Development of a numerical method to study the diffusion of SERS NPs in resected tissue for binding potential assessment and evaluation

Final Degree Project
Escola Tècnica Superior de Telecomunicacions de Barcelona
(ETSETB)

Author: Eric Navarro Comes
Advisor: Kenneth M. Tichauer, PhD

July 19th, 2016
Abstract

An intraoperative approach for breast tumor-margin assessment is presented, using a kinetic model for surface-enhanced Raman scattering nanoparticles (SERS NPs). A numerical method is developed to understand the behavior of SERS NPs in healthy and cancerous tissue, which is able to validate and study diverse binding potential (BP) models, determine what protocol is best for the tumor-margin assessment, the sensitivity of each parameter and estimate the diffusion coefficient of tissue using experimental data.

Resum

En aquesta tesi es presenta una tècnica intraoperativa per valorar el marge dels tumors de càncer de mama. Això es fa mitjançant un model cinètic per a SERS NPs. Un mètode numèric és desenvolupat per entendre el compartament d’aquestes partícules en teixit sa i malalt, que també podrà validar i estudiar diversos models per calcular el binding potential, ajudar a triar un protocol adient en cada cas, estudiar la sensibilitat dels paràmetres i estimar el coeficient de difusió de certs experiments.

Resumen

En esta tesis se presenta una técnica intraoperativa para valorar el margen de los tumores de cáncer de mama. Esto se lleva a cabo mediante el estudio y desarrollo de un modelo cinético para SERS NPs. Se desarrolla un método numérico para entender el comportamiento de las partículas en tejido enfermo y sano; que también puede validar y estudiar diversos modelos para el cálculo del binding potential, ayudar a escoger un protocolo para cada caso, estudiar la sensibilidad de los parámetros y estimar el coeficiente de difusión de ciertos experimentos.
Contents

1. Introduction........................................................................................................5
   1.1 Current technology state.............................................................................6
   1.2 Outline and aim of the thesis.................................................................8
2. Approach...........................................................................................................9
   2.1 SERS NP.................................................................................................9
      2.1.1 Raman Spectroscopy.................................................................10
      2.1.2 NP..............................................................................................11
   2.2 Protocols..................................................................................................12
   2.3 Diffusion..................................................................................................13
   2.4 Compartment model..............................................................................14
   2.5 Conclusion.............................................................................................15
3. Implementation and results............................................................................17
   3.1 System setup...........................................................................................17
   3.2 Forward model......................................................................................19
   3.3 Backward models...................................................................................30
   3.4 Diffusion effect on estimating BP.........................................................32
   3.5 MPSA......................................................................................................34
   3.6 Experimental data...................................................................................39
   3.7 D estimation............................................................................................42
4. Conclusions and future work........................................................................46
List of figures

Figure 1. Intraoperative imaging technique .................................................................6
Figure 2. Dual-paired imaging ......................................................................................7
Figure 3. SERS NPs workflow, imaging system and diagram .....................................10
Figure 4. SERS NPs diagram .....................................................................................11
Figure 5. Protocol steps .............................................................................................12
Figure 6. Impact of D in diffusion ............................................................................13
Figure 7. Compartment model for targeted NPs .....................................................14
Figure 8. System setup for data analysis ..................................................................18
Figure 9. System setup for data as an input ............................................................19
Figure 10. Discretized mesh ....................................................................................20
Figure 11. C-N calculations .....................................................................................21
Figure 12. C-N implementation for staining ............................................................24
Figure 13. C-N implementation for imaging ............................................................27
Figure 14. C-N implementation for rinsing ..............................................................30
Figure 15. Effects of D on BP estimation (staining protocol) ....................................32
Figure 16. Effects of D on BP estimation (rinsing protocol) ....................................33
Figure 17. K-S score diagram .................................................................................36
Figure 18. K-S score (staining protocol) .................................................................36
Figure 19. K-S score (imaging protocol) .................................................................37
Figure 20. K-S score (single agent) .......................................................................38
Figure 21. Raw data for control agent in healthy tissue ..........................................40
Figure 22. Profile concentration in healthy tissue ..................................................41
Figure 23. Profile concentration in cancerous tissue ..............................................41
Figure 24. Experimental workflow .......................................................................42
Figure 25. D fitting for healthy tissue ....................................................................43
Figure 26. D fitting for cancer tissue ......................................................................45
Chapter 1

Introduction

Lumpectomy and partial mastectomies (breast-conserving surgeries) are the surgical procedures to apply in most of the cases when the breast tumor is in an early stage, since clinical trials have proven that they have similar success rate as complete removal of the breast (mastectomy). However, there is a big upside down for this conserving techniques: anywhere from 20 to 60% of lumpectomy patients must undergo additional surgery if there has been an incomplete removal of the tumor mass [1].

Typically, tumor-margin assessment is carried out post-surgery using standard histopathology. There are some rules to determine if the patient must undergo further surgery. The margin is defined to be positive if there are cancer cells on the outer surface of the tumor; close if they are between 1 and 3 mm from the margin and negative if there are no cells close to the surface [2]. Even though there is some controversy about the margin criteria, there is no doubt that re-excision is necessary when cells are found on the surface.

Since the assessment is performed post-surgery, if re-excision is needed the patient would have to assume the health risks of going under an oncology surgical procedure again, as well as the extra economic cost. Therefore, there is a need for an intra-operative tumor-margin assessment that is fast and reliable.

Some imaging techniques have been developed greatly last decade, such as radiography, computed tomography (CT), magnetic resonance imaging (MRI) or positron emission tomography (PET). However, these techniques are very time consuming and have a great economic cost, therefore there is a need for other techniques.
Other approaches have been developed for intraoperative surgical guidance, such as frozen section histology, light reflectance spectroscopy (LRS) or auto fluorescence lifetime measurement (AFLM) [3]. These are good techniques, but in the case of breast cancer, the fat content of the healthy tissue surrounding the tumor makes the application of such techniques very difficult.

To solve all this issues, a new intra-operative method is developed using the molecular-imaging technology. This technology approach has the potential to identify tumors with high sensitivity [3]. A new approach, widely backed, is using surface-enhanced Raman spectroscopy nanoparticles conjugated with biological agents.

