Blowing in the Wind:
Regularizations and Outlier Removal

THÈSE N^O 7630 (2017)

PAR

Marta MARTINEZ-CAMARA

EPFL

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2017

To Pavle, who augments my world.
To music, which makes me free.

Acknowledgments

I never would have been able to produce the more than a hundred pages that follow on my own. In fact, there is little that comes just from me on those pages: they are the product of all the interactions I had with the people who surrounded me during the last five years.

Martin, thanks for taking me on board. Doing a PhD was not really in my plans until I met you. Taking that detour was one of the best decision of my life (in spite of Lausanne's cold weather and its lack of good bars). It made me grow in all dimensions: the technical ones and, above all, the human ones. Thank you for never failing to listen to even a single one of my crazy ideas, and for supporting me through all of them. Thanks for your patience with my inability to plan anything more than two days in advance. Thanks for *always* being there. When I grow up I hope to spend my life helping others, like you do.

I would like to thank my co-advisor, Benjamin, for his thought-provoking input which always challenged me and made me see my work from a new angle. His attention to detail compensated for my sometimes less than meticulous approach. Above all, I thank him for his capacity to remain calm and constant even through the roughest of patches.

The problems I worked on during these years were on the crossroads of several fields. This brought me to Norway to work with Andreas Stohl, an incredibly open-minded person that, from the very beginning, believed that engineers can bring interesting contributions to air research. Then, outliers led me to Abdelhak, and Abdelhak took me to Michael. Michael is a researcher/artist who, in spite of working with too many students, also had time to work with me. I still do not know how he does it. I am really happy to have met him and his beautiful family. The portrait that Lillie drew of me always makes me smile and cheers me up.

Thanks a lot to the other members of my thesis committee, Patrick Thiran and Sabine Süsstrunk, who had the patience to read these pages and gave me invaluable feedback.

Alas, there is no official name for the crucial role Paolo played in this thesis. I know many days he did his best to make me not think so, but working with him is a real pleasure. His exquisite taste and attention to detail pushed my skills one level higher. Thanks for challenging me every day with your powerful dialectic, for being my writing and teaching master, and for rescuing me from that black hole. Above all, thanks for (most of the time) believing in me.

I was the luckiest person in the world to have shared these years with the extraordinary human beings that are my colleagues. Mitra, my office mate for most of these years, is a wonder woman and one of the best mathematicians I have ever met (I felt my math skills were kindergarten-level when I worked with her). We grew up together over the years, and she understood me better than anyone else, in both the good and the bad moments. Now she lives far away, and I miss her a lot. Golnoosh now sits in front of me. Golnoosh is starting her PhD, but has a level of maturity that I still hope to attain. She is the one bestowing excellent advice on me much more often than the other way around.

Thanks Ivan for being the source of energy in the lab. You are always pushing yourself to your

boundaries, which inspires us to do the same. Miranda, with whom I share all my best projects, is my partner in crime. The peace that always emanates from Robin helps me compensate my chaotic behaviour. Gilles is the living example of Swiss efficiency, and he always reminds me that being organized takes you farther than being a disaster. Lionel, thank you for bringing good taste into the lab. Niranjan and Zhou always will go in a tangle for me. Amina was my first senior advisor in the lab, and since then, she has been my older sister, guiding me through the adversities and giving me precious pieces of advice. Reza has mastered the art of listening, and he is always there to help others. I would like to thank Adam, Mihailo, Feng, Runwei, Juri, Hanjie, Dalia, Michaelina, Jay, Yann, and Frederike for all the discussions, trips, coffees, and parties we had together.

Thanks Jacqueline and Heidi for making my life easier every day, for teaching me how to navigate through EPFL administration, and helping us handle the boss – the days when you were on holidays and Martin was trying to organize the agenda on his own were quite apocalyptic.

The Tiko team adopted me in Olten during six months. I had lots of fun with them, and I learned a lot. Thomas and Boris are one of the most extraordinary engineers and team workers I have ever met, and I am proud to call them friends now. Thanks Stephane for taking me on a tour of the real world and showing me how business works.

Thanks Guillermo, Laurene, and Vlad from the Swisscom Digital Lab for letting me write this thesis in your amazing office. That, and the free coffee, made me write twice as fast and are the only reason this thesis was submitted on time!

Going back home, I want to thank my oldest friends, Virginia, Contre, Violeta, Vero, Dani, Silvia, Maria, Isa and all the rest. I know that, in spite of the distance, you are just one phone call away. Your power to cheer me up always amazes me. I miss you!

This was like a second PhD for Pavle, but without doing it himself. This is much worse than doing your own PhD. I saw many times the frustration on his face when I was struggling and, of course, I could not let him help me. In addition he had to schlep saxophones all around the world. In spite of all this, he is still by my side and, I hope, he will stay there for many more years. He augments my world, and he makes me a better person.

Last, but not least, I want to thank my family, which includes of course Vitamina, Coco, and Oscar. They are the origin of all the basic values that guide me through life, and they are a source of unconditional love. Special thanks goes to my grandmother Boni, whose system of tying her chickens to a brick gave me a fresh insight on regularizations.

Abstract

Every day tons of pollutants are emitted into the atmosphere all around the world. These pollutants are altering the equilibrium of our planet, causing profound changes in its climate, increasing global temperatures, and raising the sea level. The need to curb these emissions is clear and urgent. To do so, it is first necessary to estimate the quantity of pollutants that is being emitted. Hence, the central challenge of this thesis: how can we estimate the spatio-temporal emissions of a pollutant from many later observations of the concentration of that pollutant at different times and locations?

Mathematically speaking, given such observations and an atmospheric dispersion model, this is a linear inverse problem. Using real datasets, we show that the main difficulties in solving this problem are ill-conditioning and outliers. Ill-conditioning amplifies the effect of additive noise, and each outlier strongly deflects our estimate from the ground truth. We proceed in two different ways to design new estimation methods that can handle these challenges.

In the first approach, we enhance traditional estimators, which are already equipped to deal with ill-conditioning, with a preprocessing step to make them robust against outliers. This preprocessing step blindly localizes outliers in the dataset to remove them completely or to downgrade their influence. We propose two ways of localizing outliers: the first one uses several transport models, while the second one uses random sampling techniques. We show that our preprocessing step significantly improves the performance of traditional estimators, both in synthetic datasets as well as in real-world measurements.

The second approach is based on enhancing existing robust estimators, which are already equipped to deal with outliers, with suitable regularizations, so that they are stable when the problem is ill-conditioned. We analyze the properties of our new estimators and compare them with the properties of existing estimators, showing the advantages of introducing the regularization. Our new estimators perform well both in the presence and in the absence of outliers, making them generally applicable. They have good performance with up to 50 % of outliers in the dataset. They are also stable when the problem is ill-conditioned. We demonstrate their performance using real-world measurements. Two different algorithms to compute the new estimators are given: one is based on an iterative re-weighted least squares algorithm and the other on a proximal gradient algorithm.

Software implementations of all our proposed estimators, along with sample datasets, are provided as part of our commitment to reproducible results. In addition, we provide LinvPy, an open-source python package that contains tested, documented, and user-friendly implementations of our regularized robust algorithms.

Keywords: inverse problems, regularization methods, outlier detection, emissions of pollutants, atmospheric dispersion.

Résumé

Chaque jour, tout autour du monde, des tonnes de polluants sont relâchés dans l'atmosphère. Ces polluants altèrent l'équilibre de notre planète, causent des changements profonds à son climat, augmentent les températures globales et élèvent le niveau de la mer. Le besoin de contrôler ces émissions est clair et urgent. Pour ce faire, il est en premier lieu nécessaire d'estimer la quantité de polluants émis. Ainsi cette thèse propose d'étudier des méthodes d'estimation des émissions spatio-temporelles de polluants, à partir de nombreuses observations de la concentration de ceux-ci à différents instants et endroits.

Mathématiquement parlant, à partir de telles observations et d'un modéle de dispersion atmosphérique, le problème est ramené à un problème inverse linéaire. En utilisant des ensembles de données réelles, nous démontrons que les difficultés principales dans la résolution de ce problème sont le mauvais conditionnement et les valeurs aberrantes. Le mauvais conditionnement amplifie l'effet d'un bruit additif et chaque mesure aberrante dévie notre estimation de la réalité sous-jacente. Nous procédons de deux manières différentes pour concevoir des nouvelles méthodes d'estimation aptes à relever ces défis.

Dans une première approche, nous améliorons des estimateurs existants, qui sont déjà équipés pour lutter contre le mauvais conditionnement avec une étape de traitement antérieur pour les rendre robustes aux mesures aberrantes. Ce traitement localise aveuglément les mesures aberrantes dans l'ensemble des données pour les supprimer ou réduire leur influence. Nous proposons deux méthodes pour identifier les mesures aberrantes, la première utilise plusieurs modèles de transport tandis que la seconde utilise des méthodes d'échantillonnage aléatoire. Nous démontrons que ce traitement antérieur améliore significativement la performance des estimateurs traditionnels, tant sur des ensembles de données synthétiques que sur des mesures réelles.

La deuxième approche consiste à améliorer des estimateurs robustes existants, qui sont déjà capables de tenir compte des mesures aberrantes, avec des régularisations appropriées, de tel sorte qu'ils soient robustes lorsque le problème est mal conditionné. Nous analysons les propriétés de nos nouveaux estimateurs et les comparons aux propriétés d'estimateurs existants, démontrant ainsi l'avantage d'utiliser la régularisation. Nos nouveaux estimateurs fonctionnent bien tant en présence qu'en absence de mesures aberrantes, ce qui les rend généralement applicables. Ils ont de bonne performance avec jusqu'à 50% de mesures aberrantes dans l'ensemble des données. Ils sont aussi stables lorsque le problème est mal conditionné. Nous démontrons leur performance en utilisant des données réelles. Nous donnons deux algorithmes pour calculer ces nouveaux estimateurs: l'un est basé sur un algorithme au moindre carré itératif repondéré, l'autre sur un algorithme de gradient proximal.

La mise en oeuvre logiciel de tous les estimateurs proposés, ainsi que des exemples de données, sont fournis en vertu de notre engagement pour des résultats reproductibles. De plus, nous fournissons LinvPy, un paquet Python source ouvert qui contient des mises en oeuvre testées, documentées et faciles d'utilisation de nos algorithmes robustes régularisés.

Keywords: problème inverse, régularisations, localisation de mesures aberrantes, relâchement de polluants, dispersion atmosphérique.

Contents

Chapter 1

Introduction

Many natural phenomena can be modeled as systems that produce a set of outputs (effects), given a set of inputs (causes). Imagine we observe some outputs, and wish to find the unknown inputs that caused them. Such problems are called *inverse* problems. Their name is meant to differentiate them from forward problems, in which given the causes, we want to predict the effects they will produce.

Recovering inputs may not always be easy. Think for example of the sum of two numbers, $4 + 5 = 9$. If we observe only the result, 9, and someone asks us which numbers were the inputs to the sum operation, we cannot give a unique answer. If we are told that the inputs were both positive integers, the number of possible solutions to this inverse problem is much smaller, but we can still not give a unique answer.

Inverse problems are a broad mathematical topic. The first methods for solving them were proposed by Legendre (1805) and Gauss (1809) [58], and their study is still an active field of research today. They are relevant because of their many applications. These span the whole spectrum of science and engineering, like astronomy [20], oceanography [75], geophysics [12] or signal processing [72]. In this thesis, we will focus on the applications in the air research and atmospheric dispersion community. We will especially focus on estimating releases of substances into the atmosphere.

Every day tons of pollutants are emitted into the atmosphere all around the world. These pollutants are changing the equilibrium of our planet: Earth's average temperature is rising, and the polar ice caps are melting [23, 38]. Also, there are cities with concentrations of solid particles in the air so high, that its inhabitants cannot go outdoor without a mask [24]. Scientists agree that controlling and limiting these emissions is absolutely necessary [56]. Unfortunately, the reaction from politicians and world leaders has been slow, and some of them deny the problem even exists [8].

Pollution and the rising temperature are problems that affect all the living creatures on this planet. The lack of agreement among politicians and leaders has caused concern by millions of anonymous citizens, who have decided to take action. In the last few years, several projects of citizen science were created. One of the pioneers is Safecast [6], founded in 2011, with the goal to create maps of radioactive contamination in Japan. Another example is the environmentalist Ma Jun, who created a real-time open-access pollution database to monitor corporate environmental performance [34]. This data has already provoked changes in the policies of the Chinese government.

The contributions of this thesis are in methods, rather than datasets themselves: if we have access to many observations of the concentration of a given pollutant at different times and locations, how can we estimate the spatio-temporal emissions that caused them? We will describe this inverse problem formally, as well as propose and discuss solutions to the difficulties that arise. The initial motivation for this project was the estimation of emissions of radioactive material from a known location. But, in fact, this problem can be easily extended to emissions in unknown places, as well as for different materials like CO_2 or other pollutants.

1.1   Linear inverse problem: setup

Let us note a position in space with a vector $\mathbf{r}$ and a moment in time with a scalar t. Next, suppose a source $s(\mathbf{r}, t)$ emits a certain material which is then transported by the atmosphere around the world, as illustrated in Figure 1.1. A particle in position $\mathbf{r}'$ at time t' has a probability $p(\mathbf{r}, t | \mathbf{r}', t')$ of being at position $\mathbf{r}$ at time t. Thus, the concentration y of the material at a certain point $\mathbf{r}$ and time t can be expressed as

$$\mathbb{E}\{y(\mathbf{r}, t)\} = \int_{-\infty}^{t} \int p(\mathbf{r}, t | \mathbf{r}', t') s(\mathbf{r}', t') d\mathbf{r}' dt', \tag{1.1}$$

where $\mathbb{E}\{\cdot\}$ is the expected value.

Figure 1.1: Simplified representation of the atmospheric dispersion.

The distribution $p(\mathbf{r}, t | \mathbf{r}', t')$ is in general unknown. One way of approximating it is to use virtual particles with realistic, randomly generated trajectories. At each $\mathbf{r}$ and t, a unit mass of these particles is released. Then, sample averages of them for every $\mathbf{r}, t$ are taken. The virtual particles are called Lagrangian particles, and this transport model is called the Lagrangian Particle Dispersion Model (LPDM) [78]. Figure 1.2 shows a representation of these models.

LPDMs owe their name to the reference system that they use: the observer follows each individual particle as it moves through space and time.

Another possibility to discretize the system (1.1) is to use Eulerian models [77]. The Eulerian point of view focuses on specific locations in space, through which materials flow over time. Eulerian models can be visualized as sitting on the bank of a river and watching the water pass by. Lagrangian models can be represented as sitting in a boat and drifting down a river.

LPDMs have some advantages with respect to the Eulerian ones: they can have infinite temporal and spatial resolution; they avoid the artificial initial diffusion of a point source in the corresponding cell and in the advection numerical errors; and they are computationally more efficient [78].
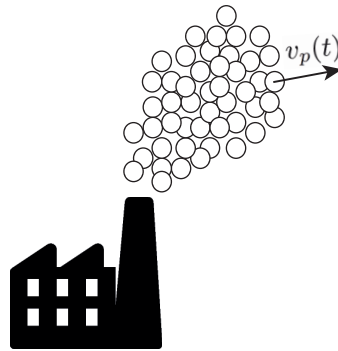

Figure 1.2: The drawing illustrates the discretization of a plume using virtual particles. Each particle is moved by pseudovelocity $v_p(t)$ at each time step. These velocities are computed to take into account the transport caused by the average wind and the turbulence terms due to wind fluctuations.

Using LPDM, Eq.(1.1) is spatially and temporarily discretized, becoming a linear system of equations, and thus, a discrete linear inverse problem

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{1.2}$$

Here $\mathbf{x} \in \mathbb{R}^{n \times 1}$ is the source term, i.e. a spatially and temporally discretised form of $s(\mathbf{r}, t)$; $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the transport model, i.e. the discretised form of $p(\mathbf{r}, t | \mathbf{r}', t')$; and $\mathbf{y} \in \mathbb{R}^{m \times 1}$ are the measurements, i.e. the discretised form of $\mathbb{E}\{y(\mathbf{r}, t)\}$.

There are only a few freely available, open source implementations of LPDMs. The Flexible Particle dispersion model (FLEXPART) [63] is one of them. It has been used and validated in a large number of studies on long-range atmospheric transport [64]. As stated in [65]:

This model was originally developed for calculating the dispersion of radioactive material from nuclear emergencies but has since been used for many other applications as well. Nuclear applications include, for instance, simulations of the transport of radioactive material from nuclear power plants and other facilities [10, 74] or from nuclear bomb tests [15]. FLEXPART is also the model operationally used at CTBTO [27] for atmospheric backtracking and at the Austrian Central Institute for Meteorology and Geodynamics for emergency response as well as CTBT verification purposes.

In this thesis we use FLEXPART to derive $\widehat{\mathbf{A}}$, which is an estimate of the unknown true transport matrix $\mathbf{A}$.

In reality, the actual measurements $\widehat{\mathbf{y}}$ are corrupted by measurement noise $\mathbf{n} \in \mathbb{R}^{m \times 1}$, so that $\widehat{\mathbf{y}} = \mathbf{A}\mathbf{x} + \mathbf{n}$. There are also errors in the generated trajectories which appear as model errors $\mathbf{N} \in \mathbb{R}^{m \times n}$ (that we assume as additive), such that $\widehat{\mathbf{A}} = \mathbf{A} + \mathbf{N}$. Therefore, the problem is typically formulated as

$$\widehat{\mathbf{y}} = \widehat{\mathbf{A}}\mathbf{x} + \mathbf{e}, \tag{1.3}$$

where $\mathbf{e} = \mathbf{n} - \mathbf{N}\mathbf{x}$ is the total error. Recall that the source $\mathbf{x}$ is unknown.

Our goal in this thesis is to use the information available to us, $\widehat{\mathbf{y}}$ and $\widehat{\mathbf{A}}$, to get an estimate $\widehat{\mathbf{x}}$ as close as possible to the ground truth $\mathbf{x}$ in the Mean Squared Error (MSE) sense, where $\mathrm{MSE}(\widehat{\mathbf{x}}) := \mathbb{E}[(\widehat{\mathbf{x}} - \mathbf{x})^2]$. This, at first sight a simple problem, involves many challenges.

1.2   Ill-conditioned Problem

In 1902 Hadamard proposed the conditions he thought a problem should have to be well-posed [36]. These are known as the Hadamard conditions, and they are still used today. They can be summarized as follows. A problem is well-posed if:

1. a solution exists,

2. the solution is unique, and

3. the solution depends continuously on the data.

If any of these conditions are not achieved, the problem becomes ill-posed.

In atmospheric dispersion problems, the model matrix $\widehat{\mathbf{A}}$ is typically tall, i.e. we have more observations than unknowns to estimate. The classic method to solve inverse problems of this type is least squares (LS), which will be explained in the next section. However, there is another important aspect in estimating emissions into the atmosphere: the matrix $\widehat{\mathbf{A}}$ typically has a large condition number. This makes the LS estimate very sensitive to additive errors (a precise bound for this sensitivity is given in (1.8)).

Figure 1.3 shows an example. The model matrix has a condition number equal to 1000, which is relatively high. Although the additive noise is quite mild - the signal-to-noise ratio (SNR) is 10dB - the LS estimate is obviously not able to recover the real source term well.

This high sensitivity to noise violates the third Hadamard condition, and makes the problem ill-posed.

To tackle an ill-conditioned problem, it is often fruitful to consider a slightly modified version of the same, where extra a priori information is introduced in the main equation. This approach is known as regularization, and many types of regularization techniques have been successfully designed and used over the years. In the next section we introduce and review the key regularization techniques that will be used in this thesis.
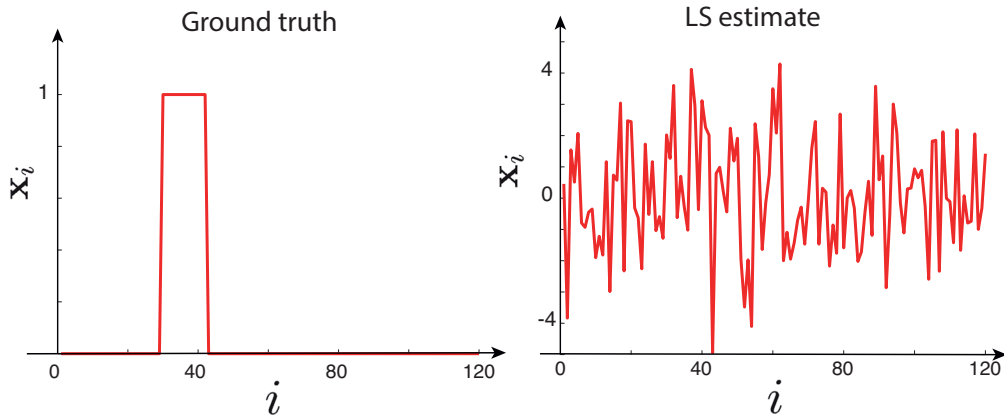
Figure 1.3: Experiment that shows the sensitivity of the LS estimate to mild additive noise when the condition number of the model matrix is large. Here, we use $\mathbf{A} \in \mathbb{R}^{300 \times 120}$ with a condition number equal to 1000. The SNR of the additive Gaussian noise with respect to the measurements is 10dB. On the left of this figure we can observe the ground truth $\mathbf{x}$. Here $\mathbf{x}_i$ is the i-th element of $\mathbf{x}$. On the right we see the LS estimate for this problem. The large condition number of the model matrix makes the system very sensitive to the noise.

1.3 Regularizations: A Geometric Perspective

In this section we will assume that there are no errors in the model matrix, i.e. $\widehat{\mathbf{A}} = \mathbf{A}$ and so $\widehat{\mathbf{y}} = \mathbf{A}\mathbf{x} + \mathbf{n}$. The vector $\mathbf{n}$ represents additive errors, and is unknown. We also assume that $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a tall (i.e. $m > n$) and full-rank matrix.

As we said before, LS is probably simplest and most widely used method to solve linear inverse problems. The LS estimate $\widehat{\mathbf{x}}_{\mathrm{LS}}$ minimizes the ℓ_2 norm between $\widehat{\mathbf{y}}$ and $\mathbf{y}(\mathbf{x}) = \mathbf{A}\mathbf{x}$,

$$\widehat{\mathbf{x}}_{\mathrm{LS}} = \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \mathbf{y}(\mathbf{x})\|_2^2 = \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2. \tag{1.4}$$

Let us define the column space of $\mathbf{A}$, $C(\mathbf{A})$, as the span of the columns of $\mathbf{A}$. Then $\|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2$ is the distance between $\widehat{\mathbf{y}}$ and the subspace $C(\mathbf{A})$. From this geometrical point of view, searching for $\widehat{\mathbf{x}}_{\mathrm{LS}}$ is the same as looking for the point $\mathbf{y}(\mathbf{x}_{\mathrm{LS}})$ in the subspace $C(\mathbf{A})$ that is closer to $\widehat{\mathbf{y}}$ than any other point in $C(\mathbf{A})$. The error vector $\widehat{\mathbf{y}} - \mathbf{y}(\mathbf{x}_{\mathrm{LS}})$ is thus perpendicular to the subspace $C(\mathbf{A})$ (see Figure 1.4). This orthogonal error vector is produced by the orthogonal projection of $\widehat{\mathbf{y}}$ onto $C(\mathbf{A})$, which is given by [66]

$$\mathbf{y}(\widehat{\mathbf{x}}_{\mathrm{LS}}) = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\widehat{\mathbf{y}}, \tag{1.5}$$

where $\mathbf{A}^T$ is the transpose of $\mathbf{A}$. From here it is easy to identify $\widehat{\mathbf{x}}_{\mathrm{LS}}$: we know that $\mathbf{y}(\widehat{\mathbf{x}}_{\mathrm{LS}}) = \mathbf{A}\widehat{\mathbf{x}}_{\mathrm{LS}}$, so

$$\widehat{\mathbf{x}}_{\mathrm{LS}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\widehat{\mathbf{y}}. \tag{1.6}$$

The matrix $\mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ is known as the pseudo-inverse of $\mathbf{A}$. Hence, we can also write $\widehat{\mathbf{x}}_{\mathrm{LS}}$ as

$$\widehat{\mathbf{x}}_{\mathrm{LS}} = \mathbf{A}^+\widehat{\mathbf{y}}. \tag{1.7}$$
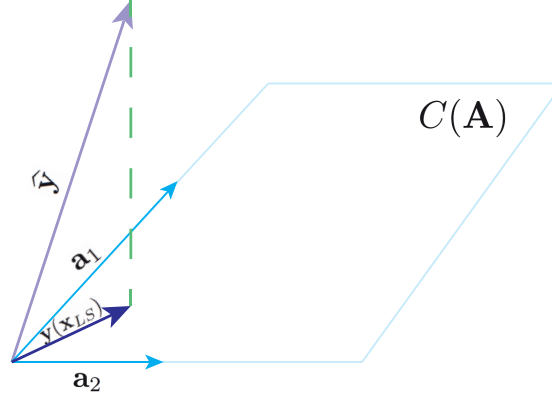
Figure 1.4: $\mathbf{y}(\mathbf{x}_{LS})$ is the orthogonal projector of the vector $\widehat{\mathbf{y}}$ onto the column space of $\mathbf{A}$, $C(\mathbf{A})$. In this example $\mathbf{A} \in \mathbb{R}^{3\times 2}$, and $\mathbf{a}_1, \mathbf{a}_2$ are the columns of $\mathbf{A}$.

To visualize this estimator in the context of linear inverse problems and regularizations, we build another low dimensional toy example with $\mathbf{A} \in \mathbb{R}^{3\times 2}$ as shown in Figure 1.5a. Let us assume that $\mathbf{x}$ lives in $\mathbb{R}^2_+$, the positive orthant of $\mathbb{R}^2$. The matrix $\mathbf{A}$ transforms the red cone $\mathbb{R}^2_+$ into the blue cone Y. Thus Y is contained in $C(\mathbf{A})$, and all the $\mathbf{y}(\mathbf{x})$ such that $\mathbf{x} \geq 0$ reside in Y. The blue vectors that delimit Y are the columns of $\mathbf{A}$: they are the images of the vectors that delimit $\mathbb{R}^2_+$, which are the unitary vectors $(1, 0)^T$ and $(0, 1)^T$.

