identify a sufficiently large and stable set of voxels meeting all the criteria of a pure arterial curve.

This finding is consistent with the partial-volume effect characterized in synthetic data in Chapter 4: at the spatial resolution of the acquisition, intracranial arterial voxels are likely to contain a variable degree of tissue contribution, so that even the best candidates cannot be assumed to represent purely arterial signal.

5.2.3 An intermediate case

Not all successful subjects produce an AIF of the quality shown for P010. Subject P005 is reported here as an intermediate-quality case: the pipeline produces an AIF, unlike the fail-safe case P006, but the result is less stable and less physiologically plausible than the representative case P010. For this subject, the arterial component is the third principal component and its correlation with the Parker reference, $r = 0.632$, is the lowest of the cohort, just above the acceptance threshold. Its extracted AIF, shown in Figure 5.7, exhibits a first pass dominated by a single narrow spike, followed by oscillations and by a pronounced rising trend in the second half of the acquisition, instead of a return toward

the baseline. The spatial selection reflects the same behavior: the candidate voxels of P005 are markedly more numerous and more diffusely scattered than those of P010 (581 versus 248 candidates, respectively), with a large fraction spread throughout the parenchyma rather than confined to the major arteries (Figure 5.8).

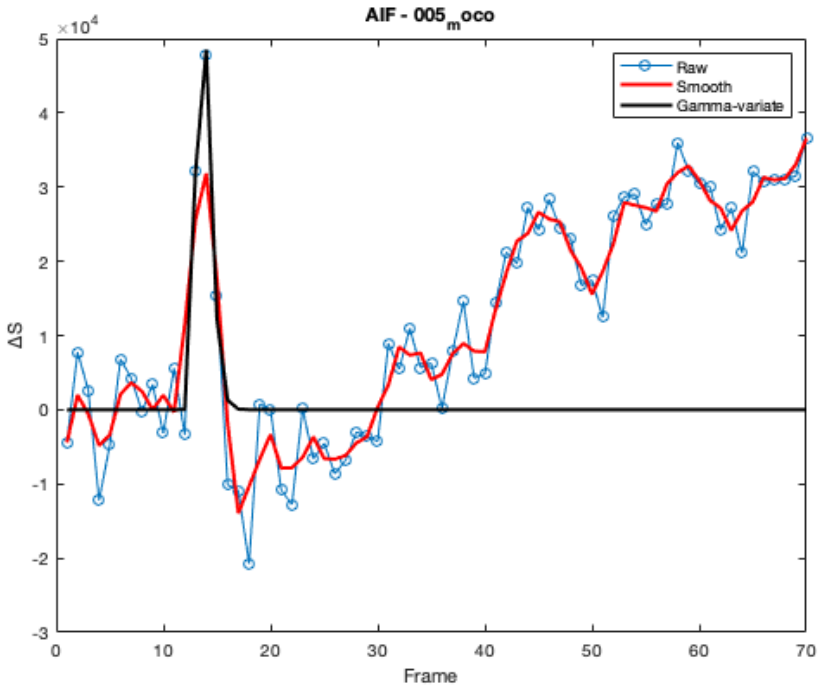
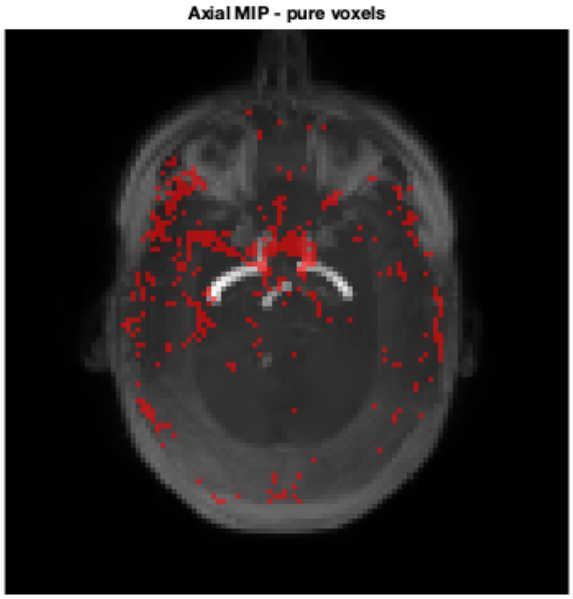
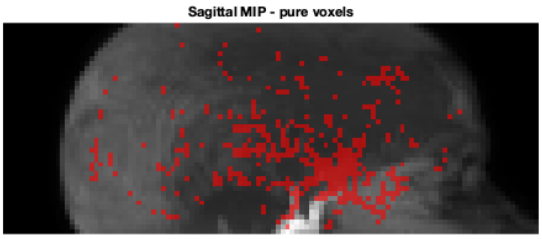


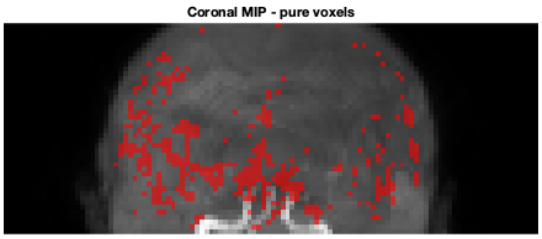
Figure 5.7: Intermediate case (P005): final extracted AIF (raw, smoothed and gamma-variate). The first pass is a single narrow spike; the curve then oscillates and rises markedly in the second half of the acquisition, instead of returning to the baseline.



(a) Axial



(b) Sagittal



(c) Coronal

Figure 5.8: Intermediate case (P005): maximum intensity projections of the candidate arterial voxels, overlaid in red on the first dynamic frame. Compared with the representative case (Figure 5.6), the candidate voxels are more numerous and more widely scattered across the parenchyma.

5.2.4 Cerebral perfusion maps

The cerebral perfusion maps were computed for the twelve subjects with a valid AIF, following the deconvolution procedure of Section 3.3. Table 5.2 reports, for each subject, the median values of CBF, CBV and MTT over the valid brain voxels, together with the cohort median.

Several observations can be drawn from Table 5.2. The cohort median CBF is approximately 35 mL/(100 g · min). This value is lower than the gray-matter values of roughly 70 mL/(100 g · min) reported for DCE-MRI perfusion by Larsson *et al.* [15]. However, this comparison should be interpreted cautiously, because the values reported here are semi-quantitative medians over valid brain voxels rather than fully quantitative measurements restricted to gray matter. The cohort median CBV (9 mL/100 g) and MTT (14 s) are of the same order as the corresponding reference values, although they tend to be higher. In all subjects, CBV is relatively uniform (from 7.8 to 11.2 mL/100 g), while CBF and MTT vary more widely; P001 has the highest CBF and the shortest MTT, and P005 and P007 the lowest CBF and the longest MTT. Within each subject, the spread of the values is large, with interquartile ranges of the same order as the median itself, reflecting a strong spatial heterogeneity of the maps.

Figure 5.9 shows the CBF, CBV and MTT maps of the representative subject P010

Table 5.2: Median values of the perfusion parameters over the valid brain voxels of each subject, and cohort median. CBF in mL/(100 g · min), CBV in mL/100 g, MTT in s. The values are semi-quantitative (Section 3.3.4).