![Figure 1](image-url)  
Figure 1. Illustration demonstrating the intraoperative imaging technique. Credit to Y.W.

### 1.1 Current technology state

The detection of cell surface receptors on cancer cells is the main interest as the way of discriminating healthy tissue from cancer tissue [4]. Therefore, cancer-targeted molecular probes were developed to measure the concentration of those receptors, in which targeted SERS NPs are
introduced in the excised tissue and then the concentration is measured. However, the signal obtained from a single molecular probe targeting cancer can be affected by many factors (i.e., diffusion, physiological factors) that are not related to receptor binding.

Most groups use now the dual-probe approach, in which there is one probe targeting the cancer receptors and one probe that’s not targeting anything, with the intention to reduce the noise factors.

Dual-paired imaging means there are two agents used: one of them is looking to bind to the cancer cell receptors, and the other is used to eliminate background and nonspecific noise. This is needed because the targeted agent can also bind to tissue parts it is not supposed to bind (nonspecific binding), therefore we could have a false-positive. Control agents (untargeted) are introduced, since those NPs would also bind in the same non-specific zones as the targeted agent, so it is easy to account for those areas [5].

![Image](image.png)

**Figure 2.** On the left, we can see the signal for the untargeted agent. On the center, the signal for the targeted agent. On the right, the signal normalized for the binding potential map. The highlighted areas in red and yellow are those who represent the highest binding potential.

When measuring the concentration of targeted and untargeted SERS NPs, two signals are acquired, related to the concentration of the NPs in the tissue. There is a diversity of models (backward models) that take this two signals and determine the binding potential (BP), which denotes how probable is it that there are still cancer cells in the tissue.

However, the most sophisticated backward models have the capability to quantify receptor concentrations, but the models do not account for diffusion.
1.2 Outline and aim of the thesis

The backward models used to estimate the binding potential in tissues do not account for the effects of diffusion. The numerical model developed in this thesis (forward model) will allow us to evaluate the impact on the diffusion parameter on the backward models. Additionally, it will allow us to replicate the diffusion of the SERS NPs and see how each backward model estimates the BP for diverse protocols.

The aim of this thesis is to develop a kinetic model that takes into account diffusion in the topical application of SERS NPs on fresh tissue, and to identify the most sensitive parameters based on the estimation for binding potential (BP), a value that is proportional to cell surface receptor concentrations.

A multiparametric sensitivity test is also performed for different backward models, determining which parameters are more sensitive than others. This can be used to determine the focus of experiments by prioritizing certain procedures among others.
Chapter 2

Approach

In order to tackle the problematic of this thesis, there is a need to understand how the SERS NPs play the key role in the molecular-imaging approach. Later I move on how to model and what equations define the diffusion phenomena in tissue, and then the final equations derived from a compartment model approach will be presented, which will be the equations describing the behavior of the SERS NPs.

Finally, the different SERS NPs application protocols are presented.

2.1 SERS NP

Surface-enhanced Raman spectroscopy nanoparticles (SERS NPs) are silica-shell encapsulated nanoparticles that can emit distinct Raman scattering signals upon being excited by a laser source (narrowband). They have attracted interest of many research groups because of their brightness, photo stability and multiplexing capabilities with laser illumination.

One of the most important properties is that they can be developed in a variety of flavors, each of which generate a unique spectral fingerprint when excited; therefore, each flavor can be easily separated from others.

By targeting different flavors of NPs to different biomarkers (for instance, one flavor to cancer-receptor targeting NPs and the other to control NPs), it is possible to carry on multiplexed molecular imaging so that the quantity of each NP flavor in a mixture can be determined through a spectral-demultiplexing software [6]. Therefore, different signal readings can be obtained by illuminating the tissue area with the same laser light.
2.1.2 Raman Spectroscopy

Raman spectroscopy is a technique used to observe low-frequency modes in a system. It is used to provide a fingerprint from which certain molecules can be identified and be differentiated from the rest. This technology is applied for the quantitative and qualitative analysis of inorganic, organic and biological system.
The theoretical foundation relies on the property of scattering of the particles upon laser excitation. The light interacts with the molecules creating a vibration resulting in energy changes so the laser photons are shifted; this shift in the energy is what provides the information about what kind of molecule the received signal is from.

Surface-enhanced Raman spectroscopy is a variation of Raman spectroscopy who happens to be more sensitive allowing for smaller molecules detection. It is widely used in conjunction with metallic nanoparticles.

2.1.3 Nanoparticles

Generally, even though SERS can happen in any metal surface, gold is the one most widely used due to its high efficiency when scattering and absorbing light. The NPs size is also very important when it comes to its optical properties, normally around 60 nm of diameter. Besides its optical properties, the bigger the NPs are, the harder it will be for them to diffuse inside the tissue.

These particles are then covered by an active layer of Raman active material, which is going to define the flavor of that nanoparticle.

![SERS NPs diagram with the biological markers (mAb) and the flavors (DyLight).](image)

**Figure 4.** SERS NPs diagram with the biological markers (mAb) and the flavors (DyLight).
2.2 Protocols

There are different protocols and methods aiming to obtain the best result for binding potential in excised tissue. Most of them are a combination of three different steps: staining, imaging and rinsing. Three protocols will be presented and, later on, simulated to see which ones are better for certain backward models (models used to estimate the binding potential).

First of all, it is important to present what staining, imaging and rinsing is.

1) **Staining**: relates to the topical application of SERS NPs on the surface of the tissue, expecting them to diffuse inside it.

2) **Imaging**: relates to the insulation of the tissue while being excited by a laser light in order to obtain the signals with the concentration of NPs in the tissue.

3) **Rinsing**: relates to the partial removal of NPs from the tissue. The tissue is submerged in a solution (i.e. PBS), allowing the NPs to diffuse outside of the tissue. Another approach for rinsing is to use a Kimwipe instead of submerging

This steps will be later discussed in Chapter 3, since they are the core of the forward model.

![Figure 5. Staining, imaging and rinsing illustrations.](image)

This thesis focuses on the following three protocols:

1) **Repeated stain protocol**: this protocol consists in staining the tissue for a given time, then imaging it; repeating this two steps for $n$ times.