In this figure we can see two examples of sources, $\mathbf{x}_1$ and $\mathbf{x}_2$, and their corresponding images in Y, $\mathbf{y}_1$ and $\mathbf{y}_2$. Notice that $\mathbf{y}_1$ and $\mathbf{y}_2$ are noiseless. The two arrows marked $\mathbf{A}$ and $\mathbf{A}^+$ indicate that, to go from $\mathbb{R}^2_+$ to Y we use $\mathbf{A}$, i.e. $\mathbf{y}_1 = \mathbf{A}\mathbf{x}_1$, but to go from Y to $\mathbb{R}^2_+$ we use $\mathbf{A}^+$, i.e. $\mathbf{x}_1 = \mathbf{A}^+\mathbf{y}_1$.

In a real problem very often we do not have access to the noiseless $\mathbf{y}_1$ and $\mathbf{y}_2$: we just know their noisy version, $\widehat{\mathbf{y}}_1$ and $\widehat{\mathbf{y}}_2$. The noisy $\widehat{\mathbf{y}}_1$ and $\widehat{\mathbf{y}}_2$ (not shown for simplicity) lie in the 3-dimensional *fuzzy* balls centered at the noiseless $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively, since we assume the noise to have mean zero. These balls represent the distribution of the noise $\mathbf{n}$. Their radius represent the standard deviation of the noise. We are assuming that this noise is i.i.d. These 3-dimensional balls intersect with $C(\mathbf{A})$. As $C(\mathbf{A})$ is a 2-dimensional subspace, the intersections between the balls and $C(\mathbf{A})$ are 2-dimensional slices, and they are shown as dark blue circles in Figure 1.5b. Notice that the blue circles can lay outside Y, like in the case of the circle centred at $\mathbf{y}_1$. The orthogonal projections of $\widehat{\mathbf{y}}_1$ and $\widehat{\mathbf{y}}_2$, $\mathbf{y}(\widehat{\mathbf{x}}_{1LS})$ and $\mathbf{y}(\widehat{\mathbf{x}}_{2LS})$, will lay inside the dark blue circles. Again, $\mathbf{y}(\widehat{\mathbf{x}}_{1LS})$ and $\mathbf{y}(\widehat{\mathbf{x}}_{2LS})$ may or may not lay inside Y.

Once we have $\mathbf{y}(\widehat{\mathbf{x}}_{1LS})$ and $\mathbf{y}(\widehat{\mathbf{x}}_{2LS})$, we use $\mathbf{A}^+$ to find $\widehat{\mathbf{x}}_{1LS}$ and $\widehat{\mathbf{x}}_{2LS}$. As we see in Figure 1.5b, $\widehat{\mathbf{x}}_{1LS}$ lies inside the cone $\mathbb{R}^2_+$ but $\widehat{\mathbf{x}}_{2LS}$ does not.
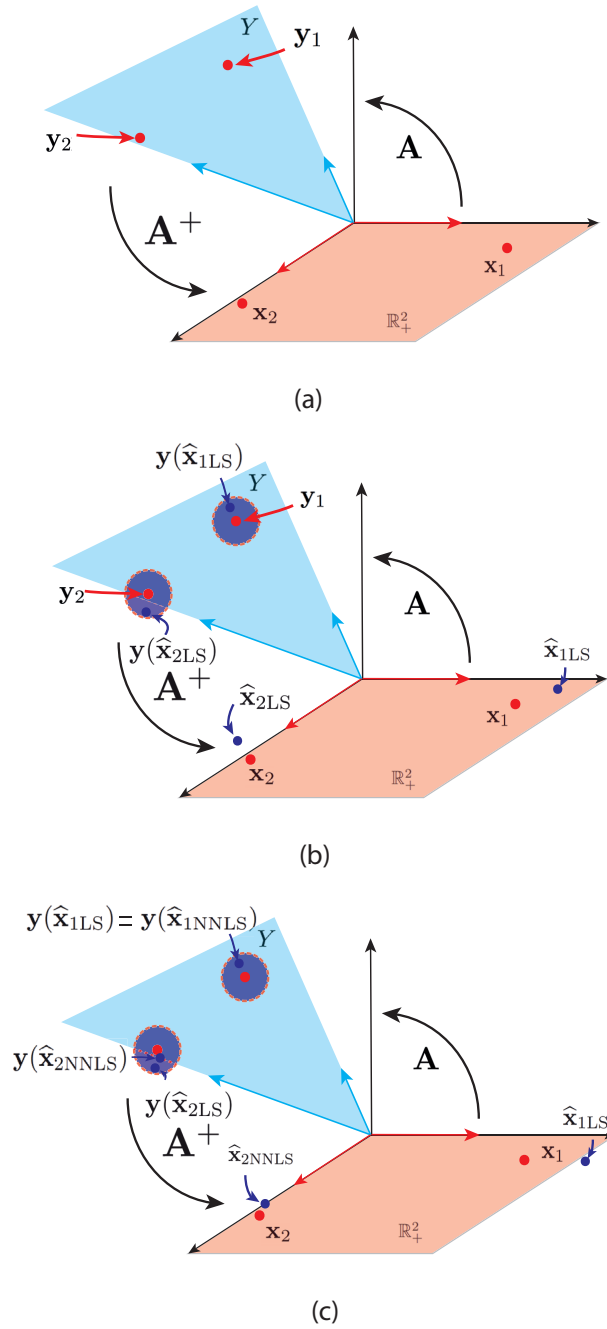
(a)

(b)

(c)

Figure 1.5: These figures show how least squares and the non-negative regularization work in a linear inverse problem $\widehat{\mathbf{y}} = \mathbf{A}\mathbf{x} + \mathbf{n}$ with additive Gaussian noise. The subfigure (a) shows the general setup. Here $\mathbf{A} \in \mathbb{R}^{3\times2}$. We assume $\mathbf{x} \geq 0$, so it lives in the non-negative 2-dimensional orthant. $\mathbf{y}_1$ and $\mathbf{y}_2$ are two particular examples of noiseless $\mathbf{y}$ such that $\mathbf{y}_1 = \mathbf{A}\mathbf{x}_1$ and $\mathbf{y}_2 = \mathbf{A}\mathbf{x}_2$. They live in the 2-dimensional cone Y. This cone is inside the 2-dimensional column space of $\mathbf{A}$, $C(\mathbf{A})$. Subfigure (b) shows the LS estimates given in (1.4), $\widehat{\mathbf{x}}_{1\mathrm{LS}}$ and $\widehat{\mathbf{x}}_{2\mathrm{LS}}$. This plot is an extension of Figure 1.4. To produce these estimates, LS has access to the noisy $\widehat{\mathbf{y}}_1$ and $\widehat{\mathbf{y}}_2$ which live in a *fuzzy* 3-dimensional balls centered at the noiseless $\mathbf{y}_1$ and $\mathbf{y}_2$. $\mathbf{y}(\widehat{\mathbf{x}}_{1\mathrm{LS}})$ and $\mathbf{y}(\widehat{\mathbf{x}}_{2\mathrm{LS}})$ are the orthogonal projections of $\widehat{\mathbf{y}}_1$ and $\widehat{\mathbf{y}}_2$ into $C(\mathbf{A})$. Subfigure (c) shows the effect of adding the non-negative constrain to LS as it is given in (1.10). If $\widehat{\mathbf{y}}(\widehat{\mathbf{x}}_{\mathrm{LS}})$ lays inside Y, as it is the case of $\mathbf{y}(\widehat{\mathbf{x}}_{1\mathrm{LS}})$, the corresponding NNLS estimate $\widehat{\mathbf{x}}_{1\mathrm{NNLS}}$ is identical to the LS estimate $\widehat{\mathbf{x}}_{1\mathrm{LS}}$. However, if $\widehat{\mathbf{y}}(\widehat{\mathbf{x}}_{\mathrm{LS}})$ lays outside Y, as it is the case for $\mathbf{y}(\widehat{\mathbf{x}}_{2\mathrm{LS}})$, the non-negative constraint take it back to Y producing $\mathbf{y}(\widehat{\mathbf{x}}_{2\mathrm{NNLS}})$, which gives $\widehat{\mathbf{x}}_{2\mathrm{NNLS}}$, a different estimate from LS.

Unfortunately, as we discussed before in Figure 1.3, the LS estimate is very sensitive to the noise $\mathbf{n}$ when the condition number of $\mathbf{A}$ is large. If we define the condition number of $\mathbf{A}$ as $\kappa(\mathbf{A})$, this sensitivity can be bound as [66]

$$\frac{\mathbf{x} - \widehat{\mathbf{x}}_{\text{LS}}}{\mathbf{x}} \leq \kappa(\mathbf{A}) \frac{(\mathbf{y} - \widehat{\mathbf{y}})}{\mathbf{y}}. \tag{1.8}$$

This means that, even if $\|\mathbf{y} - \widehat{\mathbf{y}}\|$ is small, $\|\mathbf{x} - \widehat{\mathbf{x}}_{\text{LS}}\|$ can be very large. This lack of stability violates the third Hadamard condition, and it makes the problem ill-posed. To stabilize the problem, more a priori information about the solution should be introduced. There exist several techniques for this. They are called regularizations.

1.3.1 Non-Negative Least Squares

The first fact about the solution that we can introduce into the problem is that, in general, the emissions into the atmosphere are non-negative[1]. In our toy example, this means that the source $\mathbf{x}$ lives in the red cone $\mathbb{R}_+^2$. This information can be easily added to the problem as a linear constraint

$$\widehat{\mathbf{x}}_{\text{NNLS}} = \ \underset{\mathbf{x}}{\arg\min} \ \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2, \tag{1.9}$$
$$\text{s.t.} \ \ \mathbf{x} \succeq 0$$

where $\succeq$ is the elementwise inequality. This new estimator is called Non-Negative Least Squares (NNLS), and we can consider it as the first regularization in our list.

Geometrically, Figure 1.5c shows what the linear constraint does. If $\mathbf{y}(\widehat{\mathbf{x}}_{\text{LS}})$ lies inside the cone Y, like for example $\mathbf{y}(\widehat{\mathbf{x}}_{\text{1LS}})$, the solution given by NNLS $\mathbf{y}(\widehat{\mathbf{x}}_{\text{1NNLS}})$ is equal to the $\mathbf{y}(\widehat{\mathbf{x}}_{\text{1LS}})$ given by LS. Otherwise, NNLS takes the $\mathbf{y}(\widehat{\mathbf{x}}_{\text{LS}})$ that lay outside the cone back to the cone Y to find $\mathbf{y}(\widehat{\mathbf{x}}_{\text{NNLS}})$. Examples of this are $\mathbf{y}(\widehat{\mathbf{x}}_{\text{2LS}})$ and $\mathbf{y}(\widehat{\mathbf{x}}_{\text{2NNLS}})$. All the approximations that are *taken back* to the cone lie at the edge of the cone. The corresponding $\widehat{\mathbf{x}}_{\text{2NNLS}}$ also lies on the edge of the positive orthant $\mathbb{R}_+^2$. This means that these solutions are sparse, i.e. most of their entries are equal to zero. This intuitively confirms that the non-negative constraint tends to produce sparse $\widehat{\mathbf{x}}$ [31]. Notice that, unlike LS, the NNLS estimate does not have an analytic expression, and must be computed using convex optimization iterative algorithms [17].

1.3.2 Non-Negative Tikhonov Regularization

Another realistic assumption on the source is to have a limited amount of energy. Mathematically this means that $\mathbf{x}$ has a small ℓ_2 norm. One way of introducing this information into the problem is by adding a penalty term:

$$\widehat{\mathbf{x}}_{\text{TIK}} = \ \underset{\mathbf{x}}{\arg\min} \ \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_2^2, \tag{1.10}$$

where $\lambda \geq 0$ is the regularization parameter.

This regularization is called Tikhonov regularization. It favours solutions that contain many small elements, like the vector $\mathbf{a}$ represented in Figure 1.6, rather than solutions that contain a few big elements, like the vector $\mathbf{b}$ showed in the same figure: $\mathbf{a}$ and $\mathbf{b}$ have the same ℓ_1 norm, $\|\mathbf{a}\| = \|\mathbf{b}\|$; however $\|\mathbf{b}\|_2$ is larger than $\|\mathbf{a}\|_2$.

---

1. This is not true in the case of CO_2 emissions, but we will not consider that case here.
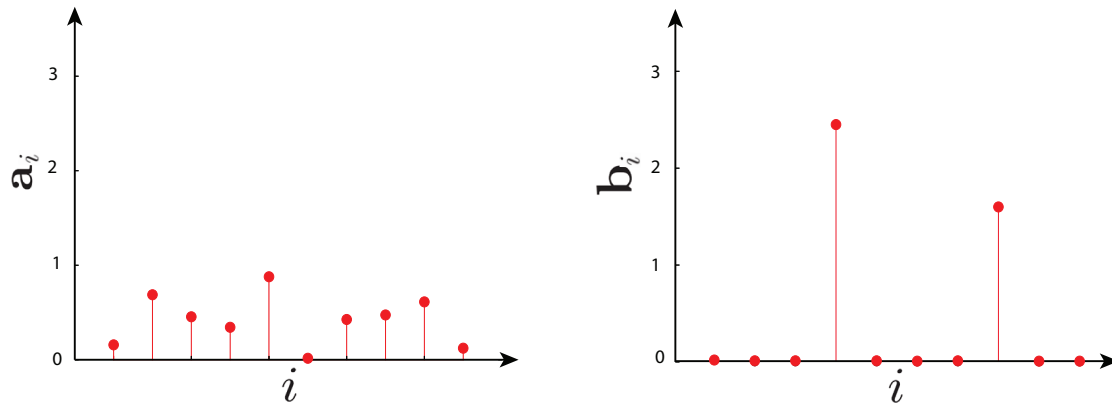
Figure 1.6: The two vectors $\mathbf{a}$ and $\mathbf{b}$ have the same ℓ_1 norm, $\|\mathbf{a}\|_1 = \|\mathbf{b}\|_1$. However the ℓ_2 norm of $\mathbf{a}$ is larger than the ℓ_2 norm of $\mathbf{b}$, $\|\mathbf{a}\|_2 < \|\mathbf{b}\|_2$.

Tikhonov is one of the most commonly used regularization methods. The reason is that, apart from its generality, this estimator has an analytic solution, which makes it easy to compute. One way of deducing this solution is to transform Equation (1.10) into an LS problem

$$\widehat{\mathbf{x}}_{\text{TIK}} = \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_2^2 = \arg\min_{\mathbf{x}} \left\| \begin{pmatrix} \widehat{\mathbf{y}} \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \mathbf{A} \\ \sqrt{\lambda}\mathbf{I} \end{pmatrix} \mathbf{x} \right\|_2^2. \tag{1.11}$$

Thus, we can use Eq. (1.6) to get the analytic expression for the Tikhonov estimate

$$\widehat{\mathbf{x}}_{\text{TIK}} = \left(\begin{pmatrix} \mathbf{A} \\ \sqrt{\lambda}\mathbf{I} \end{pmatrix}^T \begin{pmatrix} \mathbf{A} \\ \sqrt{\lambda}\mathbf{I} \end{pmatrix} \right)^{-1} \begin{pmatrix} \mathbf{A} \\ \sqrt{\lambda}\mathbf{I} \end{pmatrix} \begin{pmatrix} \widehat{\mathbf{y}} \\ \mathbf{0} \end{pmatrix} = (\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I})^{-1}\mathbf{A}\widehat{\mathbf{y}}. \tag{1.12}$$

Tikhonov regularization can be combined with the non-negative constraint,

$$\widehat{\mathbf{x}}_{\text{NNTIK}} = \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_2^2. \tag{1.13}$$
$$\text{s.t. } \mathbf{x} \succeq 0$$

We call this regularization non-negative Tikhonov regularization. Figure 1.7 shows how this new objective function affects the estimate $\widehat{\mathbf{x}}_{\text{NNTIK}}$: Figure 1.7a shows one example of a NNLS estimate. The circular arcs in $\mathbb{R}^2_+$ shown in 1.7b represent the set of $\mathbf{x}$s with the same ℓ_2 norm. These arcs are transformed by $\mathbf{A}$ into the elliptic arcs in Y. To make $\|\mathbf{x}\|_2$ small, the regularization moves $\mathbf{y}(\widehat{\mathbf{x}}_{\text{NNLS}})$ perpendicularly to the arcs in Y, as shown in 1.7c, towards the inside of the cone, to get to $\mathbf{y}(\widehat{\mathbf{x}}_{\text{NNTIK}})$. The higher the λ, the more movement there is. Eventually as λ grows, $\mathbf{y}(\widehat{\mathbf{x}}_{\text{NNTIK}}) \to \mathbf{0}$.

We note that the Tikhonov regularization makes many elements of $\widehat{\mathbf{x}}_{\text{NNTIK}}$ to be non-zero because it moves $\mathbf{y}(\widehat{\mathbf{x}}_{\text{NNTIK}})$ towards the *center* of the cone. The values of $\widehat{\mathbf{x}}_{\text{NNTIK}}$ become smaller as λ increases.

Notice that, as in the case of the NNLS estimate, there is no analytic solution for the non-negative Tikhonov estimate. Again, iterative convex optimization algorithms must be used to compute the estimate [17].

Another variation of the Tikhonov regularization can be used to enforce smoothness in the solution. To do so, the discrete differential operator $\mathbf{D}$ is introduced in the objective function

$$\widehat{\mathbf{x}}_{\mathrm{DTIK}} = \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{D}\mathbf{x}\|_2^2. \tag{1.14}$$

Using the same technique as in (1.11), it is straighforward to derive the analytic solution for this estimator

$$\widehat{\mathbf{x}}_{\mathrm{DTIK}} = (\mathbf{A}^T\mathbf{A} + \lambda\mathbf{D}^T\mathbf{D})^{-1}\mathbf{A}\widehat{\mathbf{y}}. \tag{1.15}$$

This regularization favours solutions whose discrete derivative has many non-zero components with relatively small values. Thus, it avoids solutions with large changes.

1.3.3   Non-Negative Sparse Regularization

In some cases – notably when the releases are produced during explosions – we can assume that the source is sparse, i.e. only a few of its elements are non-zero. This is exactly the opposite of what the Tikhonov regularization enforces. Thus, we need a different technique for this kind of solutions.

The ℓ_0 norm of a vector is defined as the number of entries of that vector that are different from zero. Thus, if a solution is sparse, then its ℓ_0 norm is also small. Hence, we can use this norm to enforce sparsity. We define a new objective function as

$$\widehat{\mathbf{x}}_{\ell_0} = \quad \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0. \tag{1.16}$$

But there is a problem with (1.16): it is not convex. This means that we cannot use convex optimization algorithms to compute $\widehat{\mathbf{x}}_{\ell_0}$. In fact, in general, the only way of minimizing (1.16) is in a combinatorial fashion. Hence, even problems with a moderate number of dimensions in $\mathbf{x}$ are already intractable. Luckily, under certain conditions (see [21] for more details), the problem can be relaxed, and solving (1.16) is equivalent to solving

$$\widehat{\mathbf{x}}_{\ell_1} = \quad \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1. \tag{1.17}$$

Solving the proxy given by (1.17) leads to the same solution as solving (1.16). The former is indeed convex, and can be solved efficiently using convex optimization tools.

Again, as we did in the case of Tikhonov, we can also add the non-negative constraint to enforce non-negative emissions

$$\widehat{\mathbf{x}}_{\mathrm{NN}\ell_1} = \quad \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1.$$
$$\text{s.t. } \mathbf{x} \succeq 0 \tag{1.18}$$

Figure 1.8 shows how adding the ℓ_1 regularization changes the NNLS estimate. Figure 1.8a shows the sparse ground-truth $\mathbf{x}_1$ and its NNLS estimate $\widehat{\mathbf{x}}_{\mathrm{NNLS}}$. In Figure 1.8b you can see
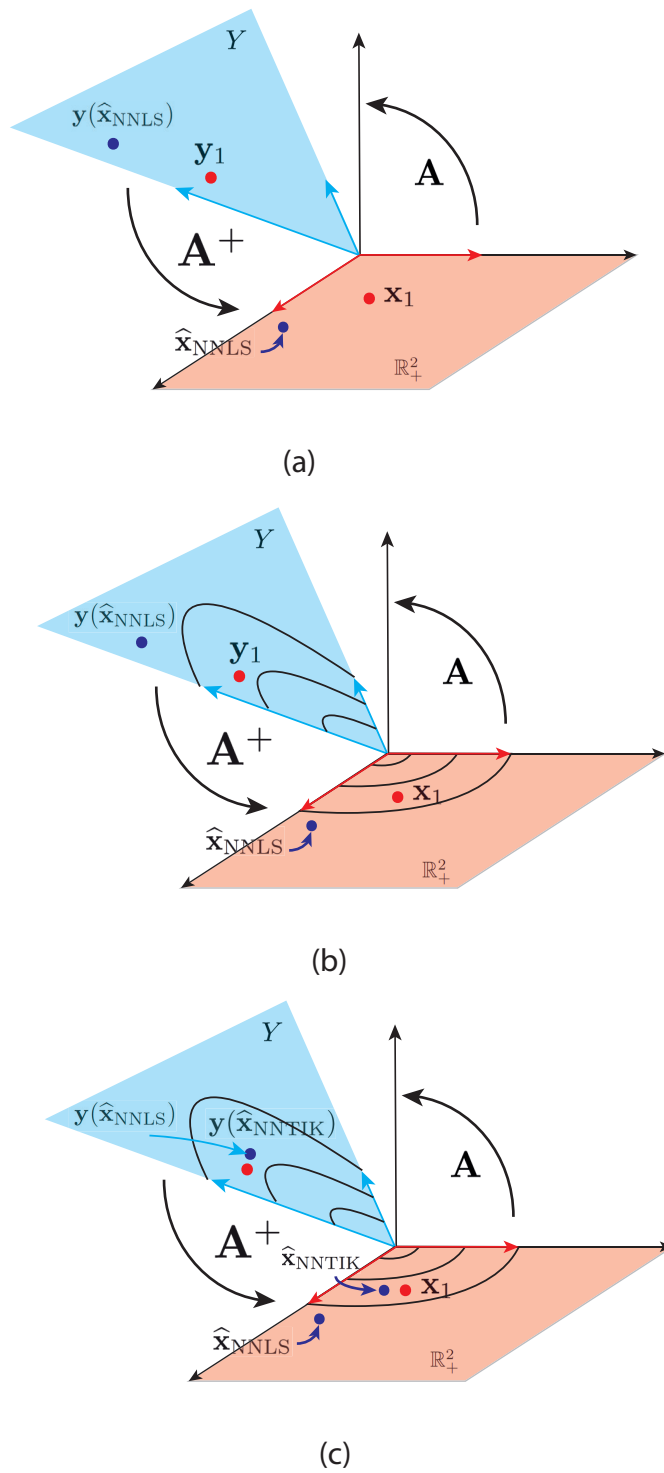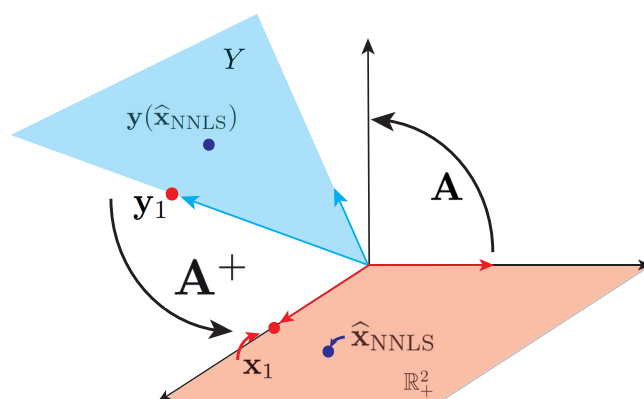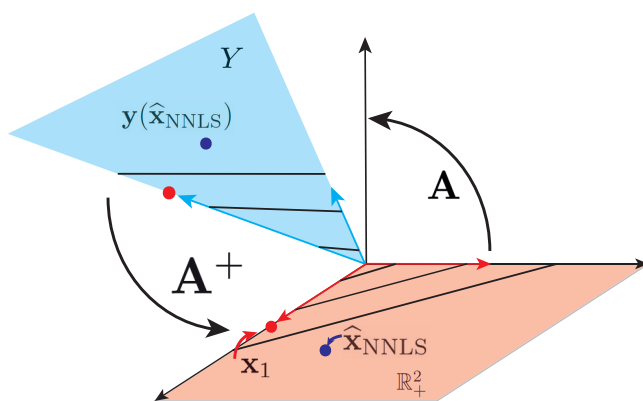
(a)

(b)

(c)

Figure 1.7: These figures show the effect of the non-negative Tikhonov regularization given in (1.13) in a linear inverse problem $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$ with additive Gaussian noise. The setup is the same as in Figure 1.5. Recall that $\mathbf{A} \in \mathbb{R}^{3 \times 2}$. Subfigure (a) shows an example of the ground-truth $\mathbf{x}_1$ and its NNLS estimate $\widehat{\mathbf{x}}_{\mathrm{NNLS}}$. The black circles in $\mathbb{R}^2_+$ in Subfigure (b) represent the $\mathbf{x}$s with the same ℓ_2 norm. The black ellipses in Y are the images of the black arcs in $\mathbb{R}^2_+$. Subfigure (c) shows how Tikhonov regularization looks for solutions with a small ℓ_2 norm, so it moves $\mathbf{y}(\widehat{\mathbf{x}}_{\mathrm{NNLS}})$ perpendicularly to the black arcs as λ grows, producing the non-negative Tikhonov estimate $\widehat{\mathbf{x}}_{\mathrm{NNTIK}}$.
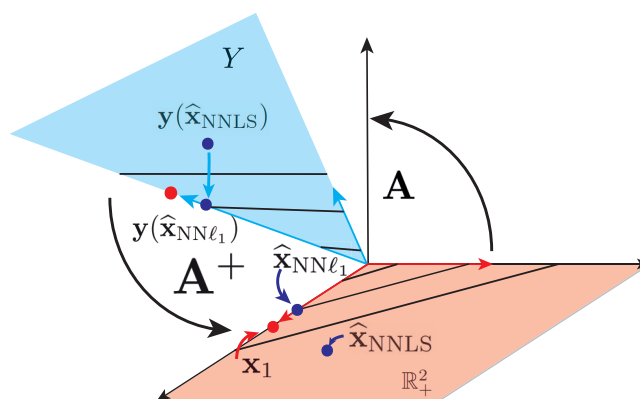
(a)

(b)

(c)

Figure 1.8: These figures show the effect of the non-negative sparse regularization given in (1.18) in a linear inverse problem $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$ with additive Gaussian noise. The setup is the same as in Figure 1.5. Recall that $\mathbf{A} \in \mathbb{R}^{3 \times 2}$. Subfigure (a) shows an example of a sparse solution $\mathbf{x}_1$ and its NNLS estimate $\widehat{\mathbf{x}}_{\mathrm{NNLS}}$. The black lines in $\mathbb{R}^2_+$ in Subfigure (b) represent the $\mathbf{x}$s with the same ℓ_1 norm. The black lines in Y are the images of the black lines in $\mathbb{R}^2_+$. Subfigure (c) shows how the sparse regularization looks for solutions with a small ℓ_1 norm, so it moves $\mathbf{y}(\widehat{\mathbf{x}}_{\mathrm{NNLS}})$ perpendicularly to the black lines as λ grows, producing the non-negative sparse estimate $\widehat{\mathbf{x}}_{\mathrm{NN}\ell_1}$.

the straight lines in $\mathbb{R}_+^2$ that represent the $\mathbf{x}$s with the same ℓ_1 norm. Because of the linearity of the operator $\mathbf{A}$, the images of these lines in $\mathbb{R}_+^2$ are also lines in Y. As in the case of the Tikhonov regularization, the ℓ_1 minimization moves $\mathbf{y}(\widehat{\mathbf{x}}_{\text{NNLS}})$ perpendicularly to the lines in Y. Figure 1.8c shows this. As λ grows, the $\mathbf{y}(\widehat{\mathbf{x}}_{\text{NN}\ell_1})$ will eventually hit an edge of the cone, i.e., one dimension of $\widehat{\mathbf{x}}_{\text{NN}\ell_1}$ will become (and stay) equal to 0. This repeats with growing λ to more and more dimensions of $\widehat{\mathbf{x}}_{\text{NN}\ell_1}$. Hence, with growing λ, $\widehat{\mathbf{x}}_{\text{NN}\ell_1}$ becomes sparser.