Subject	CBF	CBV	MTT
P001	62.6	10.5	9.3
P002	46.2	10.3	12.9
P003	36.5	7.8	11.8
P004	44.1	9.7	12.9
P005	27.5	11.2	16.9
P007	28.6	8.6	17.1
P008	32.0	8.6	13.8
P009	38.0	8.4	13.6
P010	32.9	8.9	14.5
P011	41.4	11.0	15.0
P012	31.0	7.9	13.4
P013	32.9	9.7	16.8
Cohort median	34.7	9.3	13.7

on a central axial slice. The maps show low values throughout most of the parenchyma and high values along the cortical surface and, in particular, in a band in the posterior part of the brain. This posterior band, which is visible in all subjects in the cohort, is consistent with the location of the superior sagittal sinus and other posterior venous structures. Since these structures are not explicitly excluded from the brain mask used for deconvolution, they appear as regions of apparently very high blood flow and volume.

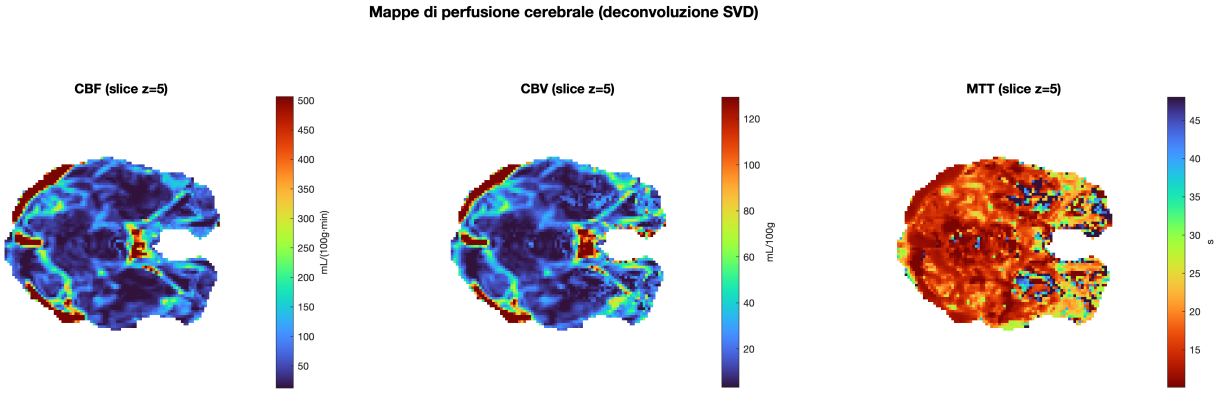


Figure 5.9: Representative case (P010): cerebral blood flow (CBF), cerebral blood volume (CBV) and mean transit time (MTT) maps on a central axial slice, obtained by SVD deconvolution with Tikhonov regularization. High CBF and CBV values are visible along the cortical surface and in the posterior venous structures.

The same spatial features are present, with greater spatial noise and heterogeneity, in the intermediate case P005, whose maps are shown in Figure 5.10: the parenchymal values

are more heterogeneous and patchy than those of P010, in line with the lower quality of the extracted AIF reported above.

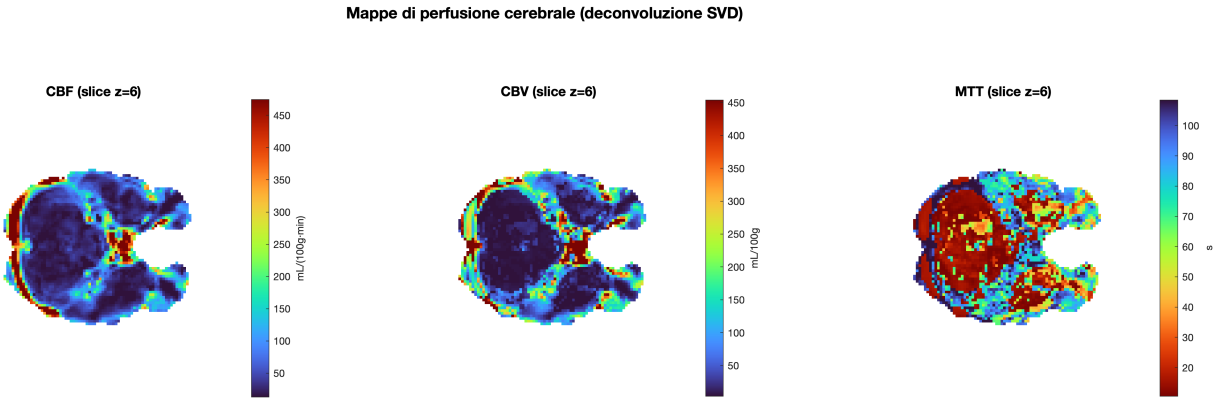


Figure 5.10: Intermediate case (P005): CBF, CBV and MTT maps on a central axial slice. Compared with the representative case (Figure 5.9), the parenchymal values are more heterogeneous and patchy.

As discussed in Section 3.3.4, these maps are semi-quantitative: the absolute scale of CBF, CBV and MTT depends on the empirical rescaling of the AIF amplitude and on the pseudo-concentration approximation. The values should therefore be interpreted as relative, intra-subject indices rather than as absolute physiological measurements. The implications of this limitation, together with the posterior venous contamination and the relationship between AIF quality and map quality, are discussed in Chapter 6.

6

Discussion

The synthetic validation and the clinical application together provide a coherent account of the proposed pipeline: a reliable recovery of the AIF waveform under controlled conditions, and a performance bounded by the partial-volume effect on clinical data, in agreement with the long-standing identification of partial volume as a principal source of error in arterial input estimation [6]. This chapter interprets that contrast: it reads the results of Chapters 4 and 5 against the method of Chapter 3, relates them to the approaches surveyed in Section 2.4, and isolates the limitations that Chapter 7 considers as directions for future work.

6.1 Robustness to noise

The noise study (Section 4.3) indicates that the shape of the estimated AIF is largely insensitive to additive noise: the peak-aligned correlation with the reference remained at $r_{\text{aligned}} \approx 0.98$ down to the lowest signal-to-noise ratio examined. This behavior is a direct consequence of the structure of the estimator. Because the arterial component is a temporal pattern shared by a large set of voxels, its estimation aggregates their contributions; under the assumption of independent, zero-mean noise, the residual fluctuation decreases approximately as $N^{-1/2}$ in the number of contributing voxels. The decomposition therefore performs an implicit denoising, which is particularly valuable given the intrinsically low signal-to-noise ratio of arterial voxels noted in the literature [6].

The voxel-selection step shows a different behavior. The introduction of noise expands the energy-fraction distribution, so that the relative threshold retains only the most unambiguously arterial voxels (72 of 720). The selection becomes stricter, but it does not introduce false positives: precision remains equal to one, meaning that every retained

voxel is genuinely arterial. For the fidelity of the estimated waveform this is a favorable regime; it also shows that the absolute number of retained voxels is not, on its own, an informative quantity.

6.2 Partial volume and the heterogeneity of contamination

The partial-volume study (Section 4.4) provides the central methodological result of this work. Because both simulated regimes span the same range of arterial fractions, the determinant of success is not the nominal contamination range, but the way contamination is distributed across the arterial population.

When the arterial voxels are predominantly pure, they are described by a single dominant temporal pattern; the selection isolates them, and the recovered input function closely matches the reference ($r_{\text{aligned}} = 0.971$). When the arterial fraction is spread over a wide range, the arterial signal is instead divided among several principal components – in the tested configuration, the first two components carried roughly 80% and 13% of the total variance – so that the component eventually selected accounts for only a small share ($f_k \approx 0.1$) of each genuine arterial voxel's own temporal variance. Low-energy tissue voxels, whose minor fluctuations happen to project onto the chosen component, then attain comparable or higher f_k values and dominate the relative ranking; as a result, the selected set is composed of tissue voxels rather than arterial ones.