2) **Single stain protocol**: same protocol as the repeated staining, but the tissue is stained and imaged only once.
3) Repeated stain and rinse protocol: this protocol approach is to stain the tissue for a certain time, and then proceed to rinse and image, repeating this last two steps for \( n \) times.

### 2.3 Diffusion

Diffusion is the movement of molecules from a region of high concentration to a region of low concentration, due to the existence of a concentration gradient.

The diffusion of SERS NPs in tissue is, principally, governed by the Fick’s second law of diffusion. This law predicts how diffusion causes the concentration to change with time in porous media, such as tissue, thus the application for this thesis. It is also used when it comes to calculate how topical medicaments diffuse inside the skin, therefore the similarity with this thesis aim is remarkable. The law is as follows:

\[
\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2}
\]

Where \( D \) refers to the diffusion coefficient, and \( C \) to the concentration. This formula contains partial differential equations, which will be solved using a numerical method, discussed later on Chapter 3.

![Figure 6](image)

**Figure 6.** Effects of different values of D in the solution of Fick’s second law.
This expression will describe the behavior of the untargeted (control) SERS NPs, but the targeted NPs need a kinetic approach to model the reactions inside the tissue.

2.4 Compartment models

A compartment model is a mathematical model used to describe the exchange of energies, concentrations, etc. between compartments of a system. In the study of this thesis, I define compartment as whatever different status the NPs can acquire. Therefore, in the case of the targeted NPs, I define two compartments: free-interstitium (C_f) and bound (C_b). A particle belongs to the free-interstitium compartment when it first comes into the tissue and has the possibility of binding, when it would change to the bound compartment. Targeted NPs can go from one compartment to another through the so called rate association or dissociation constants. An auxiliary compartment that accounts for the staining and rinsing solutions is used, but it is not included in the general form of the equation.

Figure 7. Compartment model for targeted NPs.
In the other hand, the untargeted nanoparticles will only be able to be in the free-interstitium compartment, thus it is easier to model them.

It is important to note that another compartment such as non-specific binding could be added, but it was decided that its impact would not be critical. However, the system will be scalable if this compartment is needed in the future.

### 2.5 Conclusion

Finally, we can formulate the mathematical expression hypothesis by which the system will be governed. Distinguishing by the two kinds of imaging agents:

1) **Targeted NPs**: a combination of Fick’s second law of diffusion and the compartment model kinetic analysis is needed. The nanoparticles entering from the staining solution will be in the free compartment, from which they can keep diffusing deeper or recombine to the bound compartment. In the other hand, when a NP is in the bound compartment, it can only recombine.

\[
\frac{\partial C_f}{\partial t} = D \frac{\partial^2 C_f}{\partial x^2} - k_3 C_f + k_4 C_b \\
\frac{\partial C_b}{\partial t} = k_3 C_f - k_4 C_b \\
C_t = C_f + C_b
\]
2) Untargeted NPs: only the diffusion equation will be account for here since there are no kinetic reactions.

\[ \frac{\partial C_u}{\partial t} = D \frac{\partial^2 C_u}{\partial x^2} \]

Other effects such as the convection of NPs or cancer-cell receptors saturation are not taking into account because it is not thought to be critical in the calculations.
Chapter 3

Implementation and results

In this chapter I will discuss the whole thesis workflow, how to develop the algorithms that set up the numerical method in charge of solving the partial derivative equations that account for the behavior of the SERS NPs (forward model), what models are used to calculate the binding potential (backward models).

Additionally, a multiparametric sensitivity test is presented in which it can be seen the impact of the different parameters in the backward models and protocols.

Finally, the experimental data used to tune the forward model is presented, along with the estimation of the diffusion coefficient.

3.1 System setup

The core of the setup and the thesis is without any doubt the numerical system (forward model) and its implementation. There are two ways to use the forward model:

1) To simulate the behavior of SERS NPs from input parameters such as the target BP \( BP = \frac{k_3}{k_4} \), the protocol parameters (protocol to use; stain, image and rinse times, number of repeats, the diffusion coefficient and the initial concentration of NPs in the staining solution). Given this inputs, the forward model can:

a) Estimate the BP using each backward model.

b) See how the diffusion coefficient changes the BP estimation.
c) Perform a multiparametric sensitivity analysis to see how each parameter impact the BP estimation or the total concentration of NPs.

d) Generate profile concentration curves that describe how the NPs have diffused after the simulated experiment.

2) To analyze experimental data coming from real-life situation experiments. In this case, the inputs are the profile concentration curves and the protocol used (with its respective timings and repeats values). Given this inputs, the forward model can:

a) Estimate the diffusion coefficient of the tissue used.

b) Estimate the BP using each backward model.

c) Tune the model parameters to be more realistic for future simulations.

**Figure 8.** Setup for data analysis
3.2 Forward model

Forward model is the name given to the numerical method in the thesis. The forward model uses a finite-difference method known as Crank-Nicolson [7]. This a method used to numerically solve partial differential equations by approximating them with difference equations, used widely for the solution of the diffusion or heat equation.

The main reason to use this method and not any other is because it is an implicit method (which implies that it is more accurate than other methods, like the explicit method), and it is unconditionally stable. Since the model will be fed a wide diversity of parameters, its stability was something to account for in the design phase.

The model is 1-dimensional. Even though the SERS NPs do not necessarily diffuse in a straight line inside the tissue and random walk might happen,
we decided the error would be negligible and not worth the extra computational time.

Crank-Nicolson works discretizing space and time in a grid. Therefore, the length of the tissue is divided in \( nx \) increments of length \( dx \); and the time is divided in \( nt \) increments of duration \( dt \).

![Discretized mesh](image)

**Figure 10.** Discretized mesh.

The method calculates the concentrations in certain space points (three points at a time, centered around \( i \)) in a future time step \((n+1)\), using the current time step values \((n)\), which are known.