The above are just a small set of all the existing regularization techniques. We chose to introduce these regularizations because they will be relevant in this thesis. Depending on what we know in advance about our solution and its characteristics, we will decide on the regularization that should be applied to our problem. Of course, as we did with, for example, the non-negative constraint and Tikhonov, different regularizations can be combined. The more we know about the solution, the more properties we can enforce on it. The less we know, the more general the estimator we should be applying.

1.4 Outliers

So far we have talked about how to combat the sensitivity of the estimate to the additive noise in order to get a good estimate of the source term. However, instability in the solution is not the only challenge that we may encounter.

The nature of the errors that appear in the problem may pose a significant challenge as well. This is also quite an old problem, and it was already mentioned in the first publications about LS, more than two centuries ago. Legendre wrote in 1805 [58]

If among these errors are some which appear too large to be admissible, then those observations which produced these errors will be rejected, as coming from too faulty experiments, and the unknowns will be determined by means of the other observations, which will then give much smaller errors.

Today, we refer to the *observations which produce errors too large to be admissible* as outliers. They can cause large deviations from the ground truth in the LS estimate. One example of outlier is the red point in Figure 1.9.

As shown in Eq. (1.4), the LS estimate is the vector $\mathbf{x}$ that minimizes the ℓ_2 norm between the observations $\widehat{\mathbf{y}}$ and the column space of $\mathbf{A}$. We can rewrite this estimate as

$$\widehat{\mathbf{x}}_{\text{LS}} = \underset{\mathbf{x}}{\arg\min} \, \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2 = \sum_{i=1}^{m} (\widehat{\mathbf{y}}_i - \mathbf{a}_i \mathbf{x})^2, \tag{1.19}$$

where $\mathbf{a}_i$ is the i-th row of $\mathbf{A}$. The terms $(\widehat{\mathbf{y}}_i - \mathbf{a}_i \mathbf{x})$ are typically known as *residuals*. If an observation i is an outlier, it will produce a residual much larger than the rest of the observations, and the square of its residual will dominate the sum in (1.19). Thus, to minimize the sum, LS puts a disproportionately large effort into reducing this one large residual, and it tends to ignore the rest. Figure 1.9 shows an example of this effect in a one-dimensional problem.

In the Chapter 3 we will see that our dataset indeed contains outliers. Thus, we need methods that are robust against these outliers, and are not overly affected by them.

In this thesis we will develop new methods that can simultaneously deal with outliers and make the problem well-posed. We organize these methods in two groups:
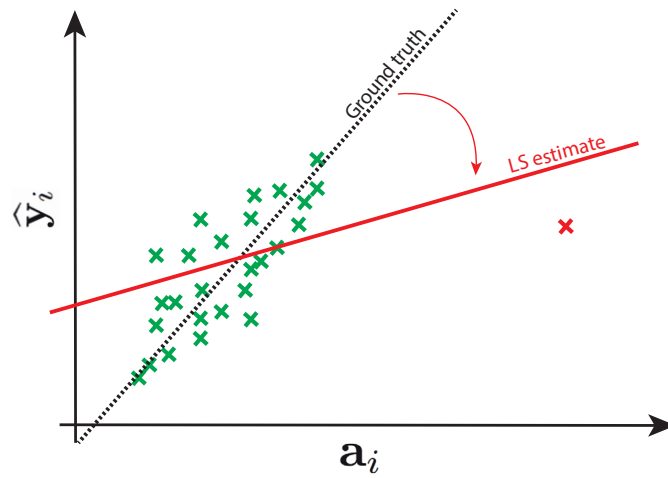
Figure 1.9: Effect of an outlier (red cross) in the LS estimate of a one dimensional problem.

1. Methods that identify the outliers, and then remove them or downgrade their influence, and then apply the traditional tools we described above to make the problem well-posed. To downgrade the influence of the outliers, a diagonal weighting matrix $\mathbf{W}$ can be introduced in the LS estimator such that

$$\widehat{\mathbf{x}}_{\mathbf{W}} = \arg\min_{\mathbf{x}} \|\mathbf{W}(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x})\|_2^2. \tag{1.20}$$

We will take a closer look at these methods and how to define the elements in $\mathbf{W}$ in Chapter 3.

2. Methods that modify the classic LS estimator by replacing the ℓ_2 loss function with a different one, thus making it robust to outliers. This is then combined with one or more regularizations that make the problem well-posed. These methods are described in Chapter 4.

1.5   Challenges

Besides the difficulties we have seen so far, related to inverse problems in general, we also have to deal with the domain-specific challenges related to estimating releases into the atmosphere:

1. **New look at the problem, from a signal processing perspective.** Meteorologists and air researchers have been working on this problem extensively in the past years. While efforts have been made to treat this problem from a signal processing perspective [57], the formulation was too simple to be applied to a real case with real data. A more universal mathematical model is needed, that allows a systematic representation of all the physically observable phenomena and all the possible estimators.

2. **Interdisciplinarity.** Bringing this problem into the realm of signal processing entails translating much knowledge and data from a completely different field of research, with its

own vocabulary and methods. This of course includes real measurement data and atmospheric transport models. This translation requires a deep understanding of atmospheric processes. We were fortunate to cooperate with one of the leading scientists in the air research field, Dr. Andreas Stohl. His help was invaluable in establishing this connection between the two domains.

3. **Lack of data.** The lack of data is still a critical difficulty in addressing inverse atmospheric dispersion problems. In addition to the scientific community, also political and military authorities are interested in estimating emissions of radioactive material into the atmosphere, because they are related not only to accidents in nuclear power plants, but also to nuclear weapons. Hence, most of the relevant measurement data is treated as secret information, and is not available freely to the scientific community.

   In the case of the Fukushima nuclear accident, good quality measurements are scarce for several reasons. First, the release of nuclear material was unexpected, and thus many measurement sets lack data in the first hours following the incident. Adding to this are the various political interests which make access to reliable data difficult. Finally, the destruction that took place around the nuclear power plant also contributed to the lack of good data.

4. **Lack of controlled experiments.** To develop new algorithms to estimate the source, we simplify the real world and make assumptions about it. In other words, we model reality with approximations. But as George Box put it in [16],

   Now it would be very remarkable if any system existing in the real world could be exactly represented by any simple model. However, cunningly chosen parsimonious models often do provide remarkably useful approximations. For example, the law $PV = RT$ relating pressure P, volume V and temperature T of an "ideal" gas via a constant R is not exactly true for any real gas, but it frequently provides a useful approximation and furthermore its structure is informative since it springs from a physical view of the behaviour of gas molecules.

   For such a model there is no need to ask the question "Is the model true?". If "truth" is to be the "whole truth" the answer must be "No". The only question of interest is "Is the model illuminating and useful?"

   To know which models are useful in our context of atmospheric dispersion, we need to evaluate objectively how good they are at estimating the source term. To do so, we test the models using very special datasets, where we know the ground truth of the source we are trying to estimate. The only way to create such special datasets is to carry out controlled experiments.

   Unfortunately there is a severe lack of controlled experiments of atmospheric dispersion of materials. This is due to, first, their high cost, and second, the fact that such experiments involve the released of harmful substances into the environment.

   Controlled experiments would help to discard quickly models that are not useful. The lack of them make the process of selecting the best source estimators quite complex.

1.6 Contributions and Outline

In this thesis we explore extensively all the problems and challenges that we described above, proposing several new solutions.

Chapter 2 focuses on the estimation of the emissions over time of Xenon (Xe-133) during the Fukushima nuclear accident. Here the lack of observations and of ground truth are the main challenges. We aim to find a regularization that uses a suitable model for the signal. The Xe-133 was released during quick bursts, and the releases are positive, so we use a non negative sparse regularization to reconstruct the source. Total releases estimated with our method match the total Xe-133 inventory existing in the power plant before the accident. Also, the times of the estimated explosions correspond to the events that took place during the accident.

In Chapter 3 we work with one of the few available datasets coming from a controlled experiment: the European Tracer Experiment (ETEX). Among other things, this dataset allows us to explore the nature of the errors. We show that they come from a heavy-tailed distribution, which implies the presence of outliers in the measurement dataset.

Thus, we propose new methods that are robust against these outliers. Our approach is to detect and remove outliers blindly, i.e. without any knowledge of the ground truth, and then to apply classic methods to the *clean* dataset. We propose two methods to remove outliers. The first algorithm is based on the comparison of different model matrices which are generated using different weather forecast models. The second algorithm is based on random sample techniques. The efficiency of these algorithms is demonstrated on the ETEX dataset.

Chapter 4 introduces a different point of view to deal with outliers: we propose a new estimator that is *robust* against outliers, is also efficient when no outliers are present, and at the same time it stabilizes the problem by making it well-posed. We call this estimator the regularized τ estimator. We provide algorithms to compute it, and provide a complete analysis of its properties. Finally we test its performance on the ETEX dataset.

In Chapter 5 we review the software that generated the results obtained in this thesis. This software implementation has two goals: first, to make our results reproducible and second, to make the new algorithms available and ready to use for other researchers and engineers.

In Chapter 6 we review my six-month internship in industry, my work there, and the lessons learnt during that time.

Chapter 2

The Fukushima Inverse Problem

Knowing what amount of radioactive material was released from the damaged Fukushima nuclear power plant in March 2011 is crucial to understanding the scope of the consequences. Moreover, this knowledge could be used in forward simulations to obtain accurate maps of deposition. But these data are often not publicly available, or are of questionable quality. In this chapter [1] we propose to estimate the emissions over time by solving an inverse problem. Previous approaches rely on a detailed expert guess of how the releases appeared, and they produce a solution strongly biased by this guess. We propose a method based on the non-negative sparse regularization described in Chapter 1, which does not need any a priori guess and is thus unbiased. Together with the atmospheric dispersion models and worldwide radioactivity measurements our method correctly reconstructs the times of major events during the accident, and gives plausible estimates of the released quantities of Xenon (Xe-133).

2.1   Goal

Nuclear power plants (NPP) provide an abundant, relatively cheap, and carbon-neutral source of energy. However, they also introduce a possibility, albeit a very remote one, of a major accident. A nuclear accident is defined by the International Atomic Energy Agency as an event having lethal consequences, environmental effects such as large radioactivity releases, and producing long-lasting facility defects such as core melts. Level 7 on the International Nuclear Events Scale is defined as *a major release of radioactive material with widespread health and environmental effects requiring implementation of planned and extended countermeasures* [42]. Two accidents have reached this level—Chernobyl in April 1986 and Fukushima Daiichi in March 2011.

The principal consequence of NPP level 7 accidents is the release of radioactive material. Transported through the atmosphere, it eventually gets widely spread, polluting the environment for centuries at a large scale. Exposure to the radioactive material causes cancer, teratogenesis, cognitive decline, and heart disease [59]. Thus, it is imperative to monitor the radioactive contamination of soil, water, and atmosphere. Unfortunately, the contamination can only be accurately measured at a limited number of survey sites due to the cost of scientific grade equipment. This suggests the need for numerical simulations of atmospheric dispersion [55, 65].

---

1. This chapter is the result of a collaboration with I. Dokmanic, J. Ranieri, R. Scheibler, M. Vetterli and A. Stohl [46].

Getting accurate concentration and deposition values through simulations requires the knowledge of the source term — how much radioactive material was released at what times. Its precise estimate is essential to the proper estimation of the contamination and implementation of risk-reducing measures. However, these data are often not publicly available or are simply unknown. In particular, some works [25, 65] challenge the data released by the Japanese government about the Fukushima accident. An alternative is to calculate the source term based on spatio-temporal samples of the concentration, that is, by solving an inverse problem.

We propose to estimate the source term by inverting the atmospheric dispersion. This would provide us with an estimate of how much radioactive material was released, and provide a solid starting point for understanding the scope of the pollution through dispersion simulation. This problem is a particular case of the atmospheric dispersion problems described in Chapter 1, so it can be modelled as a linear inverse problem

$$\widehat{\mathbf{y}} = \widehat{\mathbf{A}}\mathbf{x} + \mathbf{e}. \tag{2.1}$$

Therefore, we set out to solve the following problem:

Problem 2.1 (*Fukushima Inverse Problem*)

Given the measurements $\mathbf{y}$ collected at a number of different survey sites, and the estimated model for the atmospheric dispersion $\widehat{\mathbf{A}}$, get an estimate $\widehat{\mathbf{x}}$ of $\mathbf{x}$, the temporal variation of the release of Xenon (Xe-133).

2.2   The Fukushima dataset

During the Fukushima nuclear accident, unknown quantities of several radioactive materials, that were stored in the NPP, were released into the atmosphere. Among them was the Xenon isotope Xe-133, which was released in the form of a gas.

On the right hand side in (2.1), the source term $\mathbf{x}$ contains the rate of release of Xe-133 in Becquerels per second (Bq/s) [2] between March 10 and March 16 at the Fukushima Daiichi NPP. The temporal resolution of the source term is three hours. The height of the emissions is also considered, and it is discretized in three levels, 0-50 m, 50-300 m, and 300-1000 m. The left part of Figure 2.1 shows this discretization. The height of the emissions needs to be taken into account since the atmospheric transport of particles depends substantially on the altitude of the source [65]. The right part of Figure 2.1 shows how the unknown vector $\mathbf{x}$ is organized. Taking all into account, the vector $\mathbf{x}$ has 120 components (3 heights x 40 three-hours intervals).

The vector $\widehat{\mathbf{y}}$ in equation (2.1) contains the concentration measurements of Xe-133 in the air at stations of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) monitoring network. This network comprises 25 stations equipped with very sensitive radioactive Xenon detectors. The dataset we will use here contains measurements from 15 CTBTO and 2 non-CTBTO stations. The locations of the stations are shown in Figure 2.2. In total 858 measurements are used.

The detectors typically have three stages. The first two stages concentrate and purify the gas sample respectively, and the third one measures the activity of the final gas sample. The duration of the sampling and purification process limits the number of samples per day to a

---

2. One Becquerel is defined as the activity of a quantity of radioactive material in which one nucleus decays per second [9].
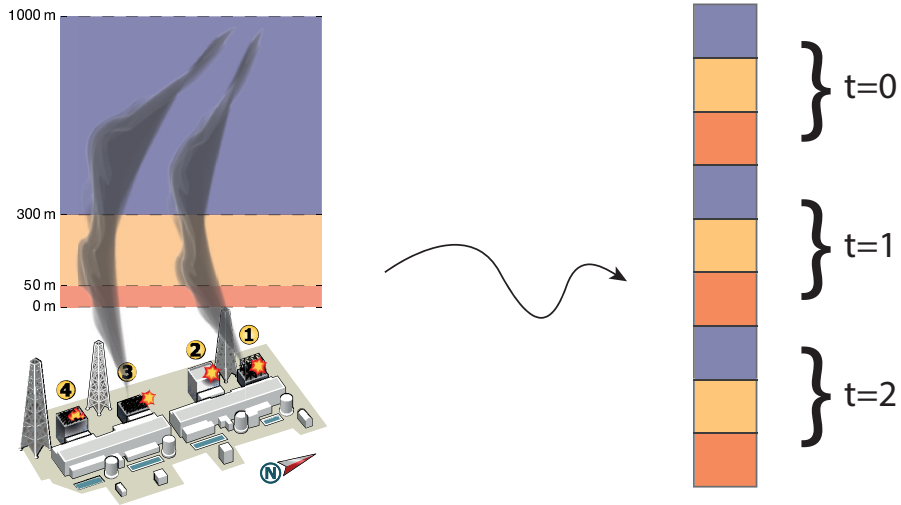
Figure 2.1: The figure shows how the unknown vector $\mathbf{x}$ is organized in the Fukushima dataset. Different times and three different discretized altitudes are considered. Thus, at each particular moment, i.e. $t = 0$, the emissions at three different altitudes is to be estimated.

maximum of three, depending on the detector model [73]. These systems measure Xe-133 up to an accuracy of 0.1 mBq m^{-3} [65]. Interestingly, measurements from the CTBTO station located in Japan could not be used because the levels of Xe-133 were over the highest detectable level of the system, saturating the detectors. All the measurements were corrected for radioactive decay. For additional information on the pre-processing of the data see [65].
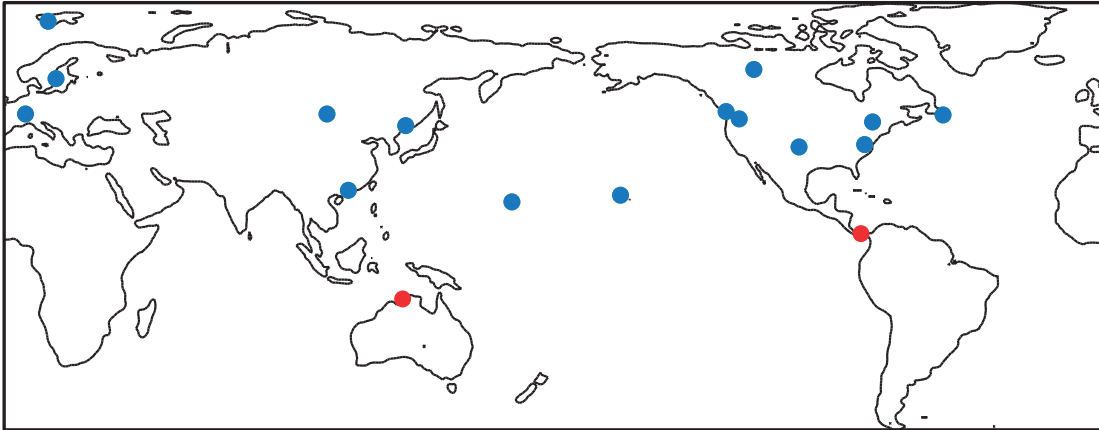

Figure 2.2: Measurement stations plotted on the world map. The red dots correspond to measurements stations that are completely left out of consideration after cleaning the matrix.

Finally, the model matrix $\widehat{\mathbf{A}}$ describes every measurement as a linear combination of source terms. In this case, the $\widehat{\cdot}$ symbol on the matrix indicates that the model is an estimation and

it may contain errors, $\widehat{\mathbf{A}} = \mathbf{A} + \mathbf{N}$. The coefficients of $\widehat{\mathbf{A}}$ are computed using FLEXPART (see Chapter 1 for more details). As the number of measurements is relatively low, a coarse temporal discretization is necessary. As we said, the length of the vector $\widehat{\mathbf{y}}$ is 858 and the length of $\mathbf{x}$ is 120, thus the model matrix has dimensions 858×120.

2.3   Prior Approaches to the Fukushima Problem

The model matrix $\widehat{\mathbf{A}}$ is a tall matrix, hence, the linear system (2.1) has more measurements than unknowns. Since it is an overdetermined system, the first approach is to use least squares (1.4) to estimate the emissions $\mathbf{x}$, but this is not a good idea, because the condition number of the model matrix is 3.4×10^5, which is very large. As we explain in Section 1.3, we might attempt to fix the conditioning issues using Tikhonov regularization, but the result would still be unsatisfactory. Regularizations introduce a priori information about the solution. In particular, Tikhonov regularization introduces the assumption that the solution has a small Euclidean norm, but, in our case, assuming that the source term has a small Euclidean norm has no justification. In fact, the contrary may be true: Xe-133 is a gas, and it was released mostly during explosions and ventings [65], which would cause a sparse source term, with a few strong peaks, and no releases between these peaks.

The dispersion-modelling community developed specific methods to cope with the lack of measurements and the large condition number of the model matrix. Typically, they aid the estimation with an a priori guess of the result based on expert knowledge. In the case of the accident in Fukushima, to build this a priori guess involves two steps. First, one estimates the existing quantity of Xe-133 before the accident per reactor and per spent-fuel pond. This is done using inventories and an estimation of how much fuel was used since the latest inventory until the date of the accident. Second, it is necessary to estimate the temporal sequence of the emissions. This is done using existing information about observed radiation increases, pressure build-up, and decay in different reactor compartments, information on relief valves open or closed, and the timing of hydrogen explosions[65].

Besides the a priori guess, more assumptions about the solution are included into the problem, like uncertainties of the a priori guess, smoothness of $\mathbf{x}$ or uncertainties of the measurements. All these assumptions are variations of the regularizations and outlier detection techniques explained in Chapter 1:

1. **A priori guess.** Denoting the a priori guess as $\mathbf{x}_a$, and defining the regularization parameter $\lambda \geq 0$, $\mathbf{x}_a$ can be included into the LS estimator modifying the Tikhonov regularization

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \widehat{\mathbf{A}}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x} - \mathbf{x}_a\|_2^2. \tag{2.2}$$

   The second term enforces the solution $\widehat{\mathbf{x}}$ to be close to $\mathbf{x}_a$. The influence of this term in the problem is controlled with λ.

2. **Uncertainties in the a priori guess.** If besides of having an a priori guess, we also know the uncertainty of each estimated element in $\mathbf{x}_a$, we can include this information into the problem. If each element in $\mathbf{x}_a$ is modelled as a random variable, its uncertainty is given by its variance σ_x^2. The larger variance an element has, the more uncertain it is. If we assume that the elements in $\mathbf{x}_a$ are independent, then its covariance matrix $\mathbf{C}_x$ is

a diagonal vector that contains the variances of each element in $\mathbf{x}_a$. $\mathbf{C}_x$ can be used to introduce the uncertainties into the problem: let us define a diagonal weight matrix $\mathbf{W}_a$ such that $\mathbf{W}_a^\top \mathbf{W}_a = \mathbf{C}_x^{-1}$. Then

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \widehat{\mathbf{A}}\mathbf{x}\|_2^2 + \|\mathbf{W}_a(\mathbf{x} - \mathbf{x}_a)\|_2^2. \tag{2.3}$$

We can see these weights as individual regularization parameters for each component in $\mathbf{x}_a$: if the variance of a component is large, its weight is small, and its influence on the solution of the problem is also small.

3. **Smoothness.** In Chapter 1 we explained how to enforce smoothness in the solution by using the discrete first derivative. We can enforce strongest smoothness by using the discrete second derivative $\mathbf{D}_2$. If $\epsilon \geq 0$ is the regularization parameter, then

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \widehat{\mathbf{A}}\mathbf{x}\|_2^2 + \epsilon\|\mathbf{D}_2\mathbf{x}\|_2^2. \tag{2.4}$$

The second derivative can be seen as the derivative of the first derivative. If we force it to be small, we are enforcing smoothness in the first derivative.

4. **Uncertainties in the observations.** As we did for the a priori guess, uncertainties for the observations can be also introduced into the problem. If we have access to the covariance matrix of the observation vector $\mathbf{C}_y$, the weights can be defined in a similar manner as we defined $\mathbf{W}_a$, $\mathbf{W}_y^\top \mathbf{W}_y = \mathbf{C}_y^{-1}$. They are introduced into the problem as

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{W}_y(\widehat{\mathbf{y}} - \widehat{\mathbf{A}}\mathbf{x})\|_2^2. \tag{2.5}$$

These weights are related with the ones introduced in Chapter 1 to mitigate the effect of outliers.