The fail-safe of v2 does not intervene in this case. It acts only at the component level, declining an estimate when no principal component is plausible (Section 3.2.5). When a plausible arterial-like component does exist, as in this regime, the voxel selection still proceeds, and the relative threshold returns the top-ranked voxels even when they are tissue. There is therefore no equivalent safeguard at the voxel level capable of declaring that the selected voxel set is not physiologically credible, a point taken up among the limitations (Section 6.6).

This outcome reveals a structural property of the selection rule. The energy fraction f_k (Section 3.2.6) is a relative quantity: it measures the proportion of the voxel's own variance explained by the component, not the absolute strength with which the voxel expresses the arterial dynamics. A flat, low-energy voxel may therefore rank as highly as a strongly enhancing arterial one, and the double-percentile threshold, lacking an absolute cut-off, always returns a non-empty set. The criterion is well posed only when a coherent sub-population of near-pure voxels is present to dominate the ranking.

The two regimes were also examined with noise added on top of the partial-volume mixing, approximating the condition expected in the clinical acquisitions. The combined experiment reinforces this picture rather than altering it. When a coherent cluster of

near-pure voxels is present, the selection remains entirely arterial and the AIF correlation is unchanged across all noise levels tested ($r_{\text{aligned}} = 0.971$). When such a cluster is absent, the selection already fails in the noise-free case, and adding noise does not recover it: it remains dominated by tissue voxels at every level. Noise and partial volume thus act on different aspects of the problem: noise affects the precision of the voxel selection, whereas partial volume governs whether a correct selection is possible at all.

A final observation carries methodological weight. In the failure regime the extracted AIF still correlated moderately with the reference, and, counter-intuitively, this correlation even rose as the noise increased. The reason lies in the relative nature of the selection. Adding noise makes the energy-fraction threshold more selective, so that the number of retained voxels falls sharply (from one hundred to about ten), and only the voxels most tightly aligned with the chosen component survive. Their weighted average is therefore a cleaner rendering of that component's arterial-like shape, and its correlation with the reference improves, even though the retained voxels remain entirely tissue. The waveform is reproduced by the projection onto the component, irrespective of the spatial origin of the voxels that carry it; the rise in correlation is an artifact of the selection narrowing, not a genuine improvement. It follows that waveform similarity is, by itself, an insufficient criterion for validating an automatic AIF method, and that spatial localization and physiological plausibility are an integral part of the validation rather than an optional refinement.

6.3 From synthetic to clinical data

Applied to the clinical cohort, the pipeline produced a plausible arterial input function in almost every subject, and the fail-safe intervened where the arterial dynamics could not be recovered. What it did not do was isolate a coherent cluster of near-pure arterial voxels of the kind seen in the synthetic data: the voxel-level physiological refinement (Section 3.2.7) retained fewer than three valid voxels in almost every subject, so that the fall-back rule governed the estimate, which was then computed from the five most strongly enhancing candidates. This observation must be interpreted with care. It does not, by itself, prove that pure arterial voxels are absent, nor does it measure the degree of partial-volume contamination, which cannot be assessed directly in vivo due to the lack of a ground truth. What it does show is that no voxel met the operating criteria at the available resolution. Given the voxel size relative to the caliber of the intracranial arteries, widespread partial-volume contamination among the candidate voxels is the most plausible explanation, and it matches the behavior that the synthetic study associated with the heterogeneous contamination regime.

The spatial distribution of the selected voxels reflects the same difficulty. As shown in Chapter 5, the candidate voxels concentrate around the major intracranial arteries but also include scattered voxels in the parenchyma and near the skull that do not correspond to arterial structures. Two factors contribute to this: the brain mask is not conservative enough to exclude peri-cerebral tissue and large venous structures, so that non-arterial voxels can enter the analysis. In addition, the fall-back rule, which keeps the most strongly enhancing candidates, offers no protection against venous voxels, whose enhancement is itself strong. Therefore, the final estimate is not guaranteed to originate from purely arterial locations, a concern that the same mask limitation raises for the perfusion maps (Section 6.4).

The comparison between the two versions of the pipeline is informative in this context. In synthetic data the two are indistinguishable, since the arterial component is unambiguous and the modifications introduced in v2 never come into play. In clinical data, they act in specific, interpretable cases: the extended search and the plausibility constraints recover an arterial component lying beyond the first three (P007); the composite score discards a more correlated but tail-heavy component in favor of one with a cleaner wash-out (P011); and the fail-safe withholds an estimate when no component is plausible (P006). The intermediate case (P005), with a weakly arterial component ($r = 0.632$) and a diffuse selection spread throughout the parenchyma, marks the lower end of the quality range still admitted by the method.

For the clinical context in which the data were acquired, a study on first-episode psychosis, these observations bear on feasibility rather than on group-level inference, which was not the object of this work. A semi-quantitative, automatically extracted AIF of the quality reported here may support exploratory relative, intra-subject comparisons, but the partial-volume limitation would have to be resolved before such estimates could support quantitative between-subject conclusions.

6.4 Semi-quantitative perfusion maps

The perfusion maps must be read as relative, intra-subject indices rather than absolute measurements. Because the analysis is semi-quantitative – no T_1 mapping, and an AIF amplitude fixed by empirical rescaling – the conversion factor κ (Section 3.3.4) is a scaling convention rather than a calibration.

Within these constraints, the maps are internally consistent in their relative patterns and in the order of magnitude of the cohort-level summary values. The cohort medians (CBF ≈ 35 mL/(100 g $\cdot$ min), CBV ≈ 9 mL/100 g, MTT ≈ 14 s) are of the same order as the DCE-MRI values reported by Larsson *et al.* [15], although the cohort CBF is lower and

the MTT longer; the two are not directly comparable, since our value is a whole-brain median over all valid voxels rather than a gray-matter estimate. The recurrent region of apparently very high flow and volume along the posterior midline coincides with the superior sagittal sinus, which is not excluded from the brain mask and therefore enters the maps as a venous, rather than arterial or parenchymal, contribution. The deconvolution itself follows the SVD framework of Østergaard *et al.* [14] with the Tikhonov regularization of Calamante *et al.* [16].

6.5 Relation to previous work

These findings can be located within the methods surveyed in Section 2.4. The reference method of Sanz-Requena *et al.* [11], from which the present pipeline was derived, retained voxels whose variance was explained for at least 95% by the arterial component in prostate imaging. In the current brain dataset, the candidate voxels do not approach that level of component-specific variance explained, and the absolute criterion had to be replaced by a relative one. The discrepancy is attributable to the anatomical setting: the intracranial arteries are small relative to the voxel size, so that genuinely pure arterial voxels are far rarer than in the comparatively large prostatic vasculature sampled in the reference study.

The same consideration applies to the other data-driven families. Clustering-based selection [8, 9], including hierarchical schemes, and independent component analysis [10] share the objective that also motivates the present work: extracting the AIF automatically, without operator-dependent steps such as manual region-of-interest placement. Where they differ is in the assumption made about the data. Clustering partitions the voxels into groups and selects the one whose representative curve is most arterial; this presupposes that arterial voxels form a compact, well-separated group in the chosen feature space, and it rests on a few design choices, namely the feature representation, the number of clusters and the rule that labels one cluster as arterial. Independent component analysis makes the related assumption that the arterial signal is a statistically independent source that can be unmixed from the others.