The equations defining the system are the following:

\[
\frac{dC_f}{dt} = D \frac{d^2C_f}{dx^2} - k_3C_f + k_4C_b
\]

\[
\frac{dC_b}{dt} = k_3C_f - k_4C_b
\]

\[
C_{\text{targeted}} = C_f + C_b
\]
And for the untargeted agent:

\[
\frac{dC_u}{dt} = D \frac{d^2 C_u}{dx^2}
\]

Crank-Nicolson is a second order method in time (account for both current and future time steps), therefore the discretization for each partial derivative are as follows:

\[
\frac{\partial C_f}{\partial t} \sim \frac{1}{dt} (C_f^{n+1} - C_f^n)
\]

\[
\frac{\partial^2 C_f}{\partial t^2} \sim \frac{1}{2dx^2} (C_{f_{i+1}}^{n+1} - 2C_{f_i}^{n+1} + C_{f_{i-1}}^{n+1} + C_f^n - 2C_f^n + C_f^{n+1})
\]

And for the rate constants:

\[
k_xz \sim \frac{1}{2} k_x (C_{z_i}^{n+1} + C_{z_i}^n)
\]

Coupling the difference equations together:

\[
\frac{1}{dt} (C_{f_{i+1}}^{n+1} - C_{f_i}^n) = \frac{1}{2dx^2} (C_{f_{i+1}}^{n+1} - 2C_{f_i}^{n+1} + C_{f_{i-1}}^{n+1} + C_f^n - 2C_f^n + C_f^{n+1}) - \frac{1}{2} k_3 (C_f^{n+1} + C_f^n) + \frac{1}{2} k_4 (C_{b_i}^{n+1} + C_{b_i}^n)
\]

\[
\frac{1}{dt} (C_{b_{i+1}}^{n+1} - C_{b_i}^n) = \frac{1}{2} k_3 (C_{f_{i+1}}^{n+1} + C_{f_i}^n) - \frac{1}{2} k_4 (C_{b_{i+1}}^{n+1} + C_{b_i}^n)
\]

As we can see, the calculations are done over the mesh following this form:

![Figure 11: Crank Nicolson calculations](image-url)
The unknowns of the numerical system are those concentration values in a future time step \((n+1)\), and the knowns are those values in the current time step \((n)\). Therefore, the concentration values in three space steps in a future time step are calculated from the three space steps in the current time step, and then the recombination equations (the ones accounting for the binding and unbinding) are applied, only taking into account one space step.

Separating the unknowns from the knowns, we can rewrite the equations.

\[
-rC_{f_{i-1}}^{n+1} + 2(1 + r)C_{f_i}^{n+1} - rC_{f_{i+1}}^{n+1} + \frac{1}{2}dt(k_3C_{f_{i+1}}^{n+1} - k_4C_{b_{i+1}}^{n+1})
\]

\[
= rC_{f_{i-1}}^n + 2(1 - r)C_{f_i}^n + rC_{f_{i+1}}^n - \frac{1}{2}dt(k_3C_{f_i}^n - k_4C_{b_i}^n)
\]

\[
C_{b_{i+1}}^{n+1} \left(1 + \frac{1}{2}dtk_4\right) - \frac{1}{2}dtk_3C_{f_{i+1}}^{n+1} = C_{b_i}^n \left(1 - \frac{1}{2}dtk_4\right) + \frac{1}{2}dtk_3C_{f_i}^n
\]

\[
r = D \frac{dt}{dx^2}
\]

Applying the same principles to the simpler equation for the untargeted agent:

\[
\frac{1}{dt} (C_{f_{i+1}}^{n+1} - C_{f_{i}}^{n}) = \frac{1}{2dx^2} (C_{f_{i-1}}^{n+1} - 2C_{f_{i}}^{n+1} + C_{f_{i+1}}^{n+1} + C_{f_{i-1}}^{n} - 2C_{f_{i}}^{n} + C_{f_{i+1}}^{n})
\]

Separating the unknowns from the knowns:

\[
-rC_{f_{i-1}}^{n+1} + 2(1 + r)C_{f_i}^{n+1} - rC_{f_{i+1}}^{n+1} = rC_{f_{i-1}}^n + 2(1 - r)C_{f_i}^n + rC_{f_{i+1}}^n
\]

The optimal way to compute these equations is through a matrix form because it is easier to compute the coupled system for the targeted agent, since \(C_b\) and \(C_f\) has to be computed simultaneously. Also, the matrices will be diagonal and sparse, which means that the majority of its values will be 0, and MATLAB has efficient algorithms to deal with these matrices.

The core of the forward model are the stain, image and rinse functions, since a combination of them is what makes the protocols and generates the outputs. Each function needs to be computed in different ways because the boundary conditions are different. Boundary conditions control what happens at each side of the tissue, and, since the NPs will only diffuse about 30 micrometers inside the tissue, the boundary condition at the lower edge of the tissue is fixed to 0.
The boundary condition controlling what happens at the surface of the tissue models the application of a solution filled with SERS NPs, the removal of that solution and the effect of a kimwipe or an empty solution for rinsing.

In the next pages it is discussed how the algorithms and expressions for the targeted imaging agents diffusion are obtained. The untargeted ones are not discussed in this chapter because it is a simpler version of the targeted. However, the MATLAB implementation is included in the annex.

1) Staining function

The boundary condition in this case will model the existence of a solution filled with SERS NPs on the surface of the tissue. Therefore, a new row is created on top of the tissue mesh, that will account for the virtual space step on the surface \(x=-1\). All the positions will be shifted one position, being \(C_f^{n+1}\) the position accounting for the solution for all time steps and \(C_f^{n+1}_{nx+1}\) the position accounting for the lower edge of the tissue. It is important to note that the tissue still has \(nx\) space steps.

An important assumption is made to define the value of the upper boundary: the concentration of SERS NPs is high enough to not significantly decrease as it enters the tissue. Therefore, it will be set as a constant.

\[
\begin{align*}
C_t(-1, t) &= K \\
C_t(L, t) &= 0
\end{align*}
\]

Using difference equations notation:

\[
\begin{align*}
C_b^{n+1}_1 + C_f^{n+1}_1 &= K \\
C_b^{n+1}_{nx+1} + C_f^{n+1}_{nx+1} &= 0
\end{align*}
\]

This boundary condition is called Dirichlet.