Putting together all these pieces leads to the convex program that the authors propose in [65],

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{W}_y(\widehat{\mathbf{y}} - \widehat{\mathbf{A}}\mathbf{x})\|_2^2 + \|\mathbf{W}_a(\mathbf{x} - \mathbf{x}_a)\|_2^2 + \varepsilon\|\mathbf{D}_2(\mathbf{x} - \mathbf{x}_a)\|_2^2. \tag{2.6}$$

The authors obtain useful results with this technique. However, this estimation strongly relies on the a priori expert-based solution. To exemplify the point, we carry out the following numerical experiment: we generate synthetic measurements $\widehat{\mathbf{y}}_a$ using

$$\widehat{\mathbf{y}}_a = \widehat{\mathbf{A}}\mathbf{x}_o + \mathbf{e}, \tag{2.7}$$

where $\mathbf{x}_o$ is the a priori guess used by the authors in [65]. It is displayed as a dashed green line in Figure 2.3(A). The vector $\mathbf{e} \in \mathbb{R}^{m \times 1}$ contains i.i.d. samples from a zero-mean Gaussian distribution. The variance of the distribution is adjusted such that the signal-to-noise ratio is 10dB. The source $\mathbf{x}_o$ is estimated using $\widehat{\mathbf{y}}_a$, $\widehat{\mathbf{A}}$, and the method described in (2.6), where we set $\mathbf{x}_a = \mathbf{x}_o$. This estimate is shown as a solid red line in figure 2.3(A). Now we generate a new a priori guess $\tilde{\mathbf{x}}_a$ (which is wrong) where we relocate the two main peaks of $\mathbf{x}_a$. This new guess is plotted as a dashed green line in 2.3(B). We use it to estimate $\mathbf{x}_o$ using the measurements $\widehat{\mathbf{y}}_a$ that we generated before using $\mathbf{x}_a$, and the same model matrix $\widehat{\mathbf{A}}$. The resulting estimate is shown in red in Figure 2.3(B).
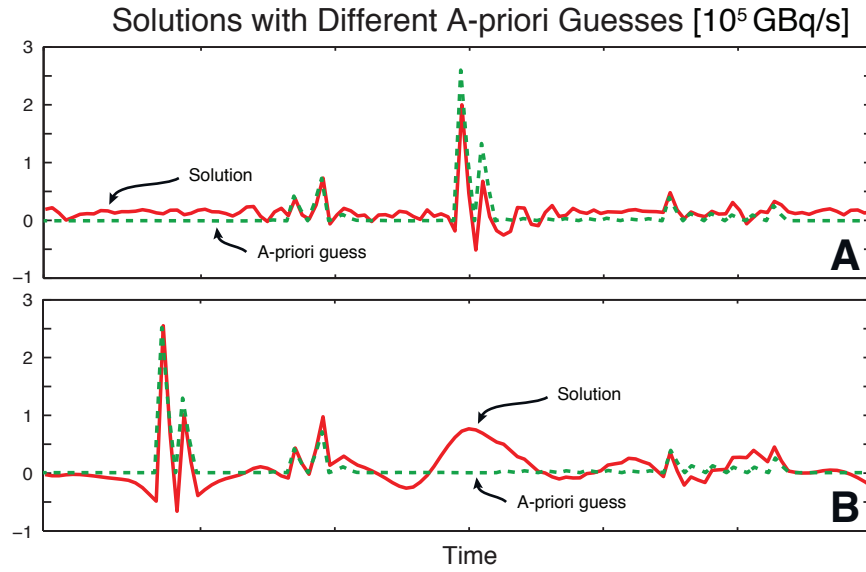
Figure 2.3: Sensitivity of the formulation in [65] to the a priori guess. (A) A priori guess used in the paper and the corresponding reconstruction. (B) Modified a priori guess with a relocated peak. The solution changes shape to match the new (in reality nonexistent) peak.

The estimate now perfectly follows the relocated peaks of $\tilde{\mathbf{x}}_a$, and the real peaks of $\mathbf{x}_o$ appear smoothed out. Our point here is that the method described in (2.6) is useful, but very sensitive to the a priori guess. If the a priori guess is wrong, our estimate will be biased strongly by this guess. In the case of this problem, as we explained above, to estimate this guess is a complex process, where many data and simulation models of nuclear power plants are involved. Because of this, it is safe to say that, most likely, this a priori guess contains errors.

What we seek is a method to detect correctly the emission times without any a priori expert knowledge. It would be wrong to just dismiss the available expert estimate, but a good method must give plausible results even without it.
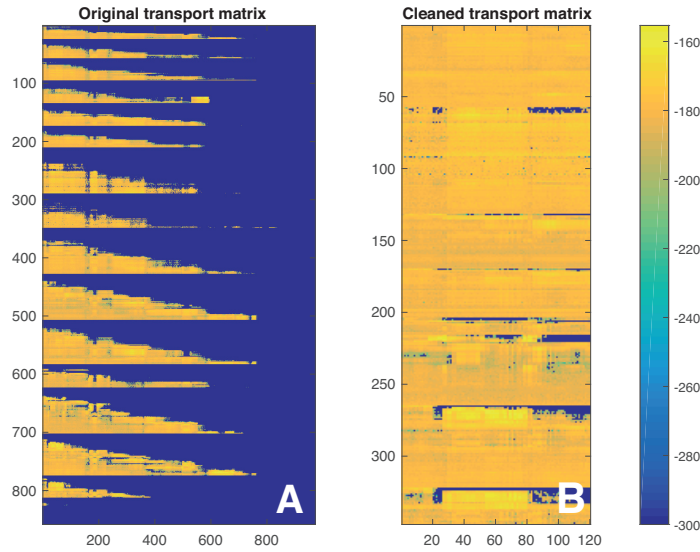
Figure 2.4: Matrix (A) before and (B) after cleaning. For a better visualization we use a logarithmic scale, so each entry in the values in the plots corresponds to $10 \log |a_{ij}|$, where a_{ij} is an entry in the matrix.

2.4 Our Approach

We will show that a successful estimation of the temporal variation of the releases comprises several ingredients:

Selecting unknowns

We want to estimate the temporal releases of the radioisotope Xe-133 during the nuclear accident. It is important to note that all the relevant events of the accident, i.e. explosions and controlled ventings, are limited to the first 5 days following the earthquake. Furthermore, Xe-133 radioisotope is a gas. If a facility where the gas is stored is damaged, the gas is released quickly. Thus, it is reasonable to assume that the Xe-133 was released during the events of the accident, i.e. over the 5 days following the earthquake, and not afterwards. Thus, we set the scope of the unknowns to this 5-day period. The temporal discretization of the dataset is 3 hours, and the altitude is discretized in 3 levels, as is described in Section 2.2. Thus, we select the first $5 \times 8 \times 3 = 120$ columns of the original matrix displayed in Figure 2.4(A), which correspond to the unknowns we want to estimate.

Cleaning the matrix

In Figure 2.4(A) you can observe the matrix used in [65]. We notice in this matrix many elements with a small magnitude, represented by the blue areas (note that the color bar is in logarithmic scale). That is, there are matrix rows that have zero or negligible norms. This means that the sensors associated to these rows barely sense releases from Fukushima at the interval of time we are interested in. These small-norm rows deteriorate the solution of the inverse problems.

This can be noted in the very unfavorable condition number of the model matrix $\widehat{\mathbf{A}}$. As we can observe in Figure 2.5, the condition number of the matrix decreases when the rows with small norms are removed. To remove the small rows, we set up the ratio $\nu(\mathbf{A})$

$$\nu(\mathbf{A}) = \frac{\min_i \sqrt{\sum_{j=1}^{n} |a_{ij}|^2}}{\max_i \sqrt{\sum_{j=1}^{n} |a_{ij}|^2}}. \tag{2.8}$$

All the rows i whose normalized norm is smaller than the defined ratio $\nu(\mathbf{A})$ such that

$$\frac{\sqrt{\sum_{j=1}^{n} |a_{ij}|^2}}{\max_i \sqrt{\sum_{j=1}^{n} |a_{ij}|^2}} < \nu(\mathbf{A}) \tag{2.9}$$

are removed.

After this process, we are left with the matrix shown in Figure 2.4(B). It is evident from the color map that the remaining rows have much narrower dynamics. The condition number of the original matrix is 3.4×10^5, and after the cleaning process it is reduced to 1303.
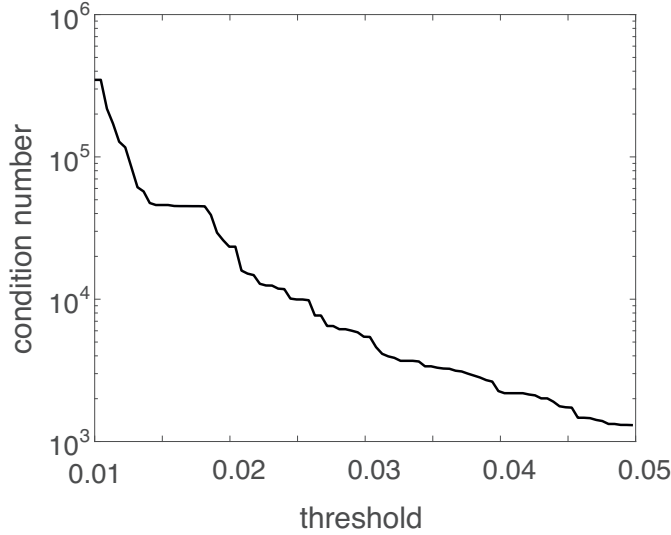

Figure 2.5: The condition number of the original matrix decreases when the rows with a small norm are removed. The threshold in this case is defined as $\min_i \sqrt{\sum_{j=1}^{n} |a_{ij}|^2}/\max_i \sqrt{\sum_{j=1}^{n} |a_{ij}|^2}$, where a_{ij} is the element of $\widehat{\mathbf{A}}$ placed in the ith row and the jth column, and n is the number of columns in $\widehat{\mathbf{A}}$.

It is interesting to notice that cleaning the matrix removes the set of measurements taken at the two southernmost stations, marked in red in Figure 2.2. The CTBTO website features a fascinating video explaining that the equator is acting as a dividing line between the northern and the southern hemisphere air masses [28]. This means that the measurement stations located on the southern hemisphere could not sense the Xe-133 cloud since it could not pass the equator. Eventually the material got transported to the southern hemisphere but in very small amounts.

Scaling

The magnitude of the entries in the original model matrix is very small. The largest value is 2.91×10^{-16}. If we observe the location of the sensors in Figure 2.2, we notice that all of them are placed far away from Fukushima. Besides, simulations showed that, once it was released, the Xe-133 cloud was taken towards the Pacific Ocean [65], where there were no sensors. When the cloud of Xenon passed over some of the sensors, it was already attenuated. We recall that, as explained in Section 1.1, the entries in the matrix $\widehat{\mathbf{A}}$ represent the probability of a particle released at Fukushima at time t_j arriving at the location of the corresponding sensor at time t_i. Because all these circumstances, these probabilities are low.

The small magnitude of the matrix entries may cause numerical problems in the source estimation algorithms, above all in CVX, the Matlab package for specifying and solving convex programs that we use [35]. Thus, to avoid these problems, we scale both the matrix $\widehat{\mathbf{A}}$ and the set of measurements $\widehat{\mathbf{y}}$. The scale factor is 10^{16}. After the scaling, we observe that the numerical problems disappear.

Proper regularization

Xe-133 is a gas and, as we already said, it is reasonable to assume that it was ejected in short bursts. This fact can be also be appreciated in the a priori estimate showed in Figure 2.3(A). This differs from other radioactive materials like Cesium, which are released during more extended periods. The temporal variation of the Xe-133 emission exhibits several peaks and many small elements. Therefore, a proper regularizations should favor signals with many zeros and a few large elements. This is in contrast with what is known about the regularization based on ℓ_2 norms, such as Tikhonov. These favor many small/moderate elements.

We should find the solution yielding something close to the observed measurements, but with as many zero elements as possible. Since minimizing the number of non-zero elements is not tractable, we use the standard relaxation based on the ℓ_1 norm. This is the sparse regularization that is explained in Section 1.3.3,

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{W}_c(\widehat{\mathbf{y}} - \widehat{\mathbf{A}}\mathbf{x})\|_2^2 + \lambda\|\mathbf{x}\|_1. \tag{2.10}$$

The cleaning and scaling of the model matrix is represented here by the diagonal matrix $\mathbf{W}_c = c\mathbf{W}$, were c is the scaling parameter. The diagonal matrix $\mathbf{W}$ selects the rows we keep after the cleaning process. The elements of its diagonal corresponding with the rows we keep after the cleaning process are equal to one. The rest of its entries are zero.

Another assumption that we should include in the problem is the non-negativity of the source: the emissions of Xe-133 cannot be negative. Thus, we enforce the non-negativity constraint in a new convex program

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{W}_c(\widehat{\mathbf{y}} - \widehat{\mathbf{A}}\mathbf{x})\|_2^2 + \lambda\|\mathbf{x}\|_1. \tag{2.11}$$
$$\text{s.t. } \mathbf{x} \succeq 0$$

This is the non-negative sparse regularization that is introduced in Section 1.3.3.

How to choose the optimal λ parameter blindly is still an active research area. In this problem, the introduction of the non-negative constraint seems to stabilize the solution. We observe that the solution is stable across a very wide range of λ values. λ is selected manually within this range.

Including as little a priori information as possible

As we explained in Chapter 1, to include a priori information about the solution into the problem may help the estimator to find a better estimate. However, if the introduced a priori information is not correct, it will bias the estimate, as Figure 2.3 shows. We need to let the measurements *talk* as much as possible, and say what they have to say. If we constrain the problem too much, we will force the measurements to tell us what we want to hear. To avoid this situation we include as little a priori information into the problem as possible, enough to get a reasonable solution. This is why, in Eq. (2.11), we do not introduce an a priori source neither any kind of uncertainty. As it is stated in [65],

The above formulation implies normally distributed, uncorrelated errors, a condition that we know to be not fulfilled. Observation errors (also model errors are subsumed in this term) may be correlated with neighbouring values, and deviations from the prior sources are likely to be asymmetric, with overestimation being more likely than underestimation as zero is a natural bound. The justification for using this approach is the usual one: the problem becomes much easier to solve, detailed error statistics are unknown anyway, and experience shows that reasonable results can be obtained.

Given that the uncertainties themselves are quite unprecise, we do not use them.

2.5   Sensitivity Test

A way to assess the robustness of different methods is to measure how sensitive they are with respect to the magnitude of errors in the measurements, $\mathbf{e}$. To this end we simulated the reconstruction by the Tikhonov-regularized program, regularization with the second derivative smoothing, and the non-negative sparse reconstruction while adding different amounts of Gaussian additive noise to the measurements. Synthetic measurements were generated using the model matrix of the Fukushima dataset and the a priori Xe-133 guess used in [65], as it is described by (2.7). Results are shown in Figure 2.6. The ℓ_1-regularized approach performs the best among the tested methods, showing that the model that it assumes is the most appropriate among all the tested ones.
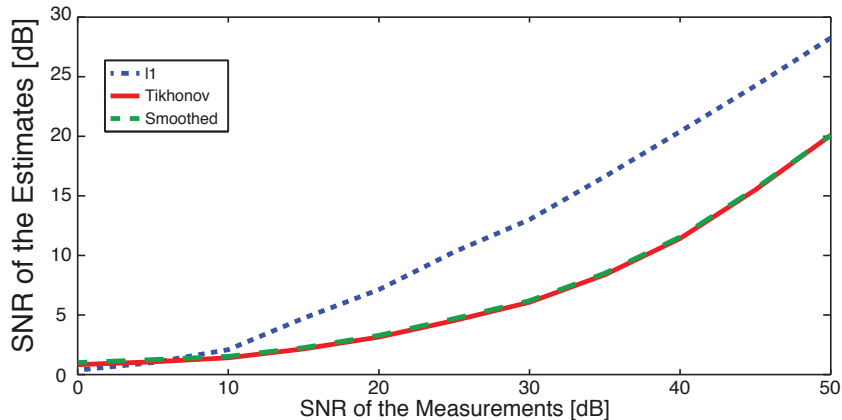
Figure 2.6: Signal-to-noise ratio (SNR) of the recovered source term as a function of the SNR of the synthetic measurement obtained by the forward-model matrix and an a priori source term from [65], with added Gaussian noise.

2.6 Experiments with real data

We apply the proposed solution (2.11) with the matrix pre-processing to real data. The measurements and the matrix that we use are the ones described in Section 2.3. This experiment is blind, not involving any prior information about the source except that it is sparse and positive.

To assess the quality of the estimate is a challenge in this case: the ground-truth is unknown. Thus, we need to use indirect methods using the available information to check that our estimate is at least reasonable. We use two approaches:

1. *Total release.* It is known that all the Xe-133 stored in the NPP before the accident was released. This quantity can be computed using initial inventories and estimating how much fuel was spent until the day of the accident. It is estimated that the total quantity of Xe-133 at the moment of the accident was 12.4×10^{18} Bq [65]. The total release estimated by our method is 9.62×10^{18} Bq. This number is of the same order, and quite close, to the estimated total inventory.

2. *Timing.* Releases of Xe-133 are likely happening during explosions and major ventings. The estimated emissions are illustrated with respect to the timeline of the Fukushima accident for the three height ranges in Figure 2.8. We see that the timing of most of the estimated emissions can be related with ventings and explosions, and in fact matches to a good extent the a priori solution mentioned earlier.

This successful estimation is quite surprising: if you look at the tracer experiment results in [64] and showed in Figure 2.7, one can see that even though there is correlation between the measured and numerically predicted values, the variability is considerable – often orders of magnitude. We would expect all conventional dispersion models to have troubles estimating anything precise.
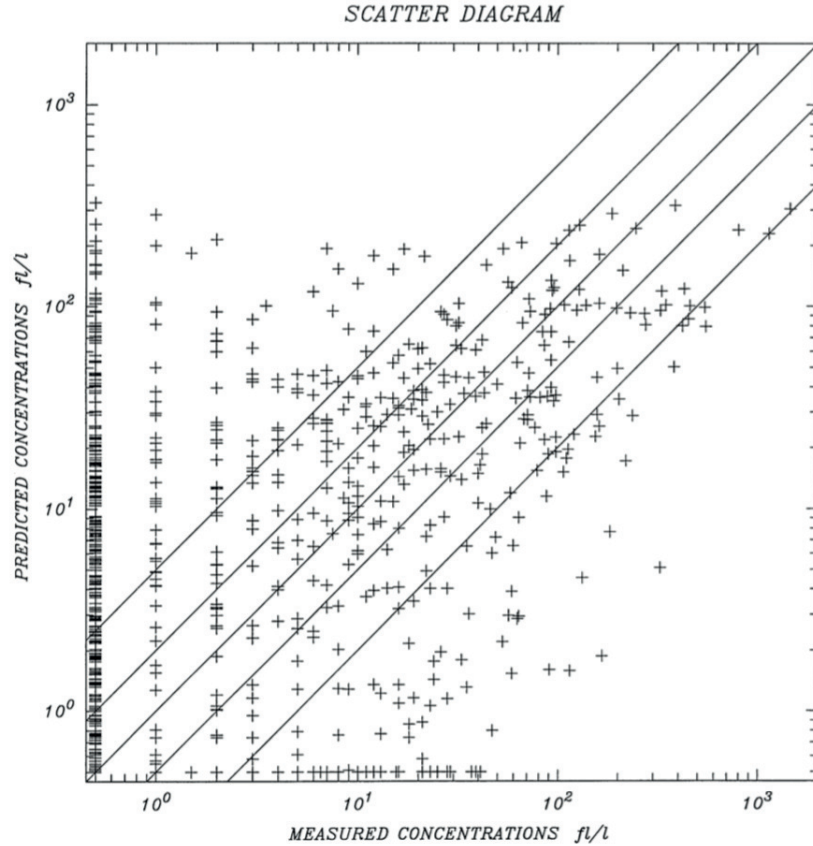
Figure 2.7: Scatter plot from [64] between observed and predicted tracer concentrations computed by FLEXPART for all CAPTEX releases. All concentrations less than $0.5 fl/l$ are plotted at that value. In this case, $1 fl/l = 1.56 \cdot 10^{-11} g/m^3$ [33]

Besides, the estimation is based on measurements taken on different continents, thousands of miles away from Fukushima.

Of course these two tests do not assure that this estimate is correct; they are just sanity checks that suggest that the estimate is reasonable, and it could be close to reality. To have more confidence in our estimates, the source estimation algorithms should be intensively tested in advance in controlled experiments where we know precisely the ground truth and the mean squared error of the estimate can be measured. Also these experiments can provide more insights about the problem, like for example about the nature of the errors in the model. This knowledge could also be included in the algorithms to improve the quality of the estimates. Unfortunately just a few experiments of this kind have being carried out. The most important of these is the European Tracer Experiment (ETEX). We will use it in the rest of this thesis to test our algorithms, and to have more insights about various aspects of this problem. In particular, we will dig into the nature of the errors **e**, and we will develop new estimators that are adapted for errors of this nature.
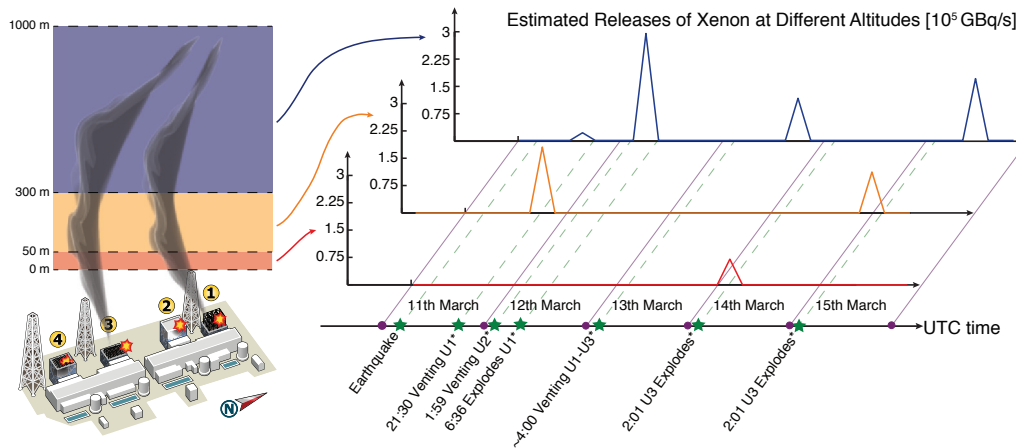
Figure 2.8: Reconstruction of the emission rates using the proposed algorithm. Reconstructed rates are shown for the three considered heights. Note that the largest peaks are in correspondence with venting events, considered to be the main cause of Xe-133 releases. The timeline indicates events at the Fukushima NPP in the temporal window of interest and is built from numerous sources, listed on the Wikipedia page for the Fukushima I Nuclear accident.

Chapter 3

Wrong memories

In the previous chapter we estimated the accidental releases of Xenon (Xe-133) from the nuclear power plant in Fukushima in March 2011. One of the challenges we faced was stabilizing the solution. To do so, we introduced a priori information about the solution using regularizations. In this chapter [1], we shift our focus from the nature of the solution to the nature of the errors contained in our measurements. What are these errors like? How can we model them? How do they affect our estimators? Can we design estimators that are particularly resistant to them? How? We will answer these questions, and in particular show that the errors contain outliers which strongly degrade the performance of traditional estimators. To combat the effect of outliers, we propose several modifications that make traditional estimators less sensitive to these outliers, thus resulting in a considerable improvement in performance.

3.1   The errors in the system

Previously proposed estimators of atmospheric emissions assume that errors are additive and have a Gaussian distribution. Hence, these estimators combine LS or weighted least squares (WLS) with different regularization functions. For instance, they may use an LS approach with Tikhonov (ℓ_2-norm) regularization that was explained in Chapter 1, or variants of this method, to recover an estimate $\widehat{\mathbf{x}}$ of the source

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\arg\min} \, \|\mathbf{A}\mathbf{x} - \widehat{\mathbf{y}}\|_2^2 + \lambda\|\mathbf{x}\|_2^2$$

where $\lambda \geq 0$ is the regularization parameter. One example is given in [61], where LS is combined with the Tikhonov regularization and a smooth first derivative constraint:

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\arg\min} \, \|\mathbf{A}\mathbf{x} - \widehat{\mathbf{y}}\|_2^2 + \lambda\|\mathbf{x}\|_2^2 + \beta\|\mathbf{D}\mathbf{x}\|_2^2, \tag{3.1}$$

where $\mathbf{D}$ is the difference operator, and $\beta \geq 0$ and $\lambda \geq 0$ are regularization parameters. Or, as we already saw in Chapter 2, WLS can be combined with an a priori guess $\mathbf{x}_a$, smoothness constraint and uncertainties as

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\arg\min} \, \|\mathbf{W}_y(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x})\|_2^2 + \lambda\|\mathbf{W}_a(\mathbf{x} - \mathbf{x}_a)\|_2^2 + \varepsilon\|\mathbf{D}_2(\mathbf{x} - \mathbf{x}_a)\|_2^2,$$

---

1. This chapter is the result of a collaboration with A. Stohl, B. Bejar Haro and M. Vetterli [49], [45].

where $\mathbf{D}_2$ is the discrete second derivative, $\mathbf{W}_y^\top \mathbf{W}_y = \mathbf{C}_y^{-1}$ and $\mathbf{W}_a^\top \mathbf{W}_a = \mathbf{C}_x^{-1}$, where $\mathbf{C}_y$ and $\mathbf{C}_x$ are the covariance matrices of the observations and of the initial guess, respectively. All these approaches minimize the energy of the disagreement between the model and the observations, while at the same time keeping the energy of the solution in check. While this is a reasonable approach, no metrics of real performance are (or can be) given in most of these studies, simply because in a vast majority of datasets, no knowledge of the ground truth is available, as we saw already in Chapter 2. This fact made it impossible to evaluate the true performance of any of these approaches.

In this chapter we consider a very special dataset for which the ground truth is known. Can we use this controlled dataset to learn more about the nature of the problem? What happens if, in fact, the errors in the problem are not Gaussian? What happens if we have outliers in our dataset? How do they influence estimators that are built on the assumption of Gaussian errors? How can we improve the estimates in these cases?

We will show that, in fact, the errors present in this source estimation problem come from a heavy-tailed distribution, which implies the presence of outliers in the measurement dataset. This makes the widely-used source estimation algorithms like (1.13) (which assume Gaussian additive errors [58]) highly sensitive to outliers. In fact, we will show that as expected, if the outliers are removed, the source estimation using (1.13) improves substantially.

Then, we follow two approaches to solve the problem. In both of them we use the ℓ_2-norm as the loss function. In the first approach, we propose to use the WLS approach, with the weights proportional to the inverse of the variance of the errors. These weights are found using a combination of several model matrices which are generated using different weather forecast models. In the second approach, we propose to discard the measurements that come from a distribution with a different variance (i.e. outliers) and use only the ones that come from the same distribution (i.e. inliers). For this selection we use random sampling techniques. The efficiency of both approaches is demonstrated on a real-world dataset, and their performance is evaluated and compared to other existing methods.

3.2 The ETEX dataset

The European Tracer EXperiment (ETEX) was established [51]

...to evaluate the validity of long-range transport models for real-time application in emergency management and to assemble a database which will allow the evaluation of long-range atmospheric dispersion models in general.

The objectives of ETEX were

1. *to conduct a long-range atmospheric tracer experiment with controlled releases under well-defined conditions,*

2. *to test the capabilities of organisations in Europe responsible for producing rapid forecasts of atmospheric dispersion to produce such forecasts in real time,*

3. *and to evaluate the validity of their forecasts by comparison with the experimental data.*

To this end, the ETEX experiment comprised two separate releases of perfluorocarbon tracers (PFT). PFTs are suitable tracer substances for experiments over long distances: they are

non-toxic, non-depositing, non-water-soluble, chemically inert, and environmentally safe. In particular, perfluoromethylcyclohexane (PMCH) was used in the first release, and perfluoromethyl-cyclopentane (PMCP) was used for the second one. Two different substances were used to avoid cross-contamination in the experiments.