The present pipeline is based on a weaker assumption. It does not require arterial voxels to be separable as a group: it treats the arterial dynamics as a dominant axis of variance and ranks the voxels continuously by their projection onto that axis, rather than assigning them to discrete clusters. When partial-volume mixing blurs the boundary between arterial and tissue voxels, which is the regime most consistent with the present clinical data, a clean separation into groups is exactly what becomes difficult to obtain, whereas a continuous ranking degrades more gracefully. On top of this, the explicit physiological plausibility constraints and the component-level fail-safe let the pipeline

reach a decision on its own, including the decision to withhold an estimate when no component is admissible. The contribution is therefore not the PCA decomposition in itself, but its combination with these constraints into a fully automatic procedure, aimed at removing operator dependence while remaining robust when variance maximization does not coincide with arterial specificity. In this sense, the present work continues, rather than supersedes, the line of automatic AIF-selection methods that these studies established.

6.6 Limitations

The preceding analysis identifies a set of limitations that are, for the most part, specific and addressable. Each is taken up as a direction for future work in Section 7.2.

1. **Relative selection criterion.** The energy-fraction criterion $f_k(v)$ has no absolute threshold. This is the root cause of the failure under heterogeneous contamination, because the selection cannot declare that no sufficiently arterial voxel is present and always returns a non-empty set.

2. **AIF amplitude not referenced to a partial-volume-free signal.** The AIF amplitude is fixed to an empirical target rather than scaled against a partial-volume-free venous output function [17, 18], leaving the peak only approximately corrected for partial-volume underestimation.

3. **Empirically set parameters.** In the current implementation, the regularization strength, the component- and voxel-level plausibility thresholds, and the peak target were tuned on the data rather than derived from an independent criterion or chosen by a data-adaptive procedure.

4. **Semi-quantitative scale.** The absence of T_1 mapping confines the analysis to pseudo-concentration, so that the perfusion maps carry an arbitrary absolute scale.

5. **Idealized synthetic model.** The digital reference object uses a regular arterial band, a single tissue partner curve, and a per-voxel arterial fraction assigned at random, and therefore does not reproduce realistic vascular geometry, spatially coherent partial-volume patterns or motion-related artifacts.

6. **Temporal resolution.** The clinical sampling interval (≈ 5 s per volume) is coarse relative to the width of the arterial first pass, which limits the precision of the recovered timing.

7. **Non-conservative brain mask.** Large venous structures, notably the superior sagittal sinus, and peri-cerebral tissue are not excluded from the analysis mask. They can therefore enter both the voxel selection and the perfusion maps, where they appear as spurious high-perfusion regions.

None of these limitations undermines the conclusions of the validation. Together, they delimit the regime within which the method is reliable and define the concrete steps required to improve its robustness, which are addressed in the next chapter.

7

Conclusions and Future Work

7.1 Conclusions

This thesis developed and tested an automatic, data-driven method for estimating the arterial input function from a DCE-MRI dataset. The pipeline applies principal component analysis to the voxel-wise time courses, identifies the component that best matches the expected arterial dynamics, and averages a small set of high-scoring voxels into a single AIF. A second version of the pipeline adds physiologically motivated plausibility constraints and a component-level fail-safe. On the synthetic data, the two versions showed equivalent behaviour in the tested conditions, indicating that the additional logic acts primarily as a safeguard rather than as a modification of the pipeline in cases where the arterial component is already unambiguous.

The method was first validated on a digital reference object, where a known ground truth allows a direct comparison. Under additive Rician noise, the arterial component was identified consistently and the recovered AIF remained close to the reference across the signal-to-noise levels examined. Noise mainly affected the voxel selection, making it more conservative, but without introducing false positives in the tested conditions. The partial-volume experiments drew a sharper line. When the arterial fraction was skewed towards nearly pure voxels, corresponding to the low-heterogeneity regime, the selection remained accurate. When the arterial fraction was spread uniformly across the arterial band, corresponding to the high-heterogeneity regime, the selection failed, returning non-arterial voxels even though the averaged curve could still correlate well with the reference. This last point is the central methodological finding: a high correlation of the recovered AIF shape does not certify, by itself, that the correct voxels were selected.

In the clinical data the pipeline produced a plausible AIF in almost every subject and the fail-safe intervened where the arterial dynamics were not recoverable. The semi-

quantitative perfusion maps were anatomically coherent, and the analysis did not isolate a coherent cluster of near-pure arterial voxels of the kind identified on the synthetic data. One possible interpretation is that the data appear closer to the heterogeneous partial-volume regime than to the more favorable one.

Taken together, the results suggest that the PCA-based approach can reliably recover the AIF waveform when a coherent set of near-pure arterial voxels is present, and that the practical limit in the clinical setting is the apparent scarcity of such voxels. The contribution of this work is therefore a complete and reproducible pipeline, together with a controlled characterization of where it succeeds and where it fails. The limitations collected in Section 6.6 are, for the most part, specific and addressable, and the next section outlines concrete directions for future work.

7.2 Future Work

The following directions address the limitations enumerated in Section 6.6. For each, the literature cited is meant as an example of a plausible route rather than as a settled solution; the aim is to indicate that the limitation is addressable, not to prescribe a single method.

Absolute voxel-selection criterion. One of the most important limitations is that the energy-fraction criterion $f_k(v)$ is purely relative and, in its current percentile-based implementation, always returns a non-empty set, so the selection cannot conclude that no sufficiently arterial voxel is present. A natural next step is to add an absolute admissibility test that a candidate voxel must pass before it is accepted, for instance a minimum agreement with an expected arterial morphology or a model-based fit to a population AIF such as the Parker form [7]. Goodness-of-fit gating of this kind is already used in cluster-based AIF selection [8], and adapting it to the present framework would allow the pipeline to abstain rather than report a spurious AIF when the data do not support one.

Amplitude scaling against a venous reference. The AIF amplitude is currently fixed to an empirical target. A more principled correction would scale the peak against a less partial-volume-affected venous output function, as is done in perfusion studies through tail scaling or related approaches [17, 18, 19]. Estimating a venous reference from a large draining vessel and using it to rescale the arterial peak is one example of how the amplitude could be tied to a measurable quantity rather than to a chosen constant.

Data-adaptive parameter selection. Several quantities – the regularization strength, the plausibility thresholds and the AIF peak target – were tuned on the data. In future

implementations, these quantities should be replaced by data-adaptive or independently justified choices. For the deconvolution step, the regularization weight could be selected per voxel using an L-curve or generalized-cross-validation criterion [16]. In parallel, delay-insensitive deconvolution schemes such as block-circulant SVD [21] could be explored to reduce sensitivity to bolus arrival delays. For the AIF extraction itself, fixed component- and voxel-level thresholds could be replaced by adaptive criteria derived from the empirical distribution of the candidate curves or from an independent validation cohort.

Absolute concentration through T_1 mapping. The analysis is confined to pseudo-concentration because no pre-contrast T_1 map was used. Acquiring or estimating the T_1 information needed for that conversion – for example, through a multiple-flip-angle acquisition, which has been used to estimate the contrast-agent concentration directly from the DCE signal [22] – would put both the AIF and the perfusion maps on a physical scale and remove the arbitrary amplitude that the semi-quantitative analysis currently carries.

More realistic synthetic models. The digital reference object used here is idealized, with a regular arterial band, a single tissue partner curve and an arterial fraction assigned at random. Validation would be strengthened by a phantom that reproduces realistic vascular geometry and spatially coherent partial-volume patterns. In silico frameworks that build a dynamic digital phantom from a detailed vascular model and a flow simulation [23] are an example of how spatially structured contamination could be generated, addressing a regime that the present synthetic model does not capture.