The NPs cannot be bound inside the solution because there are no cancer receptors there. Additionally, it can be considered that in absence of NPs in the free interstitium, there will also be absence of NPs in the bound compartment. Therefore:

\[
\begin{align*}
C_f^{n+1}_1 &= K \\
C_f^{n+1}_{nx+1} &= 0
\end{align*}
\]
In order to build the matrix, we must understand first if the equations follow a certain pattern. To do so, we will have a look on the equations around specific space steps.

It is important to note that we do not need to calculate the values in the positions $i = 1$ nor in $i=nx+1$ since they are known. Also, it is possible to simplify some of the equations, which will be the key for optimizing the algorithm:

**i=2 (surface of tissue)**

\[-rC_1^{n+1} + 2(1 + r)C_{f_2}^{n+1} - rC_{f_3}^{n+1} + \frac{1}{2} dt(k_3C_{f_2}^{n+1} - k_4C_{b_2}^{n+1})\]

\[= rC_{f_1}^n + 2(1 - r)C_{f_2}^n + rC_{f_3}^n - \frac{1}{2} dt(k_3C_{f_2}^n - k_4C_{b_2}^n)\]

\[C_{b_2}^{n+1} \left( 1 + \frac{1}{2} dt k_4 \right) - \frac{1}{2} dt k_3 C_{f_2}^{n+1} = C_{b_2}^n \left( 1 - \frac{1}{2} dt k_4 \right) + \frac{1}{2} dt k_3 C_{f_2}^n\]

We can rearrange the first equation since we know the value for $C_1^{n+1}$, which is the boundary condition.
\[
2(1 + r)C_{f_2}^{n+1} - rC_{f_3}^{n+1} + \frac{1}{2} dt (k_3 C_{f_2}^{n+1} - k_4 C_{b_2}^{n+1})
= rC_{f_1}^n + 2(1 - r)C_{f_2}^n + rC_{f_3}^n - \frac{1}{2} dt (k_3 C_{f_2}^n - k_4 C_{b_2}^n) + rC_{f_1}^{n+1}
\]

**i=3...nx-1 (inside the tissue)**

\[
\begin{align*}
& \quad rC_{b_2}^{n+1} + 2(1 + r)C_{f_3}^{n+1} - rC_{f_4}^{n+1} + \frac{1}{2} dt (k_3 C_{f_3}^{n+1} - k_4 C_{b_3}^{n+1}) \\
&= rC_{f_1}^n + 2(1 - r)C_{f_2}^n + rC_{f_3}^n - \frac{1}{2} dt (k_3 C_{f_2}^n - k_4 C_{b_2}^n) \\
&C_{b_3}^{n+1} \left(1 + \frac{1}{2} dt k_4\right) - \frac{1}{2} dt k_3 C_{f_3}^{n+1} = C_{b_3}^n \left(1 - \frac{1}{2} dt k_4\right) + \frac{1}{2} dt k_3 C_{f_3}^n \\
&\quad \ldots
\end{align*}
\]

\[
\begin{align*}
& \quad rC_{b_{nx-1}}^{n+1} + 2(1 + r)C_{f_{nx-1}}^{n+1} - rC_{f_{nx-2}}^{n+1} + \frac{1}{2} dt (k_3 C_{f_{nx-1}}^{n+1} - k_4 C_{b_{nx-1}}^{n+1}) \\
&= rC_{f_{nx-2}}^n + 2(1 - r)C_{f_{nx-1}}^n + rC_{f_{nx-1}}^n - \frac{1}{2} dt (k_3 C_{f_{nx-1}}^n - k_4 C_{b_{nx-1}}^n) \\
&C_{b_{nx-1}}^{n+1} \left(1 + \frac{1}{2} dt k_4\right) - \frac{1}{2} dt k_3 C_{f_{nx-1}}^{n+1} = C_{b_{nx-1}}^n \left(1 - \frac{1}{2} dt k_4\right) + \frac{1}{2} dt k_3 C_{f_{nx-1}}^n
\end{align*}
\]

**i = nx (around the edge)**

\[
\begin{align*}
& \quad -rC_{b_{nx}}^{n+1} + 2(1 + r)C_{f_{nx}}^{n+1} - rC_{f_{nx+1}}^{n+1} + \frac{1}{2} dt (k_3 C_{f_{nx}}^{n+1} - k_4 C_{b_{nx}}^{n+1}) \\
&= rC_{f_{nx-1}}^n + 2(1 - r)C_{f_{nx}}^n + rC_{f_{nx+1}}^n - \frac{1}{2} dt (k_3 C_{f_{nx}}^n - k_4 C_{b_{nx}}^n) \\
&C_{b_{nx}}^{n+1} \left(1 + \frac{1}{2} dt k_4\right) - \frac{1}{2} dt k_3 C_{f_{nx}}^{n+1} = C_{b_{nx}}^n \left(1 - \frac{1}{2} dt k_4\right) + \frac{1}{2} dt k_3 C_{f_{nx}}^n
\end{align*}
\]

Again, same situation as in i=2, since we know the values for the nx+1 positions. The first equation is rewritten.

\[
\begin{align*}
& \quad -rC_{b_{nx}}^{n+1} + 2(1 + r)C_{f_{nx}}^{n+1} - rC_{f_{nx+1}}^{n+1} + \frac{1}{2} dt (k_3 C_{f_{nx}}^{n+1} - k_4 C_{b_{nx}}^{n+1}) \\
&= rC_{f_{nx-1}}^n + 2(1 - r)C_{f_{nx}}^n + rC_{f_{nx+1}}^n - \frac{1}{2} dt (k_3 C_{f_{nx}}^n - k_4 C_{b_{nx}}^n) + rC_{f_{nx+1}}^{n+1}
\end{align*}
\]

Grouping all together in a matrix form.
The equations can be written as a combination of the matrices above:

\[ AB^{n+1} = CB^n + D \]
2) Imaging function

The boundary condition in the surface will model an insulation effect. This means that no NPs will be able to diffuse through the surface. The equation defining this phenomenon is the following:

\[ \frac{\partial C}{\partial t} (0, t) = 0 \]

This boundary condition is known as Neumann boundary condition, and needs a different approach when being computed.