In this thesis we will focus on the dataset obtained with the first release, for the following reasons [51],

Compared to the first release, [in the second release] much less tracer was found. Of the total number of data points obtained only 13% revealed elevated PMCP concentrations ... Not only were fewer non-zero values calculated, but also the concentrations were much lower.

Thus during the second experiment the sensors barely detected the PMCP plume, so to recover the source in this case is an extremely difficult task.
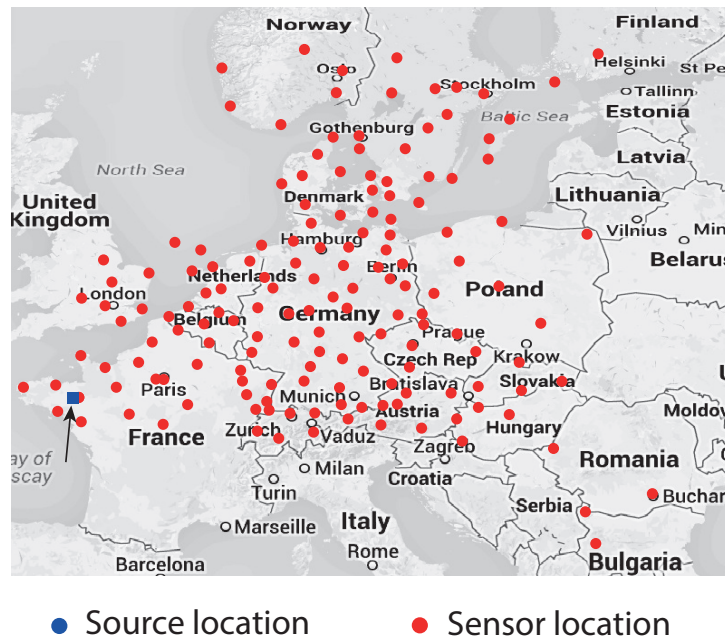

Figure 3.1: The map shows the location of the source and the locations of the measurement stations that were used during the ETEX experiment.

The first tracer release started on 23 October 1994, 16:00 UTC, and ended on 24 October 1994, 3:50 UTC. Figure 3.1 shows with a blue square the location of the release in Monterfil, France. During these 12 h a total of 340 kg of PMCH were emitted, resulting in an average release rate of 7.95 g/s. In every inverse problem, a time window must be defined, during which the activity of the source is to be recovered. In this particular case, we define a window of five days (although we in fact know that the ETEX emissions took place over only 12 hours) or $5 \cdot 24 = 120$ hours. Since the time resolution is one hour, we have 120 unknowns in the system. We decided to use the five-days window for the unknowns because typically, in a non-controlled dataset, we do not know precisely when the release took place.
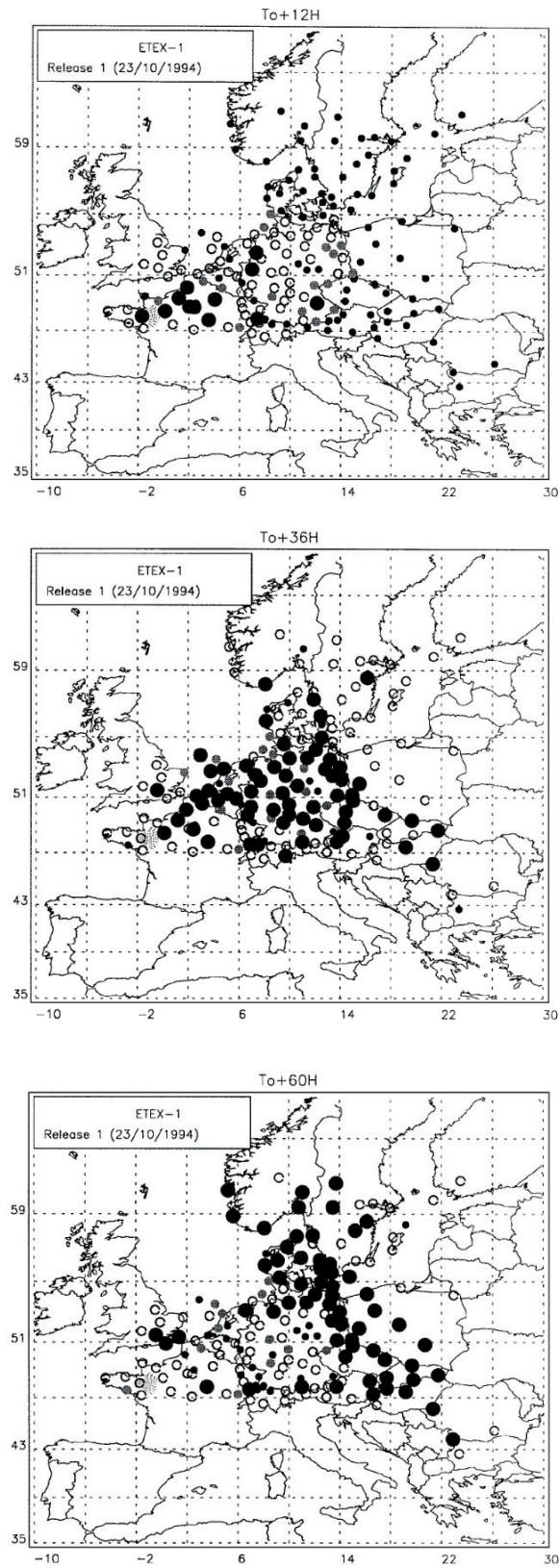
Figure 3.2: This figure appears in [51]. It shows the PMCH detected at ground-level stations in Europe, 12, 36, and 60 hours after the first release. The black dots show the stations where the tracer is found at each moment.

Air samples were taken at 168 stations in 17 European countries. Figure 3.1 shows with red dots the locations of the stations. Each sampling site was equipped with a sequential air sampler. To determine the concentration of PFCs in the atmosphere, air samples were taken for subsequent laboratory analysis. Samples were taken at each station at three-hour intervals for 96 hours from the time of release. At each site 24 consecutive samples were collected, with the initiation of sampling progressively delayed with distance from the release site, according to the expected arrival time of the tracer plume.

The plots in Figure 3.2 show how the PMCH tracer moves 12, 36, and 60 hours after the start of the release. The plume traverses Europe in a northeastern direction.

To estimate the atmospheric dispersion of the tracer, we use FLEXPART (see Chapter 1 for more information), a long-range transport model. In fact, the ETEX experiment was initially designed to validate this kind of model. However, in this thesis, we use the ETEX dataset instead to study the nature of the errors in this model and to evaluate source estimation algorithms.

3.3   ETEX reveals error properties

In the ETEX experiment we have access to the measurements $\widehat{\mathbf{y}}$, the true source $\mathbf{x}$, and the estimated transport matrix $\widehat{\mathbf{A}}$. This permits us to study the errors $\mathbf{e}$ modelled as

$$\widehat{\mathbf{y}} = \widehat{\mathbf{A}}\mathbf{x} + \mathbf{e}. \tag{3.2}$$

Let us model the components e_i of the vector $\mathbf{e}$ as random and independent and identically distributed (i.i.d.). Some degree of correlation may exist among the errors, but this correlation is unknown. The first aspect of the errors that we will study is whether they come from a Gaussian distribution. We can do so by computing the *kurtosis* of the sample distribution. Kurtosis was first introduced by Karl Pearson [54] at the beginning of the last century. It is a measure of the *tailedness* of the probability distribution of a real-valued random variable. The kurtosis of a random variable X given its mean μ is defined as

$$\mathrm{Kurt}(X) = \frac{\mathbb{E}[(X - \mu)^4]}{\mathbb{E}[(X - \mu)^2]^2}. \tag{3.3}$$

The kurtosis of any univariate normal distribution is 3. The distributions with a kurtosis greater than 3 are called *leptokurtic*. Such distributions produce more outliers than the Gaussian distribution. One example of leptokurtic distribution is the Laplace distribution.

The kurtosis of the sample distribution of $\mathbf{e}$ is $\mathrm{Kurt}(e) = 123.64$. This indicates that the underlying distribution of the errors is strongly leptokurtic. This empirical distribution clearly deviates from a Gaussian one. This is why we should consider a more complex model for the errors in the problem, that may explain better these results.

Now, let us look at the constituent components of $\mathbf{e}$. Modelling and predicting the behaviour of the atmosphere is a complicated task. Thousands of input parameters have to be taken into consideration, and the propagation of errors in the model is an unavoidable fact. On the other hand, measurements are collected with high quality instrumentation, but are still not perfect. These two different types of error - model and measurement errors - have very different characteristics. We model the former as a multiplicative error, $\mathbf{N}$, and the latter as an additive one, $\mathbf{n}$. Taking both into account, $\widehat{\mathbf{A}}$ being the estimate of the transport matrix produced by FLEXPART, the problem can be reformulated as

$$\widehat{\mathbf{y}} = (\mathbf{A} + \mathbf{N})\mathbf{x} + \mathbf{n} = \widehat{\mathbf{A}}\mathbf{x} + \mathbf{e}, \tag{3.4}$$

where **e** is now the total error.

Using the ℓ_2 norm in the loss function as in (1.10) is optimal when the additive errors are Gaussian - which is not our case. Even worse, this loss function is very sensitive to outliers. Hence, there is significant potential to improve the performance of (1.10) and its variants by additional processing, aimed at removing and/or marginalizing the outliers. In the present chapter we propose and demonstrate two novel schemes for this additional processing.

3.4   Outlier detection by combining different model matrices

Imagine that we have an oracle which reveals to us the measurements corresponding to the largest errors (i.e. the outliers). If we remove these measurements from the dataset, the performance of (1.10) in terms of the reconstruction error or mean square error (MSE) improves significantly. In order to illustrate this, since we have access to the ground truth of **x** in the case of ETEX, we remove the measurements associated with the largest errors (sorted by magnitude) and observe the effect on the MSE. Figure 3.3 shows how the MSE decreases as more and more outliers are removed. Some oscillations may occur due to outlier compensation effects. Figure 3.4 shows one example of this effect.
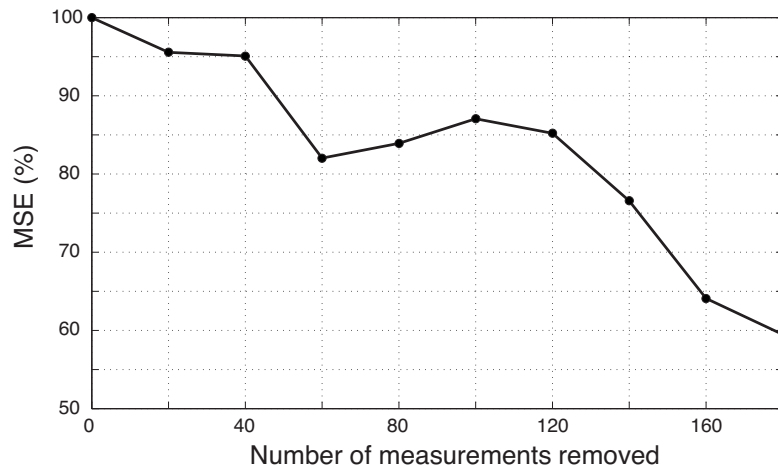

Figure 3.3: Reduction of the MSE of reconstruction obtained using (1.10). The strongest outlier measurements (the ones associated with the largest errors) have been removed manually. Notice that the MSE decreases as more outliers are removed.

However, in a real-world problem, we do not have such an oracle. The question becomes: how can we locate the outliers blindly?

A key observation we would like to make concerns the estimate of the model of the system, i.e. the matrix $\widehat{\mathbf{A}}$. To compute it, FLEXPART uses meteorological fields (analyses or forecasts) from the European Centre for Medium-Range Weather Forecasts (ECMWF) numerical weather prediction models [63]. Obviously, creating the model $\widehat{\mathbf{A}}$ using two different sets of fields will result in two slightly different matrices. The key idea is that we can use the variation between two or more of these matrices as an indication of the uncertainty in the modelling process, and
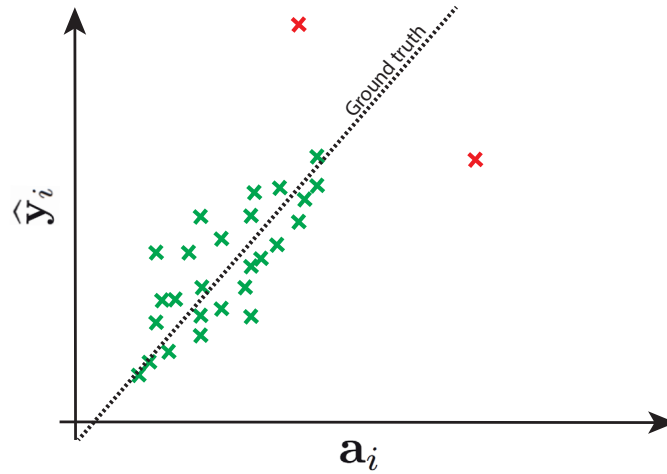
Figure 3.4: Example of two outliers (marked in red) that compensate each other. When both outliers are present, they do not deviate the LS estimate. However, if one of them is removed, the other will deviate the LS estimate.

thus the likelihood that a certain measurement is an outlier. Intuitively, if a particular entry in the model matrix $\widehat{\mathbf{A}}$ varies significantly between different models of the same process, it is likely to be estimated poorly. It will thus likely cause a large entry in its element in $\mathbf{N}$, which will increase the entry in the corresponding element in $\mathbf{e}$.

Let $\widehat{\mathbf{A}}_1$ and $\widehat{\mathbf{A}}_2$ be two model matrices generated with different meteorological fields. We can compute the matrix of absolute differences as

$$\mathbf{K} = |\widehat{\mathbf{A}}_1 - \widehat{\mathbf{A}}_2|. \tag{3.5}$$

The accumulated differences in each row $\mathbf{k}$, reveal the likelihood of an outlier occurring in each measurement

$$\mathbf{k} = \mathbf{K}\mathbf{1}, \tag{3.6}$$

where $\mathbf{1}$ is a column vector whose entries are equal to one.

Our approach is to extend the original approach shown in (1.13) by applying weights on the residuals, together with a non-negative constraint on the solution $\hat{\mathbf{x}}$. We give more weight to the most reliable samples, and vice versa. Based on (3.6) we propose to use different weighting schemes. We first consider exponential weights, such that

$$\mathbf{w}_e = e^{-\mathbf{k}}, \tag{3.7}$$

where the exponentiation is element-wise. An alternative weight function is to define thresholded weights as

$$\mathbf{w}_\beta = \begin{cases} \frac{1}{\mathbf{k}_i + 1} & \text{if } \mathbf{k}_i \geq \beta \\ 1 & \text{if } \mathbf{k}_i < \beta \end{cases} \tag{3.8}$$

where β is a predefined threshold. These weights only decrement the influence of a residual i if $\mathbf{k}_i$ is larger than a pre-defined threshold.

Our problem can then be formulated as an outlier-weighted LS with a non-negative Tikhonov regularization:

$$\widehat{\mathbf{x}} = \quad \arg\min_{\mathbf{x}} \|\mathbf{W}(\widehat{\mathbf{A}}\mathbf{x} - \widehat{\mathbf{y}})\|_2 + \lambda\|\mathbf{x}\|_2$$
$$\text{s.t. } \mathbf{x} \succeq 0, \tag{3.9}$$

where $\mathbf{W}$ is a diagonal weighting matrix, whose main diagonal is either $\mathbf{w}_e$ or $\mathbf{w}_\beta$.

To illustrate our outlier mitigation algorithms, we employ them in two different experiments: a simple synthetic example and a controlled example using real-world measurements and models.

Toy problem

Let us create a simple synthetic example. The measurements are generated using a model matrix $\mathbf{A} \in \mathbb{R}^{260 \times 10}$ with random uniform entries, a piece-wise constant source vector $\mathbf{x}$ shown in Fig. 3.5, and no additive noise

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{3.10}$$

Next, let us create two noisy estimates of the model matrix, $\widehat{\mathbf{A}}_1$ and $\widehat{\mathbf{A}}_2$. We build them by changing some of the entries of $\mathbf{A}$

$$\widehat{\mathbf{A}}_1 \quad = \mathbf{A} + \mathbf{N}_1$$
$$\widehat{\mathbf{A}}_2 \quad = \mathbf{A} + \mathbf{N}_2, \tag{3.11}$$

where $\mathbf{N}_1$ and $\mathbf{N}_2$ are random sparse matrices. The goal is to recover $\mathbf{x}$ using only $\mathbf{y}$, $\widehat{\mathbf{A}}_1$, and/or $\widehat{\mathbf{A}}_2$. Both $\widehat{\mathbf{A}}_1$ and $\widehat{\mathbf{A}}_2$ are well conditioned and thus, the Tikhonov regularization term is not necessary. We recover $\mathbf{x}$ using these variations on (3.9):

1. **LS**: $\mathbf{W} = \mathbf{I}$, $\lambda = 0$.
2. **WLS-exp**: $\mathbf{W} = \text{diag}(\mathbf{w}_e)$ as in (3.7), $\lambda = 0$.
3. **WLS-thr**: $\mathbf{W} = \text{diag}(\mathbf{w}_\beta)$ as in (3.8), $\beta = 1.5$, $\lambda = 0$.

For WLS-thr, the threshold value $\beta = 1.5$ is found experimentally as the value that minimizes the MSE of reconstruction. Fig. 3.5 shows the MSE for each of these algorithms. Clearly, the weights improve the overall performance of recovering the source, and the best performance is achieved using exponential weights.

Controlled, real-data experiment

Earlier we used the ETEX dataset [51] to characterize the noise. Now, we will use it to test the performance of our method on a real dataset.

System model matrices are now generated using FLEXPART [63]. To generate both $\widehat{\mathbf{A}}_1$ and $\widehat{\mathbf{A}}_2$, we use as input to FLEXPART atmospheric analysis fields from two different meteorological re-analysis projects: ERA 40 [70] and ERA Interim[30], both of which were developed by ECMWF. A meteorological re-analysis project [3]

...in addition to re-analysing all the old data using a consistent system, it also makes use of much archived data that was not available to the original analyses. This allows for the correction
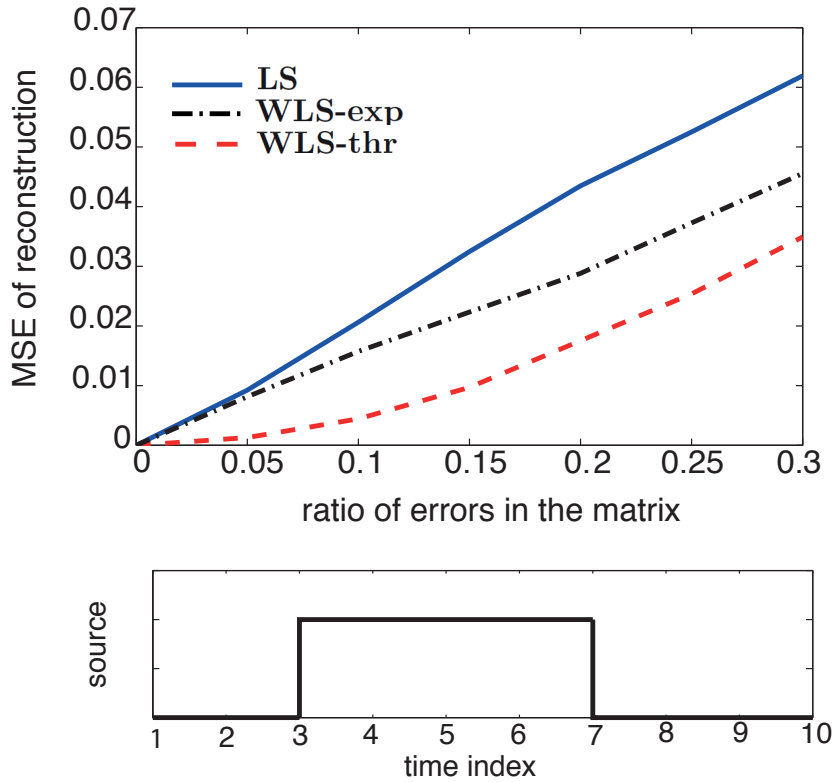
Figure 3.5: In the upper part of the figure we see the MSE in the toy problem when the solution is recovered using the three algorithms. The ratio of errors in the matrix is defined as the number of nonzero entries in $\mathbf{N}$ divided by the total number of entries in $\mathbf{N}$. In the lower part of the figure we can see the source that we use in this toy problem.

of many historical hand-drawn maps where the estimation of features was common in areas of data sparsity. The ability is also present to create new maps of atmosphere levels that were not commonly used until more recent times.

ERA 40 generates re-analyses from 1957 to 2002. ERA Interim is a new product, aimed at replacing ERA 40, and covers the period 1979 to present. Both ERA 40 and ERA Interim fields are public datasets and can be downloaded from the ECMWF website [2].

Using $\widehat{\mathbf{A}}_1$ and $\widehat{\mathbf{A}}_2$, we generate the matrix of absolute differences, $\mathbf{K}$, and the accumulated sum for each row, $\mathbf{k}$. Both are shown in Fig. 3.6. The rows with larger accumulated sum of differences can be identified easily.
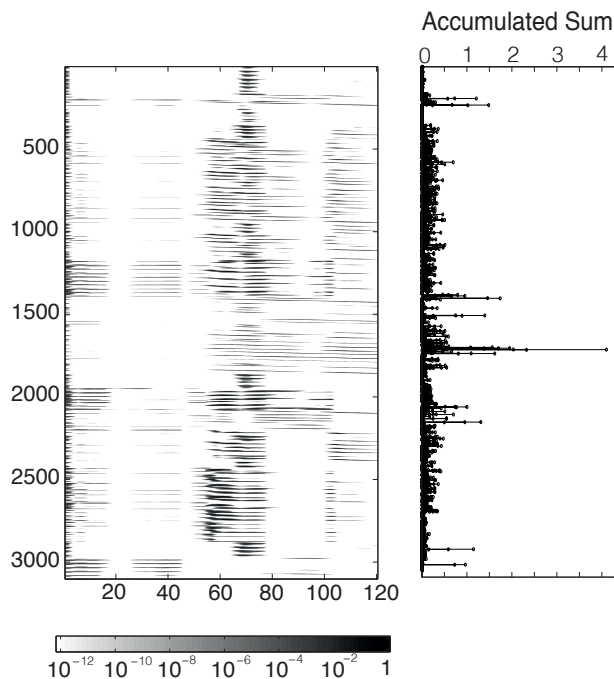
Figure 3.6: The plot on the left represents the matrix of differences $\mathbf{K}$ in the ETEX experiment. On the right we observe its corresponding accumulated row sums.

This time around, the condition numbers of both $\widehat{\mathbf{A}}_1$ and $\widehat{\mathbf{A}}_2$ are on the order of 10^{17}. Hence, we need to include the Tikhonov regularization term to recover $\mathbf{x}$. We recover $\mathbf{x}$ using these variations on (3.9):

1. **LS**: $\mathbf{W} = \mathbf{I}$, $\lambda \in [30, 400]$.
2. **WLS-exp**: $\mathbf{W} = \text{diag}(\mathbf{w}_e)$ as in (3.7), $\lambda \in [30, 400]$.
3. **WLS-thr**: $\mathbf{W} = \text{diag}(\mathbf{w}_\beta)$ as in (3.8), $\beta = 3.5 \times \text{mean}(\mathbf{k})$, $\lambda \in [30, 400]$.

For WLS-thr, we manually adjust β to 3.5 times the average of $\mathbf{k}$. Fig. 3.7 shows the comparison of results using the three algorithms. Here, just like in the toy example, the use of weights to reduce the influence of outliers improves the recovery. However, unlike in the toy example, thresholded weights perform better than the exponential ones. However, the use of the thresholded weights requires manual adjustment of the parameter β, which might be difficult to realize in practice.
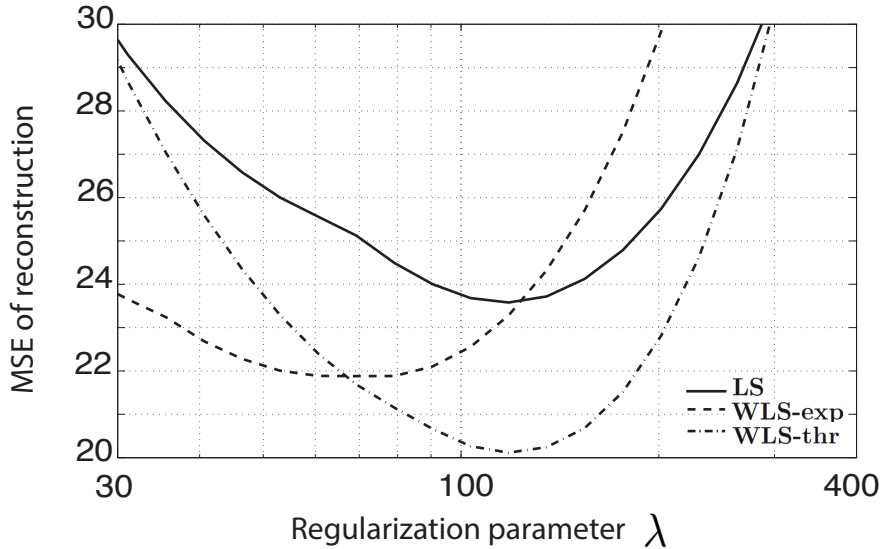
Figure 3.7: Comparison of the performance of three source estimation algorithms using the ETEX dataset: LS, WLS with exponential weights, and WLS with thresholded weights. All three algorithms also utilize Tikhonov regularization.

3.5 Outlier detection using random sampling techniques

In this section we take a different approach to outlier detection, for which we only need one model matrix. For that we will start by using the random sample consensus (RANSAC) algorithm. This is one of the simplest and most popular algorithms to localize outliers blindly. RANSAC has been widely and successfully used, mainly by the computer vision community ([62]). Figure 3.8 illustrates the operation of RANSAC, and Algorithm 3.1 describes it in pseudocode.