Higher temporal resolution. The deconvolution step is sensitive to how finely the arterial first pass is sampled, and a coarse sampling interval limits the precision of the recovered timing. In settings where the protocol can be adjusted, accelerated dynamic acquisitions offer a way to shorten the interval without sacrificing spatial coverage; golden-angle radial sparse-parallel imaging combined with compressed sensing [24] is one example of such a scheme. A finer sampling of the first pass would improve the timing estimates on which the deconvolution depends.

A more conservative analysis mask. Large venous structures and peri-cerebral tissue are not fully excluded from the current mask and can enter both the voxel selection and the perfusion maps. Two complementary routes could make the analysis mask more specific. The first is a vessel-aware preprocessing step designed to distinguish the arterial search region from venous and extra-cerebral structures. Multiscale vesselness filtering [25], possibly combined with anatomical constraints, could be used to identify large

vascular structures and exclude those that are incompatible with the expected arterial location, rather than indiscriminately removing all vessels.

The second route takes advantage of an observation made in clinical data, namely that the voxels retained after the PCA step tend to concentrate around the major intracranial arteries of the circle of Willis. This suggests restricting the analysis *a priori* to a coarse arterial region of interest, for instance through a loosely defined anatomical mask or an arterial atlas registered to the subject. Even a rough delineation would confine the search to the area where arterial voxels are expected and discard scattered, clearly extra-arterial candidates before they can be selected. Combined with the absolute selection criterion discussed above, either route would reduce the chance that non-arterial voxels are mistaken for arterial ones.

The recurring theme is that the method becomes more trustworthy as the selection is given an absolute notion of what an arterial voxel should look like and as the measurements are tied to physical, partial-volume-aware references. Pursuing these directions would move the pipeline from a semi-quantitative proof of concept towards a more robust and potentially quantitative tool for clinical DCE-MRI analysis.

A

MATLAB Implementation of the Pipeline

The complete AIF extraction pipeline and the subsequent computation of the cerebral perfusion maps were implemented in MATLAB (The MathWorks, Natick, MA, USA), using the routines of the *Tools for NIfTI and ANALYZE image* toolbox for reading and writing the imaging volumes. The listings collected in this appendix report the methodological core of each processing stage, in the order in which the data flow through the pipeline. Only the algorithmic content is reported. The listings are referenced from the corresponding sections of Chapter 3.

A.1 Signal pre-processing

The function in Listing A.1 converts the raw DCE-MRI signal into the tissue pseudo-concentration $\Delta S/S_0$ used throughout the pipeline. The scanner scaling is applied, the baseline signal S_0 is estimated as the mean of the first `Nbaseline` frames, and the resulting curves are restricted to the brain mask. The same routine is reused, with the appropriate baseline, both in the AIF extraction and in the perfusion-map computation.

Listing A.1: Computation of the tissue pseudo-concentration $\Delta S/S_0$.

```
function C_pseudo_masked = preprocess_DCE(fileDCE, fileMask, Nbaseline,
    doNotClampNeg)

if nargin < 3 || isempty(Nbaseline), Nbaseline = 10; end
if nargin < 4, doNotClampNeg = true; end

nii = load_untouch_nii(fileDCE);
raw = double(nii.img);
```

```
% Apply the scanner scaling (slope/intercept), with a safe fallback
slope = nii.hdr.dime.scl_slope;
inter = nii.hdr.dime.scl_inter;
if slope == 0, slope = 1; end
DCE = raw * slope + inter;

% Baseline signal S0 = mean of the first Nbaseline frames
S0 = mean(DCE(:,:,:,1:Nbaseline), 4);
S0(S0==0) = eps;

% Pseudo-concentration: S/S0 - 1
C_pseudo = (DCE ./ S0) - 1;

% Binarise and apply the brain mask
nii_mask = load_untouch_nii(fileMask);
mask = double(nii_mask.img) > 0;
C_pseudo_masked = C_pseudo .* mask;

% Optional clipping of negative values
if ~doNotClampNeg
    C_pseudo_masked(C_pseudo_masked < 0) = 0;
end
end
```

A.2 Reference arterial input function

Listing A.2 evaluates the population AIF model of Parker *et al.*, used as the reference arterial shape against which the principal components are compared during the identification step. The model is the sum of two Gaussians, describing the first pass and the recirculation, and an exponential decay modulated by a sigmoid, describing the wash-out.

Listing A.2: Parker population AIF model.

```
function Cp = parkerAIF_seconds(tSec, Nbaseline)
% Parker population AIF (Parker et al., 2006)

tMin = tSec(:) / 60; % the Parker model is defined in minutes
```

```
% Model parameters (Parker et al., 2006)
A1 = 0.809; A2 = 0.330;
T1 = 0.17046; T2 = 0.365;
sig1 = 0.0563; sig2 = 0.132;
alpha = 1.050; beta = 0.1685;
s = 38.078; tau = 0.483;

% Two Gaussians (first pass + recirculation)
gauss1 = (A1/(sig1*sqrt(2*pi))) * exp(-(tMin - T1).^2 / (2*sig1^2));
gauss2 = (A2/(sig2*sqrt(2*pi))) * exp(-(tMin - T2).^2 / (2*sig2^2));

% Sigmoid-modulated exponential decay (wash-out)
sigmoid = 1 ./ (1 + exp(-s*(tMin - tau)));
tail = alpha * exp(-beta * tMin) .* sigmoid;

Cp = gauss1 + gauss2 + tail;

% Optional baseline subtraction
if nargin ⩾ 2 && ~isempty(Nbaseline) && isfinite(Nbaseline) && Nbaseline
    ⩾ 1
   nb = min(round(Nbaseline), numel(Cp));
   Cp = Cp - mean(Cp(1:nb));
end
end
```

A.3 PCA-based AIF extraction (stage 1)

Listing A.3 contains the core of the first stage of the pipeline. The pseudo-concentration curves of the brain voxels are arranged in a data matrix and decomposed by principal component analysis through a singular value decomposition. The arterial component is then identified through the robust selection of version v2. Finally, the purest voxels are localized by the two-stage percentile selection on the arterial energy fraction and on the spatial score, and the candidate AIF is obtained as their energy-weighted average.