Using the second order approximation for the difference equation in this case:

\[ \frac{\partial C(0, t)}{\partial t} \approx \frac{\left( -3C_{f1}^{n+1} + 4C_{f2}^{n+1} - C_{f3}^{n+1} \right)}{2dx} = 0 \]

\[ C_{f1}^{n+1} = \frac{4}{3} C_{f2}^{n+1} - \frac{1}{3} C_{f3}^{n+1} \]

In a similar way as done in the staining function, we will have a look at the expression in both the surface and the edge. Note that there is no longer an added row modeling the solution, therefore \( i=1 \) will be the value at the surface now.

\[ C_{f1}^{n+1} = \frac{4}{3} C_{f2}^{n+1} - \frac{1}{3} C_{f3}^{n+1} \]

\[ C_{f_{nx}}^{n+1} = 0 \]

\[ C_{f0}^{n+1} = C_{o} \]

Figure 13. C-N implementation for imaging
Rewriting the equation for \( i=2 \) (one step ahead of the boundary condition).

\[
-r \left( \frac{4}{3} C_{f_2}^{n+1} - \frac{1}{3} C_{f_2}^{n+1} \right) + 2(1 + r) C_{f_2}^{n+1} - r C_{f_3}^{n+1} + \frac{1}{2} \, dt \left( k_3 C_{f_2}^{n+1} - k_4 C_{f_2}^{n+1} \right)
\]

\[= r C_{f_1}^{n} + 2(1 - r) C_{f_2}^{n} + r C_{f_3}^{n} - \frac{1}{2} \, dt(k_3 C_{f_2}^{n} - k_4 C_{f_2}^{n})\]

\[2 \left( 1 + \frac{1}{3} r \right) C_{f_2}^{n+1} - \frac{2}{3} r C_{f_3}^{n+1} + \frac{1}{2} \, dt(k_3 C_{f_2}^{n+1} - k_4 C_{f_2}^{n+1})\]

\[= r C_{f_1}^{n} + 2(1 - r) C_{f_2}^{n} + r C_{f_3}^{n} - \frac{1}{2} \, dt(k_3 C_{f_2}^{n} - k_4 C_{f_2}^{n})\]

The equations for \( i > 2 \) follow the generic expression.

Grouping it all together into a matrix form:
\[
B^{n+1} = \begin{pmatrix}
C_{f_2}^{n+1} \\
C_{f_3}^{n+1} \\
\vdots \\
C_{f_{n+1}}^{n+1} \\
C_{b_2}^{n+1} \\
C_{b_3}^{n+1} \\
\vdots \\
C_{b_{nx-1}}^{n+1} \\
C_{b_{nx}}^{n+1}
\end{pmatrix} \\
B^n = \begin{pmatrix}
C_{f_2}^n \\
C_{f_3}^n \\
\vdots \\
C_{f_{nx-1}}^n \\
C_{b_2}^n \\
C_{b_3}^n \\
\vdots \\
C_{b_{nx-1}}^n \\
C_{b_{nx}}^n
\end{pmatrix} \\
D = \begin{pmatrix}
rC_{f_1}^n \\
0 \\
\vdots \\
0 \\
rC_{f_{nx+1}}^n + rC_{f_{nx+1}}^n
\end{pmatrix}
\]

\[
AB^{n+1} = CB^n + D \\
B^{n+1} = A^{-1}(CB^n + D)
\]

The value for \(C_{f_1}^{n+1}\) is calculated after each iteration in time.

3) Rinsing function

This function follows the same principle as the staining one. This time, the boundary condition will be of type Dirichlet of value 0. This way it is possible to model the absence of NPs in the new solution. The NPs will diffuse outside the tissue through the membrane (the new solution creates a gradient) and also inside tissue towards the edge.

It is assumed that the concentration of NPs in the new solution when they diffuse from the tissue is very low, therefore it will be modeled as a constant 0.

\[
C(0, t) = 0 \\
C_{f_1}^{n+1} = 0
\]
The matrix system is identical to the staining one.

The implementation of the algorithm is discussed in the annex, along with the MATLAB scripts.

### 3.3 Backward models

Backward models have been developed by research groups as a method to determine the amount of cancer cell receptors in tissue, a value which is directly proportional to the binding potential (BP). The input for this models is the concentrations for both imaging agents.

This models have a very different behavior depending on the experimental protocol. Using the forward model, we are able to determine what model suits best every protocol.

\[
C_{f_1}^{n+1} = 0
\]

\[
C_{f_{nx+1}}^{n+1} = 0
\]

\[
C_{f_1}^1 = C_0
\]
1) **RPAM (Rinsing Paired-Agent Model)** [8]:

\[
\frac{C_t(t)}{C_t(t_0)} = \left( \frac{C_u(t)}{C_u(t_0)} \right)^{1/BP}
\]

This model was developed by Xiaochun Xu and Kenneth Tichauer while seeking for a simplified model that dealt with a rinsing protocol. It is believed to work good for low values of BP (BP < 1.5).

2) **Convolutional model:**

\[
C_t(t) - C_u(t) = k_3 C_u(t) \ast e^{-(k_4 t)}
\]

\[
BP = \frac{k_3}{k_4}
\]

This model was developed by Kenneth Tichauer and it is intended to be working only for repeated staining. Its name is just provisional since professor Tichauer still has not come with an official name.

3) **Ratiometric model:**

\[
BP = \frac{C_t - C_u}{C_u}
\]

This model is one of the first approaches used for the binding potential estimation. Even though it is not very accurate it is widely used.
3.4 Diffusion effect on BP estimation

In order to see how the changes in the diffusion coefficient change the binding potential estimate value for each model, a certain protocol is selected with a determined target BP and the input parameters are chosen so BP estimate and BP target are close for $D = 1e^{-15}$ (center of the analysis).

Then, the binding potential estimate is calculated for the range $D = (1e^{-16}, 1e^{-14})$. The mean of those values and the standard deviation is calculated. The process is performed for target BP from 1 to 4.

![Figure 15. Effect of D on BP estimation regarding the target BP for repeated staining protocol.](image)

From this figure we can assume that diffusion has a low impact on both models to estimate the binding potential, at least for low values of BP, which is the most common case. The bigger the target binding potential, a greater error will be obtained in the estimation regarding the diffusion coefficient ($D$).