Given a set of m measurements, $\widehat{\mathbf{y}}$, select randomly a subset $\mathbf{y}_s$ containing p measurements such that $n < p < m$, where n is the number of unknowns in the problem. In Figure 3.8, $m = 8$ and $p = 2$, and the subset $\mathbf{y}_s$ is shown as red diamonds. Using (1.13) and $\mathbf{y}_s$, estimate $\widehat{\mathbf{x}}_s$, and then compute the residual $\mathbf{r}_s = \mathbf{A}\widehat{\mathbf{x}}_s - \widehat{\mathbf{y}}$. Next, we want to count how many of the original samples are *inliers*. For a given tolerance η, the set of inliers is defined as

$$\mathcal{L}_s = \{q \in \{1, 2, \cdots, m\} \mid \eta \geq (\mathbf{r}_s[q])^2\}. \tag{3.12}$$

Repeat this process N times and declare the final solution $\widehat{\mathbf{x}}^*$ to be that estimate $\widehat{\mathbf{x}}_s$ which produced the most inliers. In Figure 3.8, $N = 2$.

Note that other regularizations can be used instead of (1.13). Here we use the non-negative Tikhonov regularization because it is simple, general, and most other existing approaches are
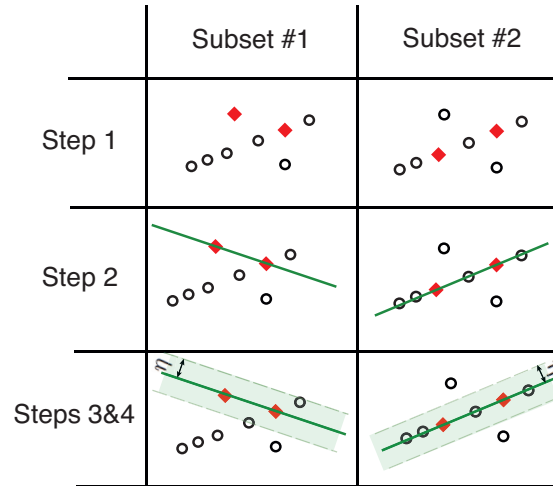
Figure 3.8: Visual representation of how RANSAC operates. Subset 1 and 2 represent two RANSAC iterations. The subset of measurements selected by RANSAC in each iteration is represented with red diamonds. Subset 1 contains one outlier. Hence, the solution corresponding with this subset generates fewer inliers than subset 2, which is free of outliers.

Algorithm 3.1 RANSAC

   INPUT: $\widehat{\mathbf{y}} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, λ, η, N, p

Require: $\lambda \geq 0, N > 0, \eta \geq 0, p \leq m$

  $\mathcal{L}^* \leftarrow \emptyset$

  $\widehat{\mathbf{x}}^* \leftarrow \mathbf{0} \in \mathbb{R}^n$

  $\mathbf{r} \leftarrow \mathbf{0} \in \mathbb{R}^m$

  $\mathbf{k} \leftarrow \mathbf{0} \in \mathbb{N}^p$

  $\mathbf{y}_s \leftarrow \mathbf{0} \in \mathbb{R}^p$

  $\mathbf{A}_s \leftarrow \mathbf{0} \in \mathbb{R}^{p \times n}$

  for $s = 1$ to N **do**

    $\mathbf{k} \leftarrow p$ unique random integers from $[1, m]$

    $\mathbf{y}_s \leftarrow \widehat{\mathbf{y}}[\mathbf{k}]$

    $\mathbf{A}' \leftarrow \mathbf{A}[\mathbf{k}, :]$

    $\widehat{\mathbf{x}}_s \leftarrow \arg\min_{\mathbf{x}} \|\mathbf{A}_s \mathbf{x} - \mathbf{y}_s\|_2^2 + \lambda \|\mathbf{x}\|_2^2$

    $\mathbf{r} \leftarrow \mathbf{A}\widehat{\mathbf{x}}_s - \widehat{\mathbf{y}}$

    $\mathcal{L} \leftarrow \{q \in \{1, 2, \cdots, m\} \mid \eta \geq (\mathbf{r}[q])^2\}$

    if $\#\mathcal{L} > \#\mathcal{L}^*$ **then**

      $\mathcal{L}^* \leftarrow \mathcal{L}$

      $\widehat{\mathbf{x}}^* \leftarrow \widehat{\mathbf{x}}$

    end if

  end for

  return $\widehat{\mathbf{x}}^*$

---

based on it. Nevertheless, if some properties of the source are known a priori (e.g., sparsity or smoothness), this step of the algorithm can be adapted accordingly.

As we will see in the following sections, if the optimal value for the threshold parameter η is known and used, using RANSAC as a pre-processing stage for outlier removal before applying (1.13) significantly improves the overall performance (compared to using only (1.13) with no outlier removal pre-processing). Unfortunately, the performance of RANSAC depends strongly on the parameter η, and finding the optimal value of η blindly is an open problem.

Now, let us remember that one possible goal for a source estimation algorithm is to find the estimate $\widehat{\mathbf{x}}$ that minimizes the MSE, $\frac{1}{n}\|\mathbf{x} - \widehat{\mathbf{x}}_s\|_2^2$. Since the ground truth $\mathbf{x}$ is generally unknown, so is the MSE. What RANSAC in fact does is use an indirect metric of the MSE: it assumes that the number of inliers is inversely proportional to the MSE. Unfortunately, this assumed inverse proportionality does not always hold in the presence of *critical measurements*, as explained in the next section.

3.5.1 Critical measurements

RANSAC operates reliably when all the measurements are of similar importance, because the inverse proportionality between MSE and the number of inliers holds. However, when critical measurements are present, this proportionality does not hold, and RANSAC does not identify all the outliers correctly. This is illustrated in Figures 3.9(a) and Figures 3.9(b).

We identify critical measurements as those which have the largest influence in the source estimation process. A quantitative measure of influence is the Cook distance [26]. This measure was introduced by R. D. Cook in 1977. He proposed a measure of influence of an observation based on how the estimate would change if we do not use that observation. We define the Cook distance as the standardized difference between $\widehat{\mathbf{x}}(i)$, the estimate obtained by omitting the i-th observation, and $\widehat{\mathbf{x}}$, the estimate obtained using all the data. If we define the residuals as $\mathbf{r} = \widehat{\mathbf{y}} - \mathbf{A}\widehat{\mathbf{x}}$, and $s = \|\mathbf{r}\|_2^2/(m-n)$, then the Cook distance for the i-th observation can be written as

$$D_i = \frac{\|\mathbf{A}(\widehat{\mathbf{x}}(i) - \widehat{\mathbf{x}})\|_2^2}{ns^2}. \tag{3.13}$$

Figure 3.10 shows the Cook distance of the measurements in the ETEX dataset. It is easy to observe the peak that identifies the critical measurements: their Cook distance is several times larger than the Cook distance of the rest of the measurements.

Let us consider again the ETEX dataset, the set of N solutions $\widehat{\mathbf{x}}_s$ that RANSAC generates, and their corresponding residuals $\mathbf{r}(\widehat{\mathbf{x}}_s)$. Figure 3.11 compares the residuals $\mathbf{r}(\widehat{\mathbf{x}}^*)$ given by the solution $\widehat{\mathbf{x}}^*$ that produces most inliers (in blue) and the residuals $\mathbf{r}(\widehat{\mathbf{x}}_m)$ given by the solution $\widehat{\mathbf{x}}_m$ that produces the minimum MSE (in red). The former is the solution selected by RANSAC, and the latter is the actually best solution in the minimum MSE sense. The main difference between $\mathbf{r}(\widehat{\mathbf{x}}^*)$ and $\mathbf{r}(\widehat{\mathbf{x}}_m)$ is that the values of $\mathbf{r}(\widehat{\mathbf{x}}^*)$ in the positions of the critical measurements (marked by two arrows) are several times larger than the rest of the values. In other words, RANSAC discarded the critical measurements as outliers. By doing so, RANSAC achieved more inliers, but completely missed the minimum MSE.
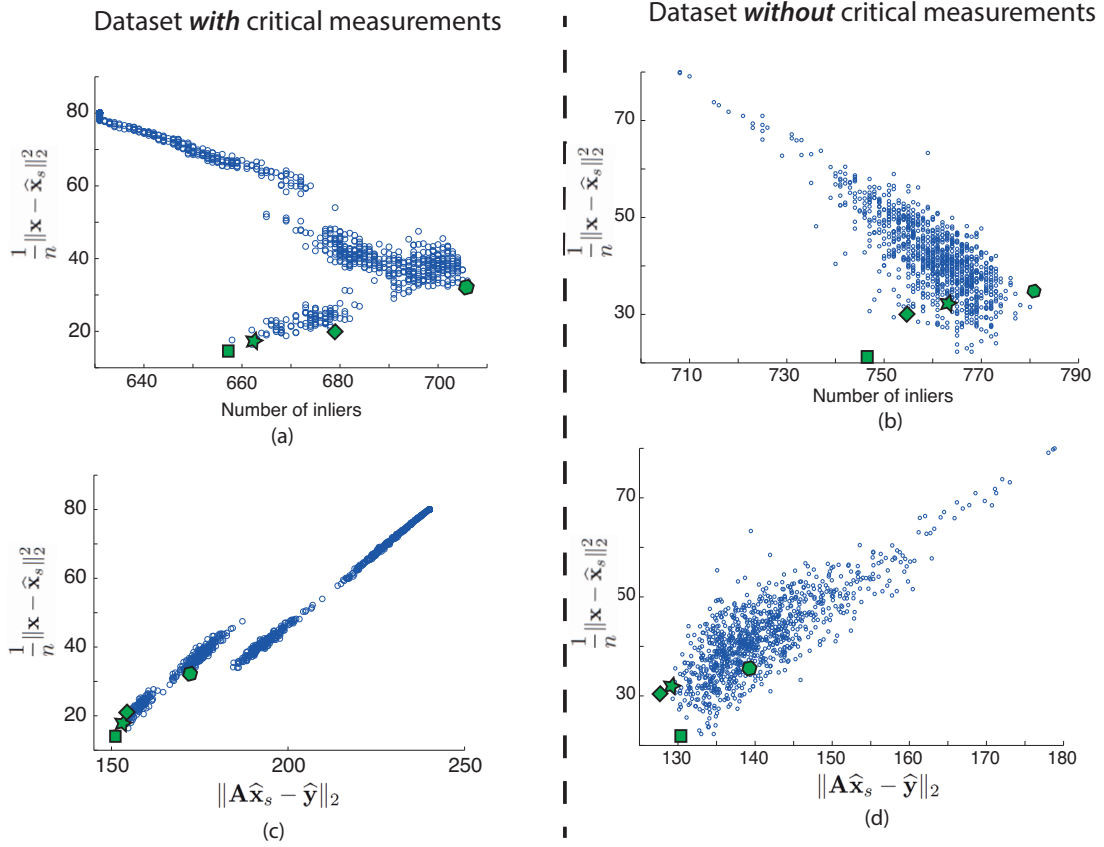
Figure 3.9: Performance of RANSAC and TRANSAC in the ETEX dataset. (a) and (b) show graphically the correlation between MSE of reconstruction and the number of inliers. (c) and (d) show graphically the correlation between MSE of reconstruction and the total residual. To build (a) and (c) the complete dataset was used, to build (b) and (d) the dataset without critical measurements was used. The diamond indicates the solution obtained by traditional Tikhonov regularization in (1.10), the star indicates the solution chosen by TRANSAC *before* the voting stage, the square indicates the final solution of TRANSAC, and the hexagon the solution chosen by RANSAC.

3.5.2   TRANSAC

In order to avoid this weakness of the standard RANSAC algorithm, we propose an alternative indirect metric to discriminate solutions with small MSE: the *total residual* $\epsilon_s = \|\mathbf{A}\widehat{\mathbf{x}}_s - \widehat{\mathbf{y}}\|_2$. We called it *total* residual because it is the residual produced by the solution $\widehat{\mathbf{x}}_s$ (which was computed using just the subset of measurements $\mathbf{y}_s$) with the complete set of measurements $\widehat{\mathbf{y}}$. We will show that using the total residual instead of the number of inliers to select the best solution gives superior results in the minimum MSE sense.

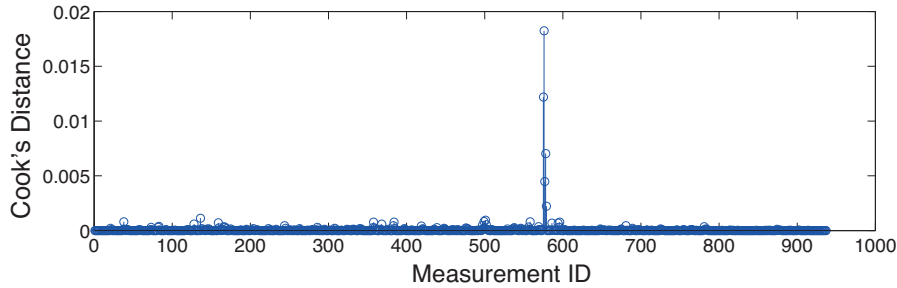Figures 3.9(c) and 3.9(d) show that the total residual is directly proportional to the MSE of

Figure 3.10: Cook distance of the measurements in the ETEX dataset.

reconstruction. Unlike the number of inliers (see Figures 3.9(a) and 3.9(b)), this proportionality is conserved also when critical measurements are present in the dataset. In fact, if we look at Figures 3.9(c) and 3.9(d), we see that the total residual and the MSE error can be related in a linear manner. If z is additive noise, and α is a non-negative scalar then

$$\|\mathbf{A}\widehat{\mathbf{x}}_s - \widehat{\mathbf{y}}\|_2 = \alpha\|\mathbf{x} - \widehat{\mathbf{x}}_s\|_2^2 + z. \tag{3.14}$$

In a real-life problem, where we do not have access to the ground truth, we do not know if critical measurements are present in a dataset. Hence, we need a robust metric like the total residual.

By replacing the number of inliers by the total residual metric, we create the first step of the Total residual RANdom SAmple Consensus (TRANSAC) algorithm. The second step consists in a *voting* stage. This step is in fact a denoising process, that mitigates the effects of z. Both steps are described in Algorithm 3.2 in pseudocode.

Because of the randomness introduced by z, even if a candidate solution has the smallest total residual, it is not guaranteed to be the solution with the smallest MSE. The intention of the voting stage is, using the candidate solutions with a total residual under a certain threshold, to come up with the best possible denoised final solution.

Intuitively, the solutions with the smallest total residual are generated using almost outlier-free random subsets of measurements $\mathbf{y}_s$. We refer to these as the *good* subsets. Outliers can appear sporadically in some of these good subsets, but the same outlier is unlikely to appear in all of them. Hence, in the voting stage we select the measurements that all the good subsets have in common, or in other words, exclude any measurements that appear very infrequently. Thus, the voting stage of TRANSAC has the following steps:

1. We first select the subsets $\mathbf{y}_s$ associated with candidate solutions $\widehat{\mathbf{x}}_s$ with a total residual ϵ_s smaller than a certain threshold β.

2. Then, for each measurement in $\widehat{\mathbf{y}}$ we count how many times (*votes*) it appears in these good subsets. For that we build a histogram.

3. Finally, we select the M measurements $\mathbf{y}_M$ with the largest frequency of occurrence.

4. We consider that these M measurements $\mathbf{y}_M$ do not contain outliers, and we use them to produce the final estimate $\widehat{\mathbf{x}}^*$ using (1.10).
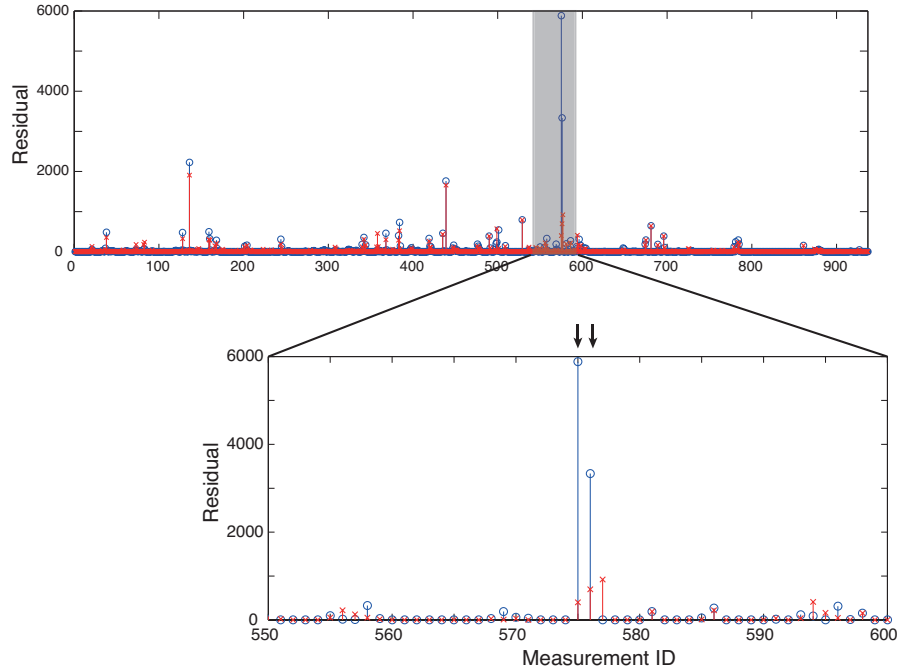
Figure 3.11: Residuals of two different source estimations: The blue peaks correspond to the residual produced by the solution $\widehat{\mathbf{x}}$ given by RANSAC in Figure 3.9(a) (the solution that produces the largest number of inliers). The red peaks correspond to the residual produced by the solution $\widehat{\mathbf{x}}$ with the smallest MSE in 3.9(a). The black arrows on the top indicate where the two most critical measurements are located. Clearly, the residuals of the RANSAC solution (blue peaks) corresponding to these two critical measurements are much larger than that of the minimum MSE solution.

Sensitivity to the paramater M

In Section 3.5.2 we confirmed the expected behaviour of the first stage of TRANSAC: we showed that the total residual is proportional to the MSE, as shown in Figures 3.9(c) and 3.9(d). Let us check now the second stage, the voting, and its sensitivity to M, the number of measurements that we select. It is important to note that $n \leq M \leq m$, since we need at least as many equations as unknowns, and we do not have more than m measurements.

Let us first perform a sanity-check experiment. We will suppose that during the first step in the voting we have access to the actual MSE of every candidate solution $\widehat{\mathbf{x}}_s$. Then, we will of course select the good subsets which in fact have the smallest MSE, and use them to build the histogram in step 2. This histogram counts how many times each measurement appears in the good subsets. In step 3 we select the M measurements that appear most often, and in the fourth step $\widehat{\mathbf{x}}^*$ is computed from those M measurements. We run this modified TRANSAC with the dataset without critical measurements. In the experiments we use the value of β which produces the minimum MSE at $M = n$. Figure 3.12(a) shows the MSE of the $\widehat{\mathbf{x}}^*$ obtained for different values of the parameter M.

Algorithm 3.2 TRANSAC

    INPUT: $\widehat{\mathbf{y}} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, λ, N, p, M, β

Require: $\lambda \geq 0, N > 0, p \leq m, n \leq M \leq m, \beta \geq 0$

    $\boldsymbol{\epsilon} \leftarrow \mathbf{0} \in \mathbb{R}^N$

    $\mathbf{k} \leftarrow \mathbf{0} \in \mathbb{N}^p$

    $\mathbf{K} \leftarrow \mathbf{0} \in \mathbb{N}^{p \times N}$

    $\mathbf{y}_s \leftarrow \mathbf{0} \in \mathbb{R}^p$

    $\mathbf{A}_s \leftarrow \mathbf{0} \in \mathbb{R}^{p \times n}$

    $\mathcal{G} \leftarrow \emptyset$

    $\mathbf{h} \leftarrow \mathbf{0} \in \mathbb{R}^m$

    $\mathbf{b} \leftarrow \mathbf{0} \in \mathbb{R}^M$

    for $s = 1$ to N **do**

        $\mathbf{k} \leftarrow p$ unique random integers from $[1, m]$

        $\mathbf{y}_s \leftarrow \widehat{\mathbf{y}}[\mathbf{k}]$

        $\mathbf{A}_s \leftarrow \mathbf{A}[\mathbf{k}, :]$

        $\widehat{\mathbf{x}}_s \leftarrow \arg\min_{\mathbf{x}} \|\mathbf{A}_s \mathbf{x} - \mathbf{y}_s\|_2^2 + \lambda \|\mathbf{x}\|_2^2$

        $\boldsymbol{\epsilon}[s] \leftarrow \|\mathbf{A}\widehat{\mathbf{x}}_s - \widehat{\mathbf{y}}\|_2$

        $\mathbf{K}[:, s] \leftarrow \mathbf{k}$

    end for

    $\mathcal{G} \leftarrow \{q \in \{1, 2, \cdots, N\} \mid \boldsymbol{\epsilon}[q] \leq \beta\}$

    $\mathbf{K}_{\mathcal{G}} \leftarrow \mathbf{K}[:, \mathcal{G}]$

    $\mathbf{h}[k] \leftarrow$ how many times k appears in $\mathbf{K}_{\mathcal{G}}$, $\forall k \in \{1, 2, \cdots, m\}$

    $\mathbf{b} \leftarrow$ indices of the M largest elements of $\mathbf{h}$

    $\mathbf{y}^* \leftarrow \widehat{\mathbf{y}}[\mathbf{b}]$

    $\mathbf{A}^* \leftarrow \mathbf{A}[\mathbf{b}, :]$

    $\widehat{\mathbf{x}}^* \leftarrow \arg\min_{\mathbf{x}} \|\mathbf{A}^* \mathbf{x} - \mathbf{y}^*\|_2^2 + \lambda \|\mathbf{x}\|_2^2$

    return $\widehat{\mathbf{x}}^*$

The dashed vertical line on the right of the figure indicates the use of the whole measurement dataset, i.e. no outlier removal. The red horizontal line indicates the MSE of the solution obtained by using the whole measurement dataset, which is the MSE produced by using just (1.13) without any pre-processing. The dashed vertical line on the left indicates the most aggressive possible removal of outliers, so that $M = n$, and corresponds to using as many measurements as unknowns.

We can observe that the MSE of the solution increases as M increases. This is to be expected: as M grows, more outliers are included in the dataset $\mathbf{y}_M$ that is used to obtain $\mathbf{x}^*$, and its MSE increases. We note that the resulting curve is non-decreasing: remember that in this particular experiment we have access to the MSE and the histogram is built from the actual best candidate solutions. In other words, we know precisely which are the best samples to drop.

Next let us examine the performance of the actual TRANSAC algorithm, still without critical measurements. We run the same experiment as above, but we do not consider any more that we have access to the MSE of the solutions $\widehat{\mathbf{x}}_s$ produced with random sampling. Instead we use their total residuals. The results are shown in Figure 3.12(b). We observe that the MSE again increases as M increases, as before, and the maximum MSE still occurs close to $M = m$. Notice

that the minimum MSE again occurs when $M = n$. Important differences with respect to the previous case are that the MSE is not as low as before and is not a strictly increasing function.

Finally, let us examine now the whole dataset, including the critical measurements. Figure 3.12(c) shows the results. We can observe that, again, the maximum MSE occurs at $M = m$. In other words, outlier removal is still beneficial. On the other hand, the minimum MSE does not occur at $M = n$, but rather at $M = 330$. When blind outlier removal is extremely aggressive, i.e. when M is close to its minimal value, the critical measurements are removed from the selected dataset $\mathbf{y}_M$. As soon as they are included, i.e. $M > 300$, the solution improves significantly.

Sensitivity to the parameter β

The parameter β regulates which subsets $\mathbf{y}_s$ are considered as *good* subsets. Here we check the sensitivity of TRANSAC to this parameter. For that we use the complete dataset and we fix M to 330, which is the value that minimizes the MSE, and we vary the value for β. As shown in Figure 3.13, the performance of the algorithm changes for different values of β. However, outlier removal is almost always beneficial, i.e. for practically any value of β there is an improvement in performance compared to no outlier removal.
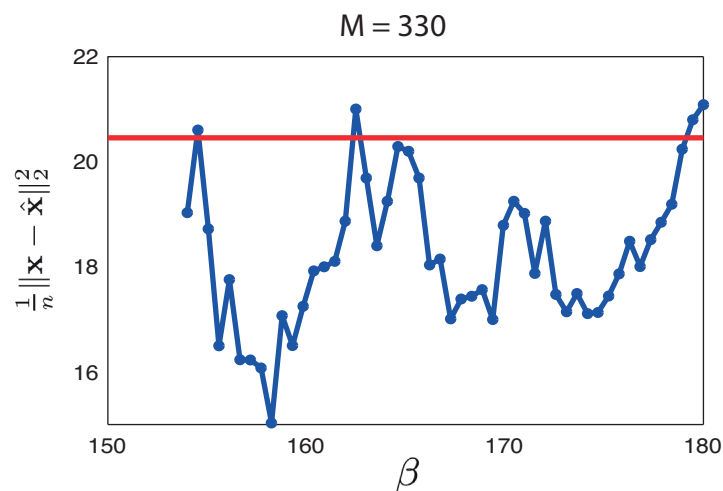

Figure 3.13: Sensibility of TRANSAC combined with LS with Tikhonov regularization to the parameter β. The red line indicates the MSE given by typical LS with non-negative Tikhonov regularization (1.13). The algorithm is sensitive to β, but in general outlier removal is almost always better than no outlier removal.

More regularizations

In this section our goals are to

1. study and compare the performance of TRANSAC with different regularizations, and
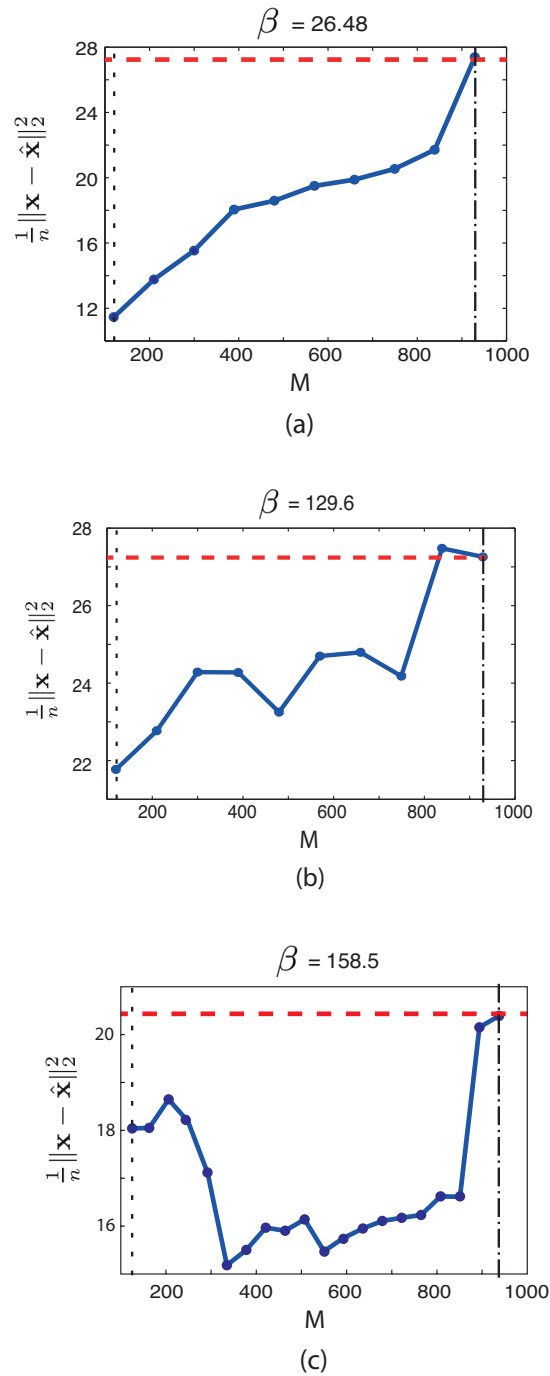2. compare the performances of RANSAC and TRANSAC.