Listing A.3: Core of the PCA-based AIF extraction (stage 1, version v2).

```
% The DCE-MRI 4D volume and the brain mask are assumed already loaded:
%  [Nx,Ny,Nz,Nt] = size(DCE), dt = frame interval in seconds,
%  fileDCE, fileMask = NIfTI paths, brain_mask = 3D logical mask.
```

```
%% 1) Pseudo-concentration (S/S0 - 1) with a 10-frame baseline
Nbaseline = 10;
C4 = preprocess_DCE(fileDCE, fileMask, Nbaseline, true);
C4D = C4 .* double(brain_mask);

%% 2) Data matrix [Nvox x Nt] and mean-centering
X = reshape(C4D(repmat(brain_mask,[1 1 1 Nt])), [], Nt);
X(~isfinite(X)) = 0;
[Nvox, Nt] = size(X);
Xc = X - mean(X,1);   % center each temporal variable

%% 3) PCA via SVD: Xc = U*S*P'
[U,S,P] = svd(Xc,'econ');
Tscore = U * S;       % voxel scores in the PC coordinate system
lambda = diag(S).^2;
explainedVar = lambda / sum(lambda);

%% 4) Reference AIF: Parker model (z-scored)
t_sec = (0:Nt-1)*dt;
parker0 = parkerAIF_seconds(t_sec, 10);
aif_ref = zscore(parker0);
maxLagFrames = round(0.15 * Nt); % max acceptable PC/reference peak lag

%% 5) Robust arterial-component selection (v2)
% Extended search (8 PCs), plausibility constraints, composite score
% r*(1-tail), and a fail-safe if no component qualifies.
nPC  = min(8, size(P, 2));
R_MIN = 0.60;

bestCorr   = -inf(nPC, 1);
bestLagFrames = zeros(nPC, 1);
PC_aligned = zeros(Nt, nPC);
PC_sign    = ones(nPC, 1);
tail_PC    = ones(nPC, 1); % tail/peak ratio of each PC
score_PC   = -inf(nPC, 1); % composite score
plausible_PC = false(nPC, 1); % plausibility flag
```

```
42 [~, refPeak] = max(aif_ref);

44 for i = 1:nPC
45    pc_raw = P(:, i);
46    % the PCA sign is arbitrary: test both polarities
47    pc_candidates = [zscore(pc_raw(:)), zscore(-pc_raw(:))];

49    r_best_i    = -Inf;
50    lag_best_i  = 0;
51    sign_best_i = +1;
52    pc_best_shifted = nan(Nt, 1);
53    pc_best_raw = nan(Nt, 1);

55    for sgn = 1:2
56       pc = pc_candidates(:, sgn);
57       [~, pcPeak] = max(pc);
58       lag_try = pcPeak - refPeak;
59       if abs(lag_try) > maxLagFrames
60          continue;
61       end
62       pc_shifted = shift_with_nan(pc, -lag_try);
63       valid = isfinite(pc_shifted) & isfinite(aif_ref);
64       if nnz(valid) ⩾ 2
65          R = corrcoef(pc_shifted(valid), aif_ref(valid));
66          r_try = R(1,2);
67       else
68          r_try = -Inf;
69       end
70       if r_try > r_best_i
71          r_best_i    = r_try;
72          lag_best_i  = lag_try;
73          pc_best_shifted = pc_shifted;
74          pc_best_raw = pc;
75          if sgn == 1, sign_best_i = +1; else, sign_best_i = -1; end
76       end
77    end
```

```
79    bestCorr(i)  = r_best_i;
80    bestLagFrames(i) = lag_best_i;
81    PC_aligned(:, i) = pc_best_shifted;
82    PC_sign(i)   = sign_best_i;

84    % tail of the PC (sign-corrected shape): mean of the last 30% of the
85    % curve normalized to its peak; penalizes venous/tissue-like plateaus
86    pc_form = pc_best_raw(:);
87    pc_form(~isfinite(pc_form)) = 0;
88    [pk, ~] = max(pc_form);
89    if pk > 0
90       tail_start_idx = max(2, round(0.7 * Nt));
91       tail_PC(i) = mean(pc_form(tail_start_idx:end)) / pk;
92    else
93       tail_PC(i) = 1;
94    end
95    tail_PC(i) = max(0, min(1, tail_PC(i)));

97    % physiological plausibility constraints
98    [~, peakIdxPC] = max(pc_best_raw);
99    is_plausible = true;
100   if r_best_i < R_MIN
101      is_plausible = false;       % too weakly correlated
102   elseif lag_best_i ⩽ 0
103      is_plausible = false;       % PC peaks before the reference
104   elseif peakIdxPC < round(0.10 * Nt) || peakIdxPC > round(0.70 * Nt)
105      is_plausible = false;       % peak outside the [10%, 70%] window
106   end
107   plausible_PC(i) = is_plausible;

109   % composite score: correlation weighted by a clean wash-out
110   if is_plausible
111      score_PC(i) = r_best_i * (1 - tail_PC(i));
112   else
113      score_PC(i) = -inf;         % rejected PCs cannot win
114   end
```

```
115 end

117 % fail-safe: abstain if no component is physiologically plausible
118 if ~any(plausible_PC)
119     error(['No principal component among the first %d satisfies the ' ...
120           'plausibility constraints (r ⩾ %.2f, lag > 0, peak in window). '
                  ...
121           'No reliable arterial component is identified for this subject.'],
                  ...
122           nPC, R_MIN);
123 end

125 [~, bestPC] = max(score_PC);

127 %% 6) Double PCA-based selection of the purest voxels
128 % (1) high arterial energy fraction frac_k;
129 % (2) high positive spatial score on the selected component.

131 % arterial rank-one reconstruction of every voxel
132 xk = Tscore(:,bestPC) * (PC_sign(bestPC) * P(:,bestPC))'; % [Nvox x Nt]

134 E_tot = sum(Xc.^2, 2); % total voxel energy
135 E_k  = sum(xk.^2, 2); % energy explained by the selected PC

137 frac_k = zeros(size(E_tot));
138 valid = E_tot > 0;
139 frac_k(valid) = E_k(valid) ./ E_tot(valid);

141 % spatial score of the selected PC for each voxel (sign-corrected)
142 selectedSpatialMap = Tscore(:,bestPC) * PC_sign(bestPC);
143 posSpatial = selectedSpatialMap > 0;

145 % Stage 1: relative threshold on frac_k (top 5%)
146 fracPrctile = 95;
147 thr_frac = prctile(frac_k(valid), fracPrctile);
148 mask_frac = false(size(frac_k));
149 mask_frac(valid) = frac_k(valid) ⩾ (thr_frac - 1e-12);
```

```
% Stage 2: relative threshold on the spatial score (top 10%),
% applied only to stage-1 voxels with a positive score
spatialPrctile = 90;
candidateStage2 = mask_frac & posSpatial;
mask_spatial = false(size(selectedSpatialMap));
if any(candidateStage2)
    thr_spatial = prctile(selectedSpatialMap(candidateStage2),
        spatialPrctile);
    mask_spatial(candidateStage2) = ...
       selectedSpatialMap(candidateStage2) ⩾ (thr_spatial - 1e-12);
end
pureMask = mask_spatial;

%% 7) Candidate AIF: frac_k-weighted average of the selected voxels
if any(pureMask)
    Xpure = X(pureMask, :);
    wpure = frac_k(pureMask);
    wpure = wpure(:) / sum(wpure);
    AIF_final = wpure' * Xpure;
else
    error('No voxels selected by the double PCA-based threshold.');
end

%% Local function: shift a vector by k frames, padding with NaN
function y = shift_with_nan(x, k)
    x = x(:);
    n = numel(x);
    y = nan(size(x));
    if k == 0
       y = x;
    elseif k > 0
       if k < n, y((k+1):end) = x(1:end-k); end
    else
       k = abs(k);
       if k < n, y(1:end-k) = x((k+1):end); end
    end
```

```
end
```

A.4 Physiological refinement and final AIF (stage 2)

Listing A.4 reports the second stage, which refines the candidate arterial mask produced by stage 1 on the basis of the first-pass shape of the individual voxel curves. The retained voxels are then averaged into the final AIF, together with a lightly smoothed version and a gamma-variate fit of the first pass.