In the other hand, from the rinse approach figure we can see that the variance on the diffusion coefficient ($D$) causes great variances in the binding potential estimated by the RPAM backward model, but not for the ratiometric approach. Therefore, we can conclude that the RPAM model is the most sensitive model to diffusion.
Figure 16. Effect of D on BP estimation regarding the target BP for repeated staining protocol.
3.5 MPSA

MPSA (Multi-parametric sensitivity analysis) is used to determine the relative importance of factors influencing the diffusion of nanoparticles in the binding potential estimation for different backward models.

It is widely used in the parametrization tasks of biological systems models, such as identifying certain parameters for prioritizing research, model reduction, analyzing the robustness of a model and optimal experiment design [9,10].

The inner working of MPSA relies on classifying the model output of a simulation corresponding to the parameter sets into two classes, generally higher or lower than the average output. Samples are sorted according to each parameter and cumulative frequency distributions (CDF) within the two classes (higher or lower than average) are computed.

The biggest difference between these two distributions (called Kolmogorov-Smirnov distance) is a metric indicating how strong the simulation output correlates to that specific parameter (measure of sensitivity). The uncertainties of the parameters are included in the analysis by random selecting values from probability distributions. This is implemented by a Monte Carlo strategy in which the forward model is run repeatedly using those parameter sets. In order to generate evenly distributed samples of the parameter probability distribution, the latin hypercube sampling (LHS) is used, and this way we achieve a better representation of the parameter probability distribution [11].

The procedure used in this MPSA is the following:

1) Carefully select the parameters to be tested. This parameters are $k_3$, $k_4$, $D$, $C_0$, time of stain, time of image/rinse and number of repeats.

2) The range of probabilities from which the samples are extracted are also important, so it is key to give feasible ranges. In the case of the constants, they range from $0.01 \text{ min}^{-1}$ to $0.3 \text{ min}^{-1}$. $D$ ranges from $10^{-16}$ to $10^{-14}$, $C_0$, which control the % of the initial solution going into the tissue, ranges from 1 to 100. Staining time range from 5 to 40 minutes, whereas the image and rinsing time range from 1 to 5 minutes. Finally, the number of repeats range from 3 to 10. The ranges are expressed in the logarithmic domain.
3) For each parameter, generate $n$ independent and random values from the probability distributions. This generates a matrix of sampled parameters (dimensions are $n \times m$, where $m=7$ being the number of parameters, and $n=1000$ (different values). This sampling and combination is performed with the LHS algorithm, implemented by MATLAB through the `lhsdesign` function, achieving the optimal coverage of parameters.

4) Simulate the model for each set of parameters (row of the matrix) to calculate the target outputs. In this case the target outputs are:
   
   a. For the repeated staining protocol, the estimation of the BP using the backward models Convolutional Model and Ratiometric model.
   
   b. For the repeated rinsing protocol, the estimation of the BP using the backward models RPAM and Ratiometric model.

5) Set the average value of the output functions as a threshold. Then, determine whether the parameter set is greater or lesser than the threshold. For each parameter, compute the cumulative frequency distribution (CFD) for both higher and lesser values.

6) The maximum vertical separation from those two CFD is termed Kolmogorov-Smirnov (KS) score and is used as the sensitivity measure of the output with respect to the parameter. The greater the difference, the more sensitive that parameter is towards the output. Hence:

$$KS_i = \sup |S_{>a_i} - S_{>a_i}|$$

Where $S$ represent the CFD functions corresponding to the greater than and lesser than case, for each parameter value $i$.

7) This process is repeated 5 times to be able to achieve a standard deviation value for more accuracy.
Figure 17. K-S score diagram.

1) Repeated staining protocol

Figure 18. K-S score for the repeated staining protocol
In the repeated staining protocol, the convolutional backward model and the ratiometric approach are analyzed. By looking at the figure, we can see that parameters having less impact in the output for both models are the number of repeats, the imaging time and the initial concentration. Also, the diffusion coefficient (D) has a very low impact on the output, meaning that perhaps a forward model accounting for diffusion it is not needed for the assessment of binding potential using the repeated staining protocol. This is a positive result since that means it is possible to use simpler models.

On the other hand, $k_3$ and $k_4$ for both backward models are the most sensitive parameters, which is not a surprise because they have a very direct relation with $BP\ (BP = \frac{k_3}{k_4})$.

2) Repeated rinsing protocol

We clearly see strong similarities with the repeated staining protocol for the initial concentration, number of repeats, initial concentration, and time of staining. However, the clearest difference is the variation of the impact of the diffusion coefficient (D) from the ratio approach to the RPAM model. While D is very low for the ratio approach, it is quite significant for the
RPAM approach, to the point that it is needed to account for this factor if the assessment is calculated using this model.

Additionally, for the ratiometric approach, $k_3$ and $k_4$ have a big impact, but it is curious that whereas $k_4$ was more significant than $k_3$ in the repeated staining protocol, it is the other way around.

3) Single imaging agent (one stain)

![Figure 20. K-S score for single imaging agent concentration](image)

This figure was created to study how each parameter affect the diffusion of the imaging agents regardless of the backward models, especially the diffusion coefficient ($D$), therefore the total quantity of each agent in the tissue is the output.

It is no surprise the high impact of the initial concentration on the total concentration, nor the impact of the staining time. However, $D$ has a high impact on the total concentrations of NPs in the tissue, which is also pretty obvious since that coefficient works regulating a gradient.
The low values of the rate constants for the untargeted agent are because there is a statistical error (very small), since they should be totally 0.

### 3.5 Experimental data

The experimental data used in this thesis was obtained due a collaboration Dr. Tichauer’s lab has with Dr. Jonathan TC Liu, from University of Washington, director of the molecular bio photonics laboratory. I worked with one of his PhD students, Soyoung Kang, in order to obtain the experimental data for the experiments I was analyzing. Additionally, we are co-authoring a paper on the effects of diffusion that is aimed to be published in the Journal of Biomedical Optics.