Figure 3.12: Performance of TRANSAC combined with non-negative Tikhonov regularization. In the three plots, the red horizontal dashed line indicates the estimation error given by typical LS with non-negative Tikhonov regularization (1.13). The dashed line on the right indicates $M = m$, the one on the left indicates $M = n$. Plot (a) shows the results of the sanity check. As the selected number of measurements M increases, the MSE of the estimation increases. Notice that the maximum MSE corresponds to $M = m$. Plot (b) shows the results of applying TRANSAC to the ETEX dataset without critical measurements. Again, the MSE increases in general with M, and the maximum MSE appears close to $M = m$. Plot (c) shows the results of applying TRANSAC to the whole ETEX dataset, critical measurements included. In this case the MSE does not always increase with M, but the maximum MSE still corresponds to $M = m$.

We have shown the capacity of TRANSAC to improve the performance of LS with Tikhonov regularization. However, as explained in Section 3.4, TRANSAC is a blind outlier detection algorithm that can be combined with any regularization in order to improve its results. In this section we also combine TRANSAC with the regularization that we introduced in (3.1). This regularization enforces smoothness in the solution, and it has been previously used in the literature.

Each of the two base algorithms (given in 1.10 and 3.1) is used in three variations:

1. base algorithm with no outlier removal

2. RANSAC + base algorithm

3. TRANSAC + base algorithm

That makes for a total of six algorithm variations to compare. The parameters for each of the variations (λ, η, β, M) are set to their optimal values i.e. the values that minimize the MSE. These values are found using linear search. Thus, in this section we compare the best performance of each algorithm.

Figure 3.14 shows the MSE of the source estimated by the different algorithms. The blue bars correspond with the base algorithms (1.13 and 3.1), and the green ones indicate that TRANSAC is used for outlier removal. The plot on the left was generated using the whole ETEX dataset. The plot on the right was generated using the ETEX dataset without critical measurements. We can observe that, with and without critical measurements, TRANSAC improves the performance of *both* regularizations.

We also run this experiment using RANSAC, and get its optimal performance for every case. The results are also shown in Figure 3.14. The violet bars represent the results given when RANSAC is used for outlier removal. We can observe that RANSAC, in all the cases, also improves the performance of all the regularizations. Hence, our idea of removing outliers indeed does lead to improved performance, regardless of critical measurements or type of regularization. Another fact is that, in all cases TRANSAC shows higher performance than RANSAC. Therefore, our proposed modification of the metric, and the addition of the voting stage result in improved performance.

Figure 3.15 shows the results more qualitatively, by illustrating the estimated sources in every case. To observe the shapes of the estimates can be misleading, because our eye cannot give a quantitative assessment of the quality of the estimate. Nevertheless, we can identify some patterns that repeat in all the estimates:

1. In all the cases, the solution given by TRANSAC is less attenuated than the rest of the solutions.

2. Also, the value of the TRANSAC estimate outside the support of the ground truth is in general smaller than the rest of the estimates.

3. The regularization that uses the derivative and enforces smoothness (3.1) in this particular case helps more than Tikhonov to find a solution closer to the ground truth. However remember that the goal of this section is to show that adding a preprocessing stage to the base algorithms improves their performance. We are not looking for a regularization tailored for this particular problem. We would like to keep a general setup.
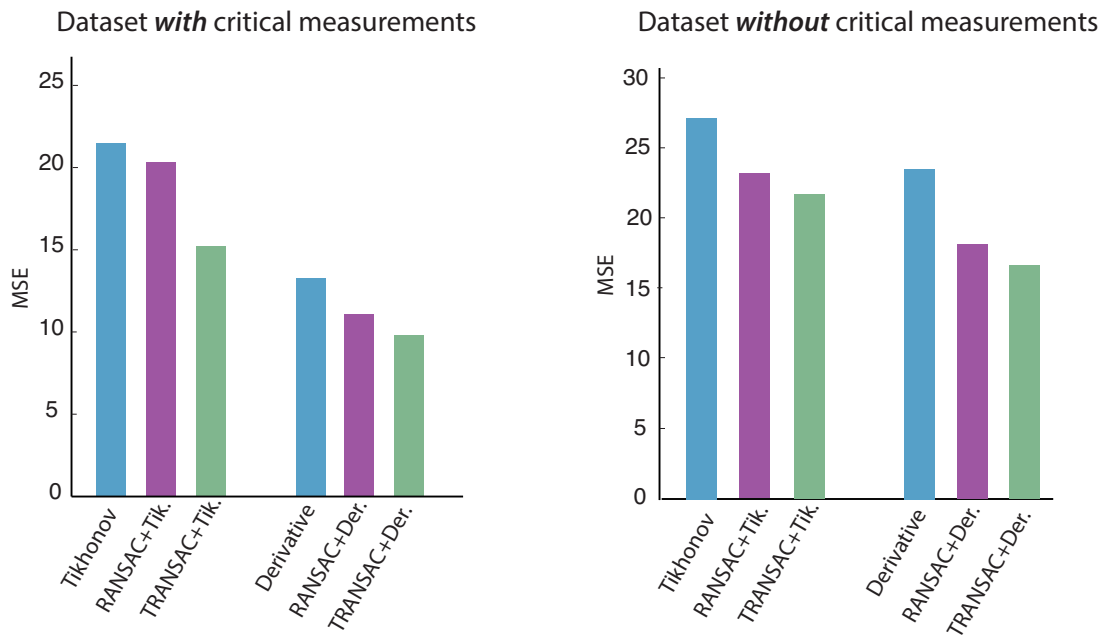
Figure 3.14: MSE of source estimated by different algorithms. The blue bars correspond to the base algorithms with no outlier removal (1.13 and 3.1). The violet bars indicate that RANSAC is used for outlier removal, and the green ones that TRANSAC is used for outlier removal. The plot on the left was generated using the whole ETEX dataset. The plot on the right was generated using the ETEX dataset without critical measurements. TRANSAC gives the best results in both cases.
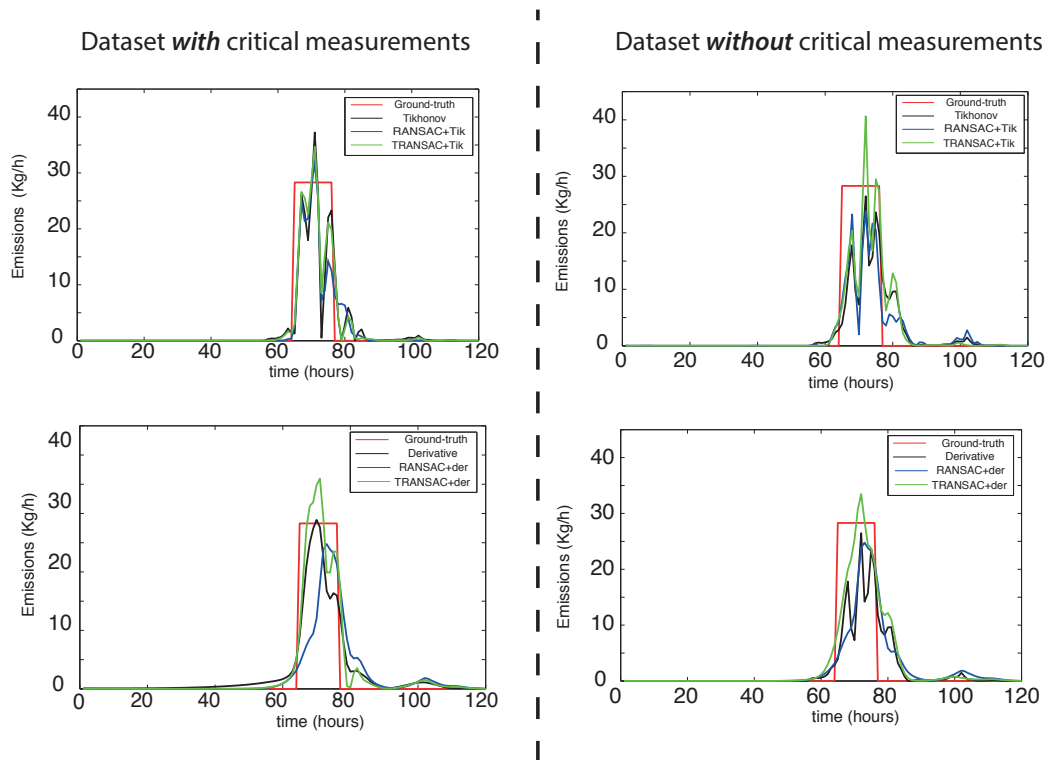
Figure 3.15: Source reconstructions given by the different algorithms. The difference between the plots on the top and on the bottom is the regularization used. The plots on the top were generated by combining (1.13) with RANSAC and TRANSAC. The plots on the bottom were generated by combining (3.1) with RANSAC and TRANSAC. The plots on the right were generated using the ETEX dataset without critical measurements. The plots on the left were generated using the whole ETEX dataset.

Chapter 4

Making the best of our mistakes

In the previous chapter we showed that our dataset indeed contains outliers. Thus, we need methods that are robust against these outliers, and are not overly affected by them. So, we proposed methods that find these outliers, and then remove them or downgrade their influence, as a preprocessing step to the traditional tools based on LS.

In this chapter [1] we go one step further and, instead of adding a pre-processing step, we modify the classic LS estimator itself, by replacing the ℓ_2 loss function with a different one, thus making it robust to outliers. We then combine this with one or more regularizations that make the problem well-posed.

4.1 Robust estimators

John W. Tukey was an american mathematician [19] and a true genius. His ideas were so original and their applicability so wide, that Tukey's work revolutionized many fields of study and even created new ones. For example, he invented the term *bit* as a contraction for a *binary digit*; he coined the word *software* and gave it the usage that we know today; he invented the box plot (see Figure 4.1); and he gave birth to the field of exploratory data analysis. His main contributions focused on statistics.

During World War II, Tukey worked in the Fire Control Research Office. There, he had access to many real datasets. After spending some time exploring the datasets, Tukey started to think of the outliers in his data. Before then, common practice was just to removed outliers manually. However his idea was that the estimators should be *robust* to the outliers. He realized the importance of designing methods for statistical analysis that have a reasonable performance even if some of their assumptions are violated [69].

After the War he founded the Princeton Study group with other statisticians [11]. Together they showed, using Monte Carlo simulations, that the outliers in the dataset may indeed deviate traditional estimators like the sample mean or LS. They also started to develop new estimators that

1. Downgrade the influence of contaminating measurements if they are present.

---

1. This chapter is the result of a collaboration with M. Muma, B. Bejar Haro, A. M. Zoubir and M. Vetterli [48], [47].
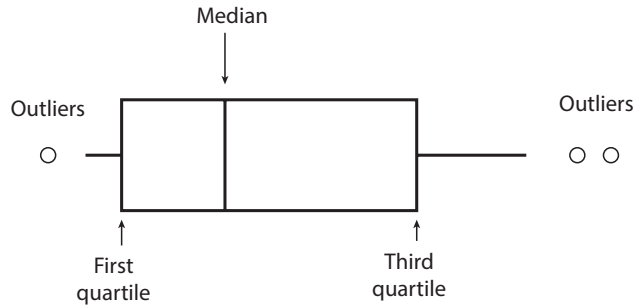
Figure 4.1: The box plot invented by J.W. Tukey.

2. Still produce correct answers if no contaminating measurements are present.

All these ideas gave birth to the new field of *robust statistics*. As a first step towards understanding this field, let us look at the prototypical robust estimator called the *M estimator*.

4.2   The M estimator

One of the simplest and most widely-used robust estimators is called the M estimator. To understand it, let us start from LS. We defined LS in (1.19) as

$$\widehat{\mathbf{x}}_{\text{LS}} = \arg\min_{\mathbf{x}} \|\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_2^2 = \arg\min_{\mathbf{x}} \sum_{i=1}^{m} (\widehat{y}_i - \mathbf{a}_i\mathbf{x})^2, \tag{4.1}$$

where $\mathbf{a}_i$ is the i-th row of $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m]^T$. The terms $(\widehat{y}_i - \mathbf{a}_i\mathbf{x})$ are commonly referred to as *residuals* r_i. Notice that each residual in the sum (1.19) is squared. This is the reason why LS is very sensitive to outliers. To see this, let us define $\rho(u) = u^2$, so we can write

$$\widehat{\mathbf{x}}_{\text{LS}} = \arg\min_{\mathbf{x}} \sum_{i=1}^{m} \rho(r_i). \tag{4.2}$$

In this case the $\rho(u)$ function is the square function depicted in Figure 4.2. Now imagine that there is an outlier y_o in our dataset (for example the red point in Figure 4.3) that produces a residual r_o whose magnitude $|r_o|$ is at least several times larger than the rest of the residuals, i.e. $|r_o| \geq \alpha|r_i|$. Then $\rho(r_o)$ is α^2 times larger than $\rho(r_i)$, $\rho(r_o) \geq \alpha^2\rho(r_i)$. This is represented in Figure 4.2. This means that the term $\rho(r_o)$ will dominate the sum in (4.2), as showed in Figure 4.4, and it will deviate $\widehat{\mathbf{x}}_{\text{LS}}$ from the ground truth, as Figure 4.3 shows.

One possible solution to avoid this effect is to replace the square ρ function. This is the main idea behind M estimator. For example, we can use the bi-square ρ function plotted in Figure 4.2. This function is very similar to the square function when the value of $|u|$ is small, but it is constant for large values of $|u|$. Figure 4.2 shows an example of this function, and it compares the value of the ρ square function and the ρ of a bisquare function for an outlier and for a regular residual.
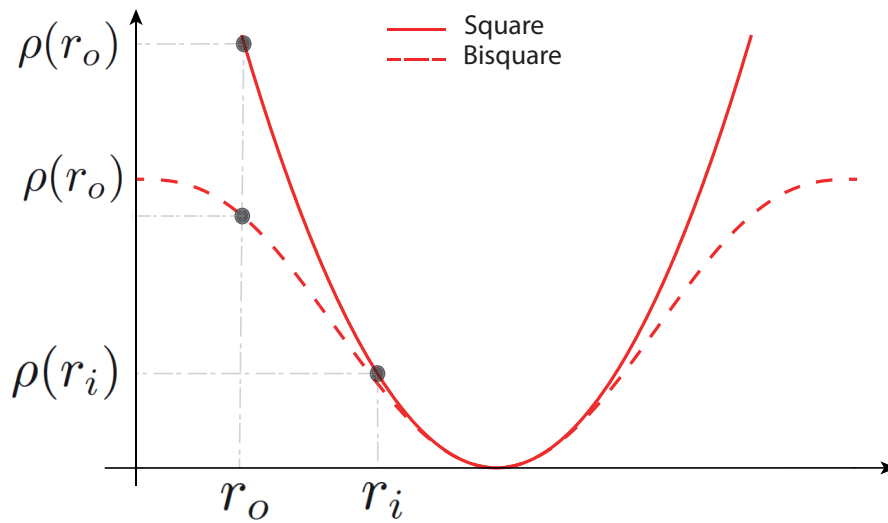
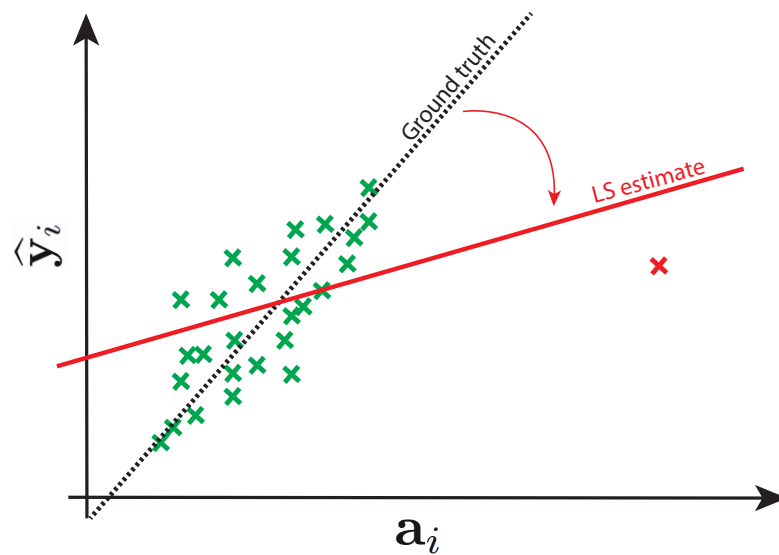**Figure 4.2:** Two different $\rho(u)$ functions.

Figure 4.3: Effect of one outlier in the LS estimate of a one dimensional problem.

The bisquare function was proposed by J.W. Tukey, and its analytic expression is given by

$$
\rho(u) = \begin{cases} \frac{c^2}{6}\left(1 - \left[1 - \left(\frac{u}{c}\right)^2\right]^3\right) & |u| \leq c \\ \frac{c^2}{6} & |u| > c. \end{cases} \tag{4.3}
$$

The parameter c, usually known as the *clipping parameter*, controls when the function becomes constant. To illustrate this point, in Figure 4.5 we show two bisquare functions with different clipping parameters, c_1 and c_2.
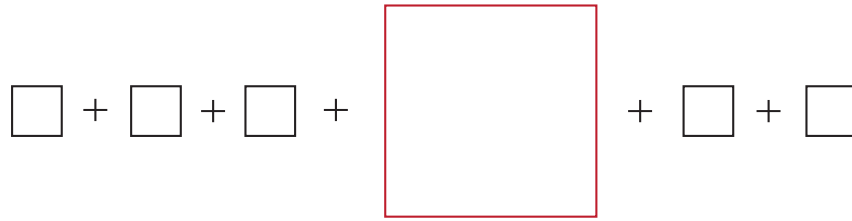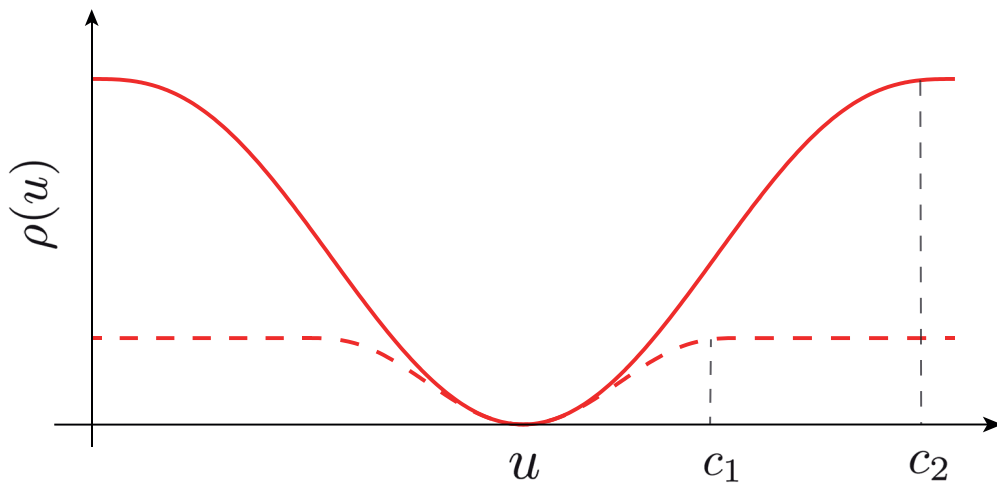
Figure 4.4: Outlier in the LS sum.

Figure 4.5: Two bisquare functions with different clipping parameters.

The idea of the bisquare function is that the residuals produced by the outliers lay in the part of the function that is constant, so they cannot dominate the sum in (4.2). The only question that remains is: how do we find the best value for c?

This depends on how dispersed the residuals produced by the *clean* data are. The dispersion of a set of points can be measured by its *scale* σ (also known as the standard deviation). Then, to avoid having to set a different clipping parameters for every dataset, the residuals are *normalized* by their scale,

$$\widehat{\mathbf{x}}_{\mathrm{M}} = \arg\min_{\mathbf{x}} \sum_{i=1}^{m} \rho\left(\frac{r_i}{\sigma}\right). \tag{4.4}$$

This is the definition of the family of M estimators. By choosing different ρ functions, we get different estimators in the family. LS itself is just a particular member of this family.

Now that we have defined the M estimator, let us go back to the big picture of the overall problem. Given the system model, $\mathbf{A}$, and a set of measurements, $\widehat{\mathbf{y}}$, how do we go about computing $\widehat{\mathbf{x}}_{\mathrm{M}}$ given in (4.4)?

4.2.1 Computation: iterative re-weighted least squares

M estimators, and robust estimators in general, are typically computed using iterative re-weighted least squares algorithms (IRLS) [39]. To compute an M estimate $\widehat{\mathbf{x}}_M$, we need to minimize the function given in (4.4). For simplicity let us assume $\sigma = 1$. Then we follow these steps

1. Take derivatives of the objective function with respect to $\mathbf{x}$, and make them equal to zero

$$\sum_{i=1}^{m} \frac{\partial \rho(r_i)}{\partial x_j} = 0, \text{ for } j = 1, \cdots, n. \tag{4.5}$$

   If we denote the derivative of the function ρ as

$$\psi(u) = \frac{d\rho(u)}{du}, \tag{4.6}$$

   we can write the system of equations as

$$\sum_{i=1}^{m} \psi(r_i) \frac{\partial r_i}{\partial x_j} = 0, \text{ for } j = 1, \cdots, n. \tag{4.7}$$

2. Define the weights $w(u)$

$$w(u) = \frac{\psi(u)}{u}. \tag{4.8}$$

   Using $w(u)$ we can re-write again the system of equations as

$$\sum_{i=1}^{m} w(r_i) r_i \frac{\partial r_i}{\partial x_j} = 0, \text{ for } j = 1, \cdots, n. \tag{4.9}$$

3. Rewriting the derivative of the objective function in this way makes it easy to realize that (4.9) is also the derivative of the function

$$f(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^{m} w(r_i) r_i^2 = \frac{1}{2} \|\mathbf{W}(\mathbf{x})(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x})\|_2^2, \tag{4.10}$$

   where $\mathbf{W}(\mathbf{x})$ is a diagonal matrix that contains the square root of the weights $\sqrt{w(r_i)}$ in its diagonal.

4. If two functions have the same derivative, they will have their minima located at the same points. Thus, in order to compute $\widehat{\mathbf{x}}_M$, instead of finding the location of the minima of (4.4), we can find the location of the minima of (4.10).

   The function given in (4.10) defines an LS problem with some weights $\mathbf{W}(\mathbf{x})$. The difficulty here is that the weights depend on $\mathbf{x}$. Thus, it is necessary to solve this problem in an iterative manner: given an initial solution $\mathbf{x}_0$, in each iteration k an estimate $\widehat{\mathbf{x}}_k$ is computed

$$\widehat{\mathbf{x}}_k = \underset{\mathbf{x}}{\arg\min} \|\mathbf{W}(\widehat{\mathbf{x}}_{k-1})(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x})\|_2^2. \tag{4.11}$$

This algorithm is part of the class of optimization algorithms known as IRLS. The main application of this class is in robust statistics, but there are also other interesting applications [29] like computing the estimate given by sparse regularization (1.17).

We now know what M estimators are, and how to compute them. Since these new robust estimators are extensions to basic estimators such as LS, we need a new vocabulary to describe their additional robustness properties. As we will see in the following section, there are four key robustness properties that describe and quantify the performance of any estimator.

4.3 Metrics of robustness and efficiency

As we said above, robust estimators have two goals: to downgrade the influence of outliers in the dataset, if any, and to produce correct estimates if the dataset is not contaminated. To evaluate these two properties in different estimators we need specific metrics. In this section we give a brief review of the metrics that we will use along this chapter.

Sensitivity curve

The sensitivity curve (SC) of an estimator was introduced by J.W. Tukey [40]. It aims to measure how much an estimator $\widehat{\mathbf{x}}$ can be deviated if a dataset of m measurements $\mathbf{y}$ is contaminated with one outlier y_0. The new contaminated measurement set is thus $\mathbf{y} \cup \{y_0\}$. If we add a new measurements to the dataset, we should also add a new row $\mathbf{a}_0$ to the model matrix $\mathbf{A}$. Thus the SC also considers outliers in the model matrix $\mathbf{A}$. If $\widehat{\mathbf{x}}(\mathbf{M}, \mathbf{v})$ indicates that the estimator uses the model matrix $\mathbf{M}$ and the set of measurements $\mathbf{v}$ in the linear regression, then the SC is defined in this case as

$$\mathrm{SC}(\mathbf{a}_0, y_0) = \frac{\widehat{\mathbf{x}}(\mathbf{A} \cup \{\mathbf{a}_0\}, \mathbf{y} \cup \{y_0\}) - \widehat{\mathbf{x}}(\mathbf{A}, \mathbf{y})}{1/(m+1)}. \tag{4.12}$$

Influence function

The SC considers finite datasets. We can extend its definition to the asymptotic case, i.e. when the number of measurements in the dataset m is infinite, but the number of unknowns in $\mathbf{x}$, n, is still finite. This is known in statistics as the asymptotic regime. In this regime the SC becomes the *Influence Function* (IF), originally proposed by Hampel [37]. It indicates how an estimate changes when there is an infinitesimal proportion of outliers ϵ in the data. To give the exact definition of the IF we need to formulate the problem in the asymptotic regime. We will do it in the next sections.

For some particular estimators it was proven than the SC converges to the IF [44]. However we cannot generalize this assumption.