Listing A.4: Core of the physiological refinement and final AIF estimation (stage 2).

```
% Inputs assumed loaded: D = 4D DCE volume, M = candidate arterial mask
% (3D logical, output of stage 1), with [X,Y,Z,T] = size(D).

%% Parameters
Nbaseline  = 7;
max_tail_ratio = 0.25;
max_TTP    = 3;     % frames beyond the median time-to-peak
min_upslope = 0.1;

%% Baseline subtraction on the candidate voxels
D_resh  = reshape(D, [], T);
mask_vec = M(:);
vox_curves = D_resh(mask_vec, :);
baseline = mean(vox_curves(:,1:Nbaseline), 2);
deltaS  = bsxfun(@minus, vox_curves, baseline);
nvox    = size(deltaS, 1);

%% First-pass shape descriptors for each voxel
[peak_val, peak_idx] = max(deltaS, [], 2);

% arrival time = first frame exceeding 5% of the peak
arrival = zeros(nvox,1);
for v = 1:nvox
   thr = 0.05 * peak_val(v);
   idx = find(deltaS(v,:) > thr, 1, 'first');
   if isempty(idx), arrival(v) = NaN; else, arrival(v) = idx; end
end
```

```
% upslope = mean rate of increase between arrival and peak
TTP = peak_idx;
upslope = zeros(nvox,1);
for v = 1:nvox
   a = arrival(v); p = TTP(v);
   if isnan(a) || p ⩽ a
      upslope(v) = 0;
   else
      upslope(v) = (deltaS(v,p) - deltaS(v,a)) / max(1,(p-a));
   end
end

% tail ratio = area after the peak / area before the peak
tail_ratio = zeros(nvox,1);
for v = 1:nvox
   p = TTP(v);
   if p == 0 || isnan(p), tail_ratio(v) = Inf; continue; end
   a = arrival(v); if isnan(a), a = 1; end
   area_before = sum(deltaS(v,a:p));
   n_after  = min(T, p + round(2*(p-a)));
   area_after = sum(deltaS(v,p+1:n_after));
   tail_ratio(v) = area_after / (area_before + eps);
end

%% Physiological selection of arterial voxels
median_peak = round(median(TTP(~isnan(TTP))));
sel = (TTP ⩽ median_peak + max_TTP) & ...
     (upslope ⩾ min_upslope) & ...
     (tail_ratio ⩽ max_tail_ratio) & ...
     (peak_val > 0);

% fall-back: if fewer than 3 voxels survive, keep the 5 strongest by peak
if sum(sel) < 3
   [~, ord] = sort(peak_val, 'descend');
   sel = false(nvox,1);
   sel(ord(1:min(5,nvox))) = true;
end
```

```
sel_curves = deltaS(sel, :);

%% Final AIF: raw average, light smoothing, gamma-variate fit
AIF_raw = mean(sel_curves, 1);
AIF_smooth = smooth(AIF_raw, 3);

[~, peak_frame] = max(AIF_raw);
fp_start = max(1, peak_frame - 3);
fp_end = min(T, peak_frame + round((peak_frame - fp_start)*2));
t = (fp_start:fp_end)';
y = AIF_raw(fp_start:fp_end)';

gamma_fun = @(p,x) p(1)*((x-p(2)).^p(3)) .* exp(-(x-p(2))/p(4)) .* (x>p(2));
p0 = [max(y), fp_start, 2, 2];
lb = [0, 0, 0.1, 0.1]; ub = [inf, fp_start+5, 10, 20];
try
    opts = optimset('Display','off');
    p_est = lsqcurvefit(gamma_fun, p0, t, y, lb, ub, opts);
    AIF_gamma = gamma_fun(p_est, (1:T)');
catch
    AIF_gamma = AIF_smooth; % fall back to the smoothed curve
end
```

A.5 Cerebral perfusion maps

Listing A.5 computes the cerebral perfusion maps from the estimated AIF and the tissue pseudo-concentration curves, through a Tikhonov-regularized SVD deconvolution. The flow-scaled residue function recovered for each voxel is converted into cerebral blood flow, cerebral blood volume and mean transit time, expressed in clinical units.

Listing A.5: Core of the perfusion-map computation by SVD deconvolution.

```
% Inputs assumed available: DCE = 4D volume, AIF = stage-2 smoothed AIF,
% brain_mask and art_mask (3D), and the NIfTI paths fileDCE, fileMask.

%% Parameters
TR            = 4.94681; % repetition time [s]
Nbaseline     = 7;       % baseline frames before bolus arrival
Hct           = 0.45;    % arterial hematocrit
```

```
density_brain  = 1.04;   % brain tissue density [g/mL]
svd_thresh_rel = 0.05;  % Tikhonov threshold relative to sigma_max
target_peak    = 5.0;    % AIF peak target in delta S/S0 units
firstpass_window_s = 30; % post-peak integration window for CBV [s]

%% AIF normalisation and amplitude rescaling
S0_per_voxel = mean(DCE(:,:,:,1:Nbaseline), 4);
S0_arterial = mean(S0_per_voxel(art_mask), 'omitnan');
AIF = AIF / S0_arterial;        % normalise to delta S/S0
AIF = AIF * (target_peak / max(AIF)); % rescale peak to target

%% Tissue pseudo-concentration and tissue voxels (brain minus arteries)
C_pseudo = preprocess_DCE(fileDCE, fileMask);
[X, Y, Z, T] = size(C_pseudo);
DCE_2D  = reshape(C_pseudo, [], T);
brain_idx = find(brain_mask > 0 & ~art_mask);
Ct_voxels = DCE_2D(brain_idx, :)';
Ct_voxels = Ct_voxels - mean(Ct_voxels(1:Nbaseline, :), 1, 'omitnan');
Nvox    = numel(brain_idx);

%% Causal Toeplitz convolution matrix from the AIF
H = zeros(T);
for i = 1:T
    H(i, 1:i) = AIF(i:-1:1)' * TR;
end

%% SVD deconvolution with Tikhonov regularisation
[U, S, V] = svd(H, 'econ');
s     = diag(S);
s_inv = s ./ (s.^2 + (svd_thresh_rel * max(s))^2);
H_pinv = V * diag(s_inv) * U';

%% Voxel-wise Meier-Zierler triad
N_window = max(1, round(firstpass_window_s / TR));
F_vec = zeros(Nvox, 1);
CBV_vec = zeros(Nvox, 1);
MTT_vec = zeros(Nvox, 1);
```

```
for v = 1:Nvox
    I_curve = H_pinv * Ct_voxels(:, v);
    I_curve(I_curve < 0) = 0;    % non-negativity constraint
    [F_v, peak_idx] = max(I_curve); % CBF ~ peak of the residue function
    end_idx = min(T, peak_idx + N_window);
    CBV_v = sum(I_curve(1:end_idx)) * TR; % first-pass integral
    if F_v > 0 && CBV_v > 0
        MTT_v = CBV_v / F_v;
    else
        MTT_v = NaN;
    end
    F_vec(v) = F_v; CBV_vec(v) = CBV_v; MTT_vec(v) = MTT_v;
end

%% Conversion to clinical units (hematocrit and tissue density)
kappa = (1 - Hct) / density_brain;
CBF_vec = (F_vec / kappa) * 60 * 100; % mL/(100g*min)
CBV_vec = (CBV_vec / kappa) * 100; % mL/100g
% MTT is unaffected by the conversion and is already expressed in seconds

%% Mask non-physical values and reconstruct the 3D maps
invalid = ~isfinite(CBF_vec) | CBF_vec ⩽ 0 | ...
        ~isfinite(CBV_vec) | CBV_vec ⩽ 0 | ...
        ~isfinite(MTT_vec) | MTT_vec ⩽ 0;
CBF_vec(invalid) = NaN; CBV_vec(invalid) = NaN; MTT_vec(invalid) = NaN;

CBF_map = nan(X, Y, Z); CBF_map(brain_idx) = CBF_vec;
CBV_map = nan(X, Y, Z); CBV_map(brain_idx) = CBV_vec;
MTT_map = nan(X, Y, Z); MTT_map(brain_idx) = MTT_vec;
```

Bibliography

[1] Thomas E. Yankeelov and John C. Gore. "Dynamic Contrast Enhanced Magnetic Resonance Imaging in Oncology: Theory, Data Acquisition, Analysis, and Examples". In: *Current Medical Imaging Reviews* 3.2 (2009), pp. 91–107. DOI: 10.2174/157340507780619179.