To obtain the data, a REMI system for intraoperative assessment of surgical margins with a paired-agent imaging protocol with SERS NPs was used.

The experiment consisted in performing a single-staining protocol from different times (3, 6, 9, 12 and 15 minutes) and imaging the tissue for a minute. The protocol is as follows.

1) Stain A431 tumor xenograft with 2-flavor NP mixture.
   a. EGFR-NP conjugated to DyLight 650 (targeted agent).
   b. Isotype-NP conjugated to DyLight 550 (untargeted agent).
2) Fast rinse in PBS (2s) to remove the NPs in the outer surface (not accounted as rinsing in the forward model).
3) Snap freeze in isopentane chilled in liquid nitrogen.
4) Cryosection (10 microns’ slices)
5) Fluorescence microscopy and MATLAB image processing

Data is sent to me in a raw format, consisting in a series of vectors. Each vector corresponds to the values of concentration measured from the surface, in other words, the values of the profile concentration (depth). The experiment is repeated for other areas of the tissue, trying to cover all the surface to obtain a more accurate reading.

The experiments are carried in different sections of resected tissue, therefore the diffusion coefficient will be different in each experiment and it is
something important to have in mind while evaluating the profile concentration graphs (i.e., the concentration for lower staining times can be bigger than for greater staining times).

![Profile concentration in healthy tissue of the untargeted agent](image.jpg)

**Figure 21.** Raw data obtained by UWash describing the profile concentration in healthy tissue for the untargeted imaging agent.

It is observed that the peaks of maximum concentration are not in the surface of the tissue in this graph. This happens because there is a problem with scaling the distance from the camera to the surface of the tissue. Therefore, shifting the peaks to the surface is needed for each different time.

This raw data will be used to make an estimation for the diffusion coefficient and also to tune the parameters of the forward model to be realistic.

Below two figures are displayed showing the profile concentration in healthy and cancerous tissue for both imaging agents. As expected, in the
healthy tissue the concentration of the control and the targeted NPs is very similar.

Figure 22. Profile concentration in healthy tissue for both agents

Figure 23. Profile concentration of both targeted and untargeted imaging agents in a cancerous tissue.
In the other hand, as seen in the figure above, where I trimmed two of the cases to avoid confusion with many representations, when the tissue contains cancer receptor cells, the concentration of the targeted imaging agent is higher than the control agent, which confirms the hypothesis that BP is higher for cancerous tissue than for healthy tissue.

The conclusions obtained from this experiment are:
1) NP penetration is very limited, which helps reduce nonspecific background noise.
2) Binding site barrier may be hindering the transport of NPs at early staining time points, but not as much as time passes.

3.6 D Estimation

After this, all the vectors are averaged to obtain a single vector, which is then scaled so it can be processed by the forward model.
Using the *lsqcurvefit* MATLAB function, the forward model is ran until achieving a low error with the experimental data. The data used correspond only to the control agents, so this way there is no need to fit for $k_3, k_4$. 
The vector to fit is trimmed to a 100 positions vector in which only the exponential decay function is included (the rest of the points do not matter for the diffusion coefficient since they are very close to zero).

Since each experiment was performed on a different section of the resected tumor from the rat, the diffusion coefficient (D) varies from experiment to experiment. D is calculated for each experiment and then the average value and standard deviation is calculated.

1) Healthy tissue:

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>D (m/s²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4.6230E-15</td>
</tr>
<tr>
<td>6</td>
<td>2.0737E-15</td>
</tr>
<tr>
<td>9</td>
<td>4.9272E-15</td>
</tr>
<tr>
<td>12</td>
<td>1.4642E-15</td>
</tr>
<tr>
<td>15</td>
<td>8.0773E-16</td>
</tr>
</tbody>
</table>

Mean (m/s²) 2.78E-15
Standard deviation 1.88E-15

D estimation for 3 minutes of staining

D = 4.62e-15 (m/s²)
Values relating to perfusion rate are very similar in both tissues, meaning that the hypothesis that SERS NPs would diffuse slower in cancer tissue is false.
Figure 26. $D$ fitting for the forward model in 6 and 15 minutes of staining (cancer tissue, control agent).

In both tissues, only the control imaging agent is analyzed, since the targeted agent is very susceptible to $k_3$ and $k_4$. Since the diffusion coefficient should be the same for both agents, it is more accurate to analyze only the control agent.

It can also be appreciate that the forward model shape highly correlates with the experimental data.
Chapter 4

Conclusions and future work

We have seen how the mathematical approach we presented to study the diffusion of the SERS NPs has a strong correlation with the experimental data. Therefore, we assume our approach is correct and significantly accurate.

We can conclude that strong variations of the diffusion coefficient (an order of magnitude away from the realistic values calculated from the experimental data) have a great impact in the binding potential assessment. However, if the variation of D is kept within the mean and standard deviation values calculated in this thesis, it is not a big factor determining the binding potential. Therefore, it is possible to use models not accounting for diffusion and still have accurate values.

The best protocol appears to be the repeated staining protocol because it shows an almost linear relationship between the target BP and estimated BP, for both convolutional and ratiometric models. However, the rinsing protocol is also acceptable since normally the tissue will not have a BP above 1.5, and for certain parameters it works with reliability.

Also, this thesis has output new findings, such as the similarity of the diffusion coefficient for both cancerous and healthy tissue, pointing that the perfusion rate of the SERS NPs is almost the same in both tissues.

In the future, since we have observed a high sensitivity for the rate constants ($k_3$, $k_4$), Tichauer’s lab in conjunction with Liu’s lab will perform a series of flow cytometry experiments to determine the values of those parameters; focusing on how much they change from one tissue to another.

Also, more compartments can be included in the forward model design, such as a compartment accounting for non-specific binding, to accomplish even more accuracy.

From now on, new and old backward models can be tested and validated using the forward model, determine what parameters are the most sensitive in the estimation and analyze the diffusive properties of other
tissues, among others. It is strong, reliable and can be used in other projects involving topical application of SERS NPs.

Also, even though it has not been tested, it might be able to model the diffusion of other particles or biomarkers in intravenous applications. It will be applied in the future to a research project trying to locate certain cells in the eye.
REFERENCES