Breakdown point

The SC and IF explore how the estimator is affected by a small amount of outliers in the data. In contrast, the aim of the breakdown point (BP) is to explore the maximum number of outliers that one estimator can handle. It was also proposed by Hampel in 1968, and Donoho and Huber defined in [32] as

...roughly, the smallest amount of contamination that may cause an estimator to take on arbitrarily large aberrant values (...) The breakdown point provides only a very crude quantification of the robustness properties of an estimator. But just in this lies its generality and strength. In our opinion, the breakdown point is most useful in small sample situations, and it can be defined without recourse to a probability model...

For example, consider a set of samples $z_1, \ldots, z_m$ from a random variable Z. We can estimate its mean μ as

$$\widehat{\mu} = \frac{z_1 + \cdots + z_m}{m}. \tag{4.13}$$

The estimator $\widehat{\mu}$ has a BP equal to zero, because we can make $\widehat{\mu}$ arbitrarily large just by replacing any of the samples. In this case we say that $\widehat{\mu}$ is not robust to outliers.

The highest BP that an estimator can achieve is 0.5. This means that half of the data are outliers. We can not have a higher BP because it does not make sense to have a dataset where more than half of the data are outliers.

Relative efficiency

The aim of the metrics above is to measure the *robustness* of an estimator. We need also tools to measure how good an estimator performs if there are not outliers in a dataset. One way of measuring this is, given a dataset with errors **e** with a particular non-contaminated distribution, to compare the performance the estimator that we want to evaluate with the performance of the *optimal* estimator for that distribution. In this case we define optimal estimator as the maximum likelihood estimator (MLE) $\widehat{\mathbf{x}}_{\mathrm{MLE}}$. Then

$$\mathrm{Eff}(\widehat{\mathbf{x}}) = \frac{\mathbb{E}[(\widehat{\mathbf{x}}_{\mathrm{MLE}} - \mathbf{x})^2]}{\mathbb{E}[(\widehat{\mathbf{x}} - \mathbf{x})^2]}. \tag{4.14}$$

The relative efficiency is defined in the interval $\mathrm{Eff}(\widehat{\mathbf{x}}) \in [0, 1]$. For example, if the errors **e** have a Gaussian distribution, the MLE estimator is LS, so we use LS as the optimal estimator. We would like to design an estimator that has a performance close to LS when the distribution of the errors is Gaussian. If an estimator has an efficiency close to 1, then we say that the estimator is *efficient*.

4.4 The regularized τ estimator

Designing an estimator that simultaneously is robust against outliers and is efficient when there are no outliers in the data is very challenging. Usually, we make an estimator robust by sacrificing its efficiency, or vice-versa. This trade-off exists in both the M estimators and the S estimators [44]. However, we ideally want to have estimators with both properties simultaneously. When working with real-world data, we usually do not know a priori if they contain outliers or not.

The τ estimator [76] meets this challenge: it is simultaneously robust to outliers and efficient w.r.t. the Gaussian distribution. It has a BP of 0.5, and at the same time it performs almost as well as the LS estimator when the errors are Gaussian: its efficiency at the normal distribution is close to one.

To understand how the τ estimator achieves this, let us look at the intuition behind it. The τ estimator can be seen as an *adaptive* M estimator whose ρ function is the weighted sum of two other ρ functions [76]

$$\rho_\tau(u) = w_\tau \rho_1(u) + \rho_2(u). \tag{4.15}$$

The weight w_τ is non-negative, and it depends on the data. We will give later its precise definition, but note that w_τ adapts automatically to the distribution of the data. Imagine that we choose ρ_1 to be a robust loss function (for example the function drawn as a dashed line in Figure 4.5), and ρ_2 to be an efficient one, for example the function drawn as a solid line in Figure 4.5. Then, if there are no outliers, the weights w_τ will be approximately zero and the estimator will be efficient; if there are many outliers, w_τ will be large and the estimator will be robust.

Before we give the definition for the τ estimator, we first need to introduce the concepts of M and τ scales.

The M scale

In this chapter we have seen estimators, like LS or M estimators. Here we introduce the first estimator for the scale: the M scale estimator $\widehat{\sigma}_M(\mathbf{v})$ [44]. This estimator is robust against outliers, and is defined as follows: given a vector $\mathbf{v}$ that contains m points, we implicitly define $\widehat{\sigma}_M(\mathbf{v})$ as

$$\frac{1}{m} \sum_{i=1}^{m} \rho \left(\frac{v_i}{\widehat{\sigma}_M(\mathbf{v})} \right) = b. \tag{4.16}$$

If $f(u)$ is the probability density function of a Gaussian distribution, then b is defined as

$$b = \mathbb{E}[\rho(u)] = \int \rho(u)f(u)du. \tag{4.17}$$

The τ scale

Like the M scale, the τ scale is another robust estimator for the scale of a set of points $\mathbf{v}$, $\widehat{\sigma}_\tau(\mathbf{v})$. It uses two different ρ functions, ρ_1 and ρ_2, and is defined as

$$\widehat{\sigma}_\tau^2(\mathbf{v}) = \widehat{\sigma}_M^2(\mathbf{v}) \frac{1}{m} \sum_{i=1}^{m} \rho_2 \left(\frac{v_i}{\widehat{\sigma}_M(\mathbf{v})} \right). \tag{4.18}$$

In this case $\widehat{\sigma}_M(\mathbf{v})$ uses the ρ_1 function.

The regularized τ estimate for regression

The τ estimate [76] is defined as the solution $\mathbf{x}$ that minimizes the τ scale estimate of the residuals $\mathbf{r} = \widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}$

$$\widehat{\mathbf{x}}_\tau = \arg\min_{\mathbf{x}} \widehat{\sigma}_\tau^2(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}). \tag{4.19}$$

This estimate can be computed using an IRLS algorithm [60] and, as we pointed out before, it can also be seen as an M estimator that automatically adapts to the data. This is why, depending on the dataset, it can be robust (if there are outliers) or efficient (if the dataset is outlier-free).

As we have seen in Chapter 3, the central problem of this thesis - how to estimate emissions into the atmosphere - involves datasets that contain outliers. In addition, model matrices $\mathbf{A}$ in this problem typically have large condition numbers, which makes the problem ill-posed (see Chapter 1). The τ estimator, in spite of all its good properties, cannot deal with ill-posed problems [48]. It is designed to attenuate the effect of outliers, but it does not convert an ill-posed problem into a well-posed one.

We need estimators that, in addition to being robust and efficient simultaneously, can also deal with ill-posed problems. The table given in Figure 4.6 gives a summary of the most important estimators found in the literature, along with their properties. Included are LS, regularized LS (introduced in Chapter 1), regularized M [52], regularized S [68], and the τ estimator. It is important to note that none of them achieves the three properties that we need.

Property \\ Estimator	Regularized	Robust	Efficient
Least Squares	✗	✗	✓
Regularized LS	✓	✗	✓
Regularized M	✓	✓	✗
Regularized S	✓	✓	✗
τ	✗	✓	✓

Figure 4.6: Summary of the properties of different estimators.

To create an estimator with the three properties that we need, we extend the τ estimator by adding a regularization to its objective function. The idea is that the regularization will make the estimator suitable for ill-posed problems, while preserving all its good properties of robustness and efficiency. We propose two versions of the regularized τ estimator. In the first one we use the ℓ_2 of $\mathbf{x}$ as a regularization (see Chapter 1), so the estimator is defined as

$$\widehat{\mathbf{x}}_\tau = \arg\min_{\mathbf{x}} \widehat{\sigma}_\tau^2(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}) + \lambda\|\mathbf{x}\|_2^2. \tag{4.20}$$

The second version uses the ℓ_1 norm of $\mathbf{x}$ as a regularization (again, see Chapter 1), so the new estimator is defined as

$$\widehat{\mathbf{x}}_\tau = \arg\min_{\mathbf{x}} \widehat{\sigma}_\tau^2(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}) + \lambda\|\mathbf{x}\|_1. \tag{4.21}$$

Having defined our new estimators, we proceed to:

1. Design algorithms to compute the new estimates,

2. Analyze the properties of the estimates to understand how the introduction of the regularization affects the robustness and efficiency of the basic τ estimator, and

3. Apply our new estimator to a real-world dataset.

4.5   Computing the regularized τ estimator

To compute the regularized τ estimate, we need to minimize the objective functions given in (4.20) and (4.21). These functions are non-convex, like the example represented in Figure 4.7. We can find a minimum, but we do not know if that minimum is global unless we can compare it with the rest of minima of the function. That is why most non-convex problems are hard (if not impossible) to solve exactly in a reasonable time [22]. Hence, we propose instead heuristic algorithms that compute an approximation of this global minimum in a reasonable amount of time, althought they may or may not find it.
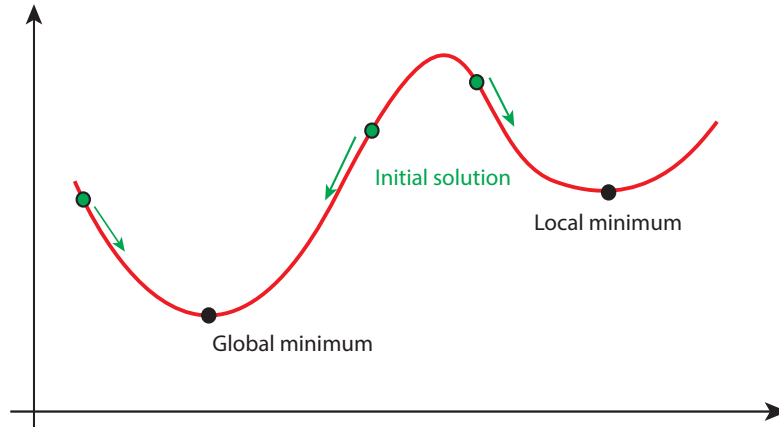
Figure 4.7: Example of a non convex function: the local minima that we find with the IRLS algorithm depend on the initial solution that we use.

We will explain the algorithms in two parts: how to find local minima, and how to approximate the global minimum.

4.5.1   Finding local minima

As we said above, the non-regularized τ estimator, as the M and many others robust estimators, can be computed using an IRLS algorithm [60]. Thus, our first idea to compute our estimates is to use also an IRLS algorithm.

Iterative Re-weighted Least Squares algorithm

To derive this algorithm we use the same procedure that we followed to obtain the IRLS algorithm to compute the M estimate in Section 4.2.1: first, we find the derivative w.r.t. $\mathbf{x}$ of the objective function, and then, by reordering and redefining some terms we identify the weighted least squares function. The details about this derivation are given in Appendix 4.8.1.

We found that the derivative of (4.20) is also the derivative of the regularized weigthed least squares function

$$\widehat{\mathbf{x}}_{\tau} = \arg\min_{\mathbf{x}} \|\mathbf{Z}^{1/2}(\mathbf{x})(\mathbf{A}\mathbf{x} - \widehat{\mathbf{y}})\|_2^2 + \lambda\|\mathbf{x}\|_2^2. \tag{4.22}$$

To define the diagonal matrix of weights $\mathbf{Z}(\mathbf{x}) \in \mathbb{R}^{m \times n}$ we introduce first some notation. Notice that the goal of these new definitions is just to have a more compact notation, so the IRLS objective function can be written in a simpler manner:

1. **The *normalized* residuals:** These are the residuals that are normalized by their scale. The scale is given in this case by the M scale estimator $\widehat{\sigma}_{\mathrm{M}}(\mathbf{v})$:

$$\tilde{\mathbf{r}}(\mathbf{x}) := \frac{\mathbf{r}(\mathbf{x})}{\widehat{\sigma}_{\mathrm{M}}(\mathbf{r}(\mathbf{x}))} = \frac{\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}}{\widehat{\sigma}_{\mathrm{M}}(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x})}. \qquad (4.23)$$

2. **The score function:** This function, denoted by $\psi(u)$, is the derivative of $\rho(u)$

$$\psi(u) := \frac{d\rho(u)}{du}. \qquad (4.24)$$

3. **Score function for the τ estimator:** As we introduced in (4.15) the τ estimator has a special ρ function

$$\rho_\tau(u, \mathbf{v}) = w_\tau(\mathbf{v})\rho_1(u) + \rho_2(u). \qquad (4.25)$$

Here we make explicit that the weight $w_\tau(\mathbf{v})$ depends on the data. It is defined as

$$w_\tau(\mathbf{v}) = \frac{\sum_{i=1}^{m} -2\rho_2(v_i) + \psi_2(v_i) v_i}{\sum_{i=1}^{m} \psi_1(v_i) v_i}. \qquad (4.26)$$

The score function $\psi_\tau(u)$ for this case is defined as

$$\psi_\tau(u, \mathbf{v}) = w_\tau(\mathbf{v})\psi_1(u) + \psi_2(u). \qquad (4.27)$$

Using the three concepts defined above we can write the weights contained in $\mathbf{Z}(\mathbf{x})$ as

$$z_i(\mathbf{x}) = \begin{cases} \frac{\psi_\tau(\tilde{r}_i(\mathbf{x}), \tilde{\mathbf{r}}(\mathbf{x}))}{2\tilde{r}_i(\mathbf{x})} & \text{if } \tilde{r}_i(\mathbf{x}) \neq 0 \\ 1 & \text{otherwise.} \end{cases} \qquad (4.28)$$

Thus when the τ estimate is regularized by the ℓ_2 norm it can be computed using an IRLS algorithm. More details are provided in Algorithm 4.1, and in the software implementation.

Algorithm 4.1 Regularized IRLS

INPUT: $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, λ, ξ, $\mathbf{x}_0$, K
OUTPUT: $\widehat{\mathbf{x}}$
constrain $\lambda \geq 0$, $\xi \geq 0$, $K \geq 0$
for $k = 0$ to $K - 1$ **do**
    $\mathbf{r}_k \leftarrow \mathbf{y} - \mathbf{A}\mathbf{x}_k$
    Compute $\widehat{\sigma}_M(\mathbf{r}_k)$
    Compute $\mathbf{Z}(\mathbf{x}_k)$
    $\mathbf{x}_{k+1} \leftarrow (\mathbf{A}^\top \mathbf{Z}(\mathbf{x}_k)\mathbf{A} + \lambda^2)^{-1}\mathbf{A}^\top \mathbf{Z}(\mathbf{x}_k)\mathbf{y}$
    if $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < \xi$ **then**
        break
    end if
end for
return $\widehat{\mathbf{x}} \leftarrow \mathbf{x}_{k+1}$

For the case when the τ estimator is regularized with the ℓ_1 norm (4.21), the derivation of the IRLS algorithm is not so straight-forward: the ℓ_1 norm of a vector is not differentiable. Nevertheless, by approximating $\|\mathbf{x}\|_1$ with a differentiable function, we show in Appendix 4.8.1 that a local minimum in (4.21) is indeed also a local minimum in

$$\widehat{\mathbf{x}}_\tau = \arg\min_{\mathbf{x}} \|\mathbf{Z}^{1/2}(\mathbf{x})(\mathbf{A}\mathbf{x} - \widehat{\mathbf{y}})\|_2^2 + \lambda\|\mathbf{x}\|_1, \tag{4.29}$$

where $\mathbf{Z}(\mathbf{x})$ is defined as in (4.22). Hence we can use a regularized IRLS algorithm to find the local minima for the regularized τ estimator that uses the ℓ_1 norm.

These IRLS algorithms are quite efficient and we observed that they converge in a few iterations. However, there are not theoretical guarantees that assure their converge to a local minimum. In fact these guarantees do not exist either for the case of the non-regularized estimator [60].

Accelerate proximal gradient algorithm

The lack of theoretical guarantees for convergence in the IRLS algoritms moved us to look for new algorithms with guarantees. Thus, changing completely our approach we now look at the class of *proximal algorithms*.

Proximal algorithms is a class of algorithms for solving optimization problems. As it is put in [53]

...much like Newtons method is a standard tool for solving unconstrained smooth minimization problems of modest size, proximal algorithms can be viewed as an analogous tool for nonsmooth, constrained, large-scale, or distributed versions of these problems.

This family of algorithms owe its name to the *proximal operator*. Given a function $h : \mathbb{R}^n \to \mathbb{R}$, and the scalar $\alpha > 0$, the proximal operator is defined as [53]

$$\mathbf{prox}_{\alpha h}(\mathbf{v}) = \arg\min_{\mathbf{x}}(h(\mathbf{x}) + \frac{1}{2\alpha}\|\mathbf{x} - \mathbf{v}\|_2^2). \tag{4.30}$$

This operator can be interpreted as follows [53]:

The definition indicates that $\mathbf{prox}_{\alpha h}(\mathbf{v})$ is a point that compromises between minimizing h and being near to $\mathbf{v}$. For this reason, $\mathbf{prox}_{\alpha h}(\mathbf{v})$ is sometimes called a proximal point of $\mathbf{v}$ with respect to h. In $\mathbf{prox}_{\alpha h}(\mathbf{v})$, the parameter α can be interpreted as a relative weight or trade-off parameter between these terms.

One of the simplest algorithms that uses this operator is the *proximal gradient algorithm* (PG). Consider the problem

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} f(\mathbf{x}) + g(\mathbf{x}), \tag{4.31}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are convex functions and f is differentiable. Then the proximal gradient method is given by

$$\mathbf{x}_{k+1} := \mathbf{prox}_{\alpha_k g}(\mathbf{x}_k - \alpha_k \nabla f(\mathbf{x}_k)). \tag{4.32}$$

There are several interpretations of this algorithm. One way of seeing the intuition behind is realizing that, at each step, the PG algorithm makes a trade-off between the point $\widehat{\mathbf{x}}_{\text{GD}}$ given by the gradient descent algorithm applied to f,

$$\widehat{\mathbf{x}}_{\text{GD}} = \mathbf{x}_k - \alpha_k \nabla f(\mathbf{x}_k), \tag{4.33}$$

and the minimum of $g(\mathbf{x})$. The step α_k controls the trade-off. This algorithm has guarantees of convergence if the two functions f and g are convex. This is not the case for the regularized τ estimator objective function. However, there are extensions of the proximal gradient algorithms that ensure convergence in a non-convex setting. We will focus on the idea proposed in [43].

Before explaining the algorithms given in [43], we introduce an extension of the proximal gradient algorithm, that will help us to understand [43]. This extension is called *accelerated proximal gradient algorithms* (APG). These algorithms add an extrapolation step to the proximal gradient algorithms

$$\begin{aligned} \mathbf{z}_{k+1} &:= \mathbf{x}_k + w_k(\mathbf{x}_k - \mathbf{x}_{k-1}) & (4.34) \\ \mathbf{x}_{k+1} &:= \mathbf{prox}_{\alpha_k g}(\mathbf{z}_{k+1} - \alpha \nabla f(\mathbf{z}_{k+1})), & (4.35) \end{aligned}$$

where $w_k \in [0,1)$ is an extrapolation parameter. The extrapolation step was introduced by Nesterov [50] . He proved that his method has a worst-case convergence rate that is superior to the standard method and that it cannot be improved further. This method, and variations of it, are used in all kinds of optimization algorithms.

After all the concepts explained above we have a suitable setup to introduce the algorithms proposed in [43]. These algorithms are called accelerated proximal gradient methods for non-convex programming. The framework for these algorithms is similar to the framework used by the APG algorithms

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} F(\mathbf{x}) = \arg\min_{\mathbf{x}} f(\mathbf{x}) + g(\mathbf{x}), \tag{4.36}$$

but in this case f is a differentiable, and maybe non-convex function, and g can be non-convex and non-smooth. Notice that the objective functions of the regularized τ estimators (4.20) and (4.21) fit in this framework if we define

$$\begin{aligned} f(\mathbf{x}) &:= \hat{\sigma}_\tau^2(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}) \\ g(\mathbf{x}) &:= \lambda\|\mathbf{x}\|_2^2 \text{ or } \lambda\|\mathbf{x}\|_1. \end{aligned} \tag{4.37}$$

In particular we focus on one of the methods proposed in [43], the monotone accelerate proximal gradient algorithm. It is defined as follows

$$\begin{aligned} \mathbf{l}_k &:= \mathbf{x}_k + \frac{t_{k-1}}{t_k}(\mathbf{z}_k - \mathbf{x}_k) + \frac{t_{k-1}-1}{t_k}(\mathbf{x}_k - \mathbf{x}_{k-1}) & (4.38) \\ \mathbf{z}_{k+1} &:= \mathbf{prox}_{\alpha_l g}(\mathbf{l}_k - \alpha_l \nabla f(\mathbf{l}_k)) & (4.39) \\ \mathbf{v}_{k+1} &:= \mathbf{prox}_{\alpha_x g}(\mathbf{x}_k - \alpha_x \nabla f(\mathbf{x}_k)) & (4.40) \\ t_{k+1} &:= \frac{\sqrt{4(t_k)^2+1}+1}{2} & (4.41) \\ \mathbf{x}_{k+1} &:= \begin{cases} \mathbf{z}_{k+1}, \text{ if } F(\mathbf{z}_{k+1}) \leq F(\mathbf{v}_{k+1}) \\ \mathbf{v}_{k+1}, \text{ otherwise} \end{cases} & (4.42) \end{aligned}$$

where $\alpha_l \in (0, 1/L]$ and $\alpha_k \in (0, 1/L]$ are the step sizes of the algorithm, and L is the Lipschitz constant of $\nabla f(\mathbf{x})$. This constant in our case is unknown, so α_l and α_k are found in each iteration using the linear search algorithm proposed in [14]. Let us take a look at the steps proposed in (4.38) – (4.42):

1. The step (4.38) is a variation of Nesterov's extrapolation. This steps gives the *accelerate* to the name of the algorithm.

2. Steps (4.39) and (4.40) are identical to the proximate gradient algorithm (given in (4.32)). It is applied at the point $\mathbf{l}_k$ that we obtained with the extrapolation, and also at $\mathbf{x}_k$. This is done because, ought to the function f being non-convex, there are not guarantees of $\mathbf{l}_k$ being a *good* extrapolation, i.e. the extrapolation may not go in the right direction.

3. The step (4.41) updates the extrapolation parameter for the next iteration.

4. $\mathbf{v}_{k+1}$ and $\mathbf{z}_{k+1}$ are our two candidates. In (4.42) we simply choose the one that minimizes the objective function.

As we said above, this algorithm can be used to minimize the objective function of the regularized τ estimator. Besides, it has theoretical guarantees that it will converge to a local minima (more details in [43]), which is something we could not prove in the case of the IRLS algorithm.

Algorithm 4.2 Accelerated proximal gradient (APG) algorithm

INPUT: $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, λ, ξ, $\mathbf{x}_0$, K
OUTPUT: $\widehat{\mathbf{x}}$
constrain $\lambda \geq 0$, $\xi \geq 0$, $K \geq 0$
$t_0 = 1$
for $k = 0$ to K **do**
    $\mathbf{l}_k = \mathbf{x}_k + \frac{t_{k-1}}{t_k}(\mathbf{z}_k - \mathbf{x}_k) + \frac{t_{k-1}-1}{t_k}(\mathbf{x}_k - \mathbf{x}_{k-1})$
    $\mathbf{z}_{k+1} = \text{prox}_{\alpha_l g}(\mathbf{l}_k - \alpha_l \nabla f(\mathbf{l}_k))$
    $\mathbf{v}_{k+1} = \text{prox}_{\alpha_x g}(\mathbf{x}_k - \alpha_x \nabla f(\mathbf{x}_k))$
    $t_{k+1} = \frac{\sqrt{4(t_k)^2 + 1} + 1}{2}$
    if $F(\mathbf{z}_{k+1}) \leq F(\mathbf{v}_{k+1})$ **then** $\mathbf{x}_{k+1} \leftarrow \mathbf{z}_{k+1}$
    else
        $\mathbf{x}_{k+1} \leftarrow \mathbf{v}_{k+1}$
    end if
    if $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < \xi$ **then**
        break
    end if
end for

Performance comparison

To compare the performance of the IRLS and APG algorithms that we propose, we run numerical experiments. In each of these experiments synthetic datasets are generated such that $\mathbf{y}_k = \mathbf{A}_k \mathbf{x}_k + \mathbf{e}_k$ where $\mathbf{x}_k \in \mathbb{R}^{n \times 1}, \mathbf{y}_k \in \mathbb{R}^{m \times 1}, \mathbf{A}_k \in \mathbb{R}^{m \times}$ and $\mathbf{e}_k \in \mathbb{R}^{m \times 1}$ are random vectors sampled from a Gaussian distribution with zero mean and variance 1. The dimensions of the problem are fixed to $m = 10$ and $n = 3$.

The experiment is run 1000 times. In each run, $\mathbf{x}_k$ is estimated using the IRLS algorithm and it is also estimate using the APG algorithm. To measure the difference between the two estimates we use

$$d = \frac{\|\widehat{\mathbf{x}}_{\text{APG}} - \widehat{\mathbf{x}}_{\text{IRLS}}\|}{\|\mathbf{x}_k\|}. \tag{4.43}$$

In the case of the τ estimator regularized with the ℓ_2 norm the maximum d that we obtained is $4.2e^{-4}$, and in the case of using the ℓ_1 regularization the maximum d is $5.2e^{-4}$. Given that these distances are normalized by the ground-truth, they can be considered negligible.

We also observed that the regularized IRLS algorithm has a faster convergence than the APG algorithm. The IRLS algorithm converges in 1 or 2 iterations, whereas the APG algorithm needs around 100 iterations.

4.5.2 Approximating the global minimum

As the objective function (4.19) is non-convex, each local minimum that we find using the IRLS and APG algorithms depends on the initial solution that we use (see Figure 4.7). This takes us to the second step of the algorithms: to approximate the global minimum. For that, we run the algorithms Q times using for each run a different random initial solution. Then, we compare the different minima that we found and pick the best one. See Algorithm 4.3 for more details.

Algorithm 4.3 Finding global minimum

INPUT: $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, λ, ξ, K, Q
OUTPUT: $\widehat{\mathbf{x}}$
constrain $\lambda \geq 0$, $\xi \geq 0$, $K \geq 0$, $Q \geq 0$
for $q = 0$ to $Q - 1$ **do**
    $\mathbf{x}_q \leftarrow$ random initial condition
    $\widehat{\mathbf{x}}_q \leftarrow$ Algorithm 4.1 $(\mathbf{y}, \mathbf{A}, \lambda, \xi, \mathbf{x}_q, K)$ or Algorithm 4.2 $(\mathbf{y}, \mathbf{A}, \lambda, \xi, \mathbf{x}_q, K)$
end for
return $\widehat