[2] Paul S. Tofts. "Modeling Tracer Kinetics in Dynamic Gd-DTPA MR Imaging". In: *Journal of Magnetic Resonance Imaging* 7.1 (1997), pp. 91–101. DOI: 10.1002/jmri.1880070113.

[3] Steven P. Sourbron and David L. Buckley. "Classic Models for Dynamic Contrast-Enhanced MRI". In: *NMR in Biomedicine* 26.8 (2013), pp. 1004–1027. DOI: 10.1002/nbm.2940.

[4] Kyung Jae Kim et al. "Dynamic Contrast-Enhanced MRI and Its Applications in Various Central Nervous System Diseases". In: *Investigative Magnetic Resonance Imaging* 26.4 (2022), pp. 256–264. DOI: 10.13104/imri.2022.26.4.256.

[5] Anna K. Heye et al. "Tracer Kinetic Modelling for DCE-MRI Quantification of Subtle Blood–Brain Barrier Permeability". In: *NeuroImage* 125 (2016), pp. 446–455. DOI: 10.1016/j.neuroimage.2015.10.018.

[6] Fernando Calamante. "Arterial Input Function in Perfusion MRI: A Comprehensive Review". In: *Progress in Nuclear Magnetic Resonance Spectroscopy* 74 (2013), pp. 1–32. DOI: 10.1016/j.pnmrs.2013.04.002.

[7] Geoffrey J. M. Parker et al. "Experimentally-Derived Functional Form for a Population-Averaged High-Temporal-Resolution Arterial Input Function for Dynamic Contrast-Enhanced MRI". In: *Magnetic Resonance in Medicine* 56.5 (2006), pp. 993–1000. DOI: 10.1002/mrm.21066.

[8] Kim Mouridsen et al. "Automatic Selection of Arterial Input Function Using Cluster Analysis". In: *Magnetic Resonance in Medicine* 55.3 (2006), pp. 524–531. DOI: 10.1002/mrm.20759.

[9] Denis Peruzzo et al. "Automatic Selection of Arterial Input Function on Dynamic Contrast-Enhanced MR Images". In: *Computer Methods and Programs in Biomedicine* 104.3 (2011), e148–e157. DOI: 10.1016/j.cmpb.2011.02.012.

[10] Jiandong Yin, Jiawen Yang, and Qiyong Guo. "Automatic Determination of the Arterial Input Function in Dynamic Susceptibility Contrast MRI: Comparison of Different Reproducible Clustering Algorithms". In: *Neuroradiology* 57.5 (2015), pp. 535–543. DOI: 10.1007/s00234-015-1493-9.

[11] Roberto Sanz-Requena et al. "Automatic Individual Arterial Input Functions Calculated from PCA Outperform Manual and Population-Averaged Approaches for the Pharmacokinetic Modeling of DCE-MR Images". In: *Journal of Magnetic Resonance Imaging* 41.1 (2015), pp. 133–140. DOI: 10.1002/jmri.24487.

[12] Paul S. Tofts et al. "Estimating Kinetic Parameters from Dynamic Contrast-Enhanced T1-weighted MRI of a Diffusable Tracer: Standardized Quantities and Symbols". In: *Journal of Magnetic Resonance Imaging* 10.3 (1999), pp. 223–232. DOI: 10.1002/(SICI)1522-2586(199909)10:3<223::AID-JMRI2>3.0.CO;2-S.

[13] Paul Meier and Kenneth L. Zierler. "On the Theory of the Indicator-Dilution Method for Measurement of Blood Flow and Volume". In: *Journal of Applied Physiology* 6.12 (1954), pp. 731–744. DOI: 10.1152/jappl.1954.6.12.731.

[14] Leif Ostergaard et al. "High Resolution Measurement of Cerebral Blood Flow Using Intravascular Tracer Bolus Passages. Part I: Mathematical Approach and Statistical Analysis". In: *Magnetic Resonance in Medicine* 36.5 (1996), pp. 715–725. DOI: 10.1002/mrm.1910360510.

[15] Henrik B. W. Larsson et al. "Measurement of Brain Perfusion, Blood Volume, and Blood–Brain Barrier Permeability, Using Dynamic Contrast-Enhanced T1-weighted MRI at 3 Tesla". In: *Magnetic Resonance in Medicine* 62.5 (2009), pp. 1270–1281. DOI: 10.1002/mrm.22136.

[16] Fernando Calamante, David G. Gadian, and Alan Connelly. "Quantification of bolus-tracking MRI: Improved characterization of the tissue residue function using Tikhonov regularization". In: *Magnetic Resonance in Medicine* 50.6 (2003), pp. 1237–1247. DOI: 10.1002/mrm.10643.

[17] Matthias J. P. van Osch et al. "Correcting partial volume artifacts of the arterial input function in quantitative cerebral perfusion MRI". In: *Magnetic Resonance in Medicine* 45.3 (2001), pp. 477–485. DOI: 10.1002/1522-2594(200103)45:3<477::AID-MRM1063>3.0.CO;2-4.

[18] Linda Knutsson et al. "Reduction of arterial partial volume effects for improved absolute quantification of DSC-MRI perfusion estimates: Comparison between tail scaling and prebolus administration". In: *Journal of Magnetic Resonance Imaging* 41.4 (2015), pp. 903–908. DOI: 10.1002/jmri.24621.

[19] Sukhdeep Singh Bal et al. “Optimal scaling approaches for perfusion MRI with distorted arterial input function (AIF) in patients with ischemic stroke”. In: *Brain Sciences* 12.1 (2022), p. 77. DOI: 10.3390/brainsci12010077.

[20] Andrew B. Gill, Ferdia A. Gallagher, and Martin J. Graves. “Open source code for the generation of digital reference objects for dynamic contrast-enhanced MRI analysis software validation”. In: *The British Journal of Radiology* 96.1147 (2023), p. 20220976. DOI: 10.1259/bjr.20220976.

[21] Ona Wu et al. “Tracer Arrival Timing-Insensitive Technique for Estimating Flow in MR Perfusion-Weighted Imaging Using Singular Value Decomposition with a Block-Circulant Deconvolution Matrix”. In: *Magnetic Resonance in Medicine* 50.1 (2003), pp. 164–174. DOI: 10.1002/mrm.10522.

[22] Ayesha Bharadwaj Das et al. “Estimation of Contrast Agent Concentration in DCE-MRI Using Two Flip Angles”. In: *Investigative Radiology* 57.5 (2022), pp. 343–351. DOI: 10.1097/RLI.0000000000000845.

[23] Chengyue Wu et al. “An In Silico Validation Framework for Quantitative DCE-MRI Techniques Based on a Dynamic Digital Phantom”. In: *Medical Image Analysis* 73 (2021), p. 102186. DOI: 10.1016/j.media.2021.102186.

[24] Li Feng et al. “Golden-Angle Radial Sparse Parallel MRI: Combination of Compressed Sensing, Parallel Imaging, and Golden-Angle Radial Sampling for Fast and Flexible Dynamic Volumetric MRI”. In: *Magnetic Resonance in Medicine* 72.3 (2014), pp. 707–717. DOI: 10.1002/mrm.24980.

[25] Alejandro F. Frangi et al. “Multiscale Vessel Enhancement Filtering”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 1998*. Vol. 1496. Lecture Notes in Computer Science. Springer, 1998, pp. 130–137. DOI: 10.1007/BFb0056195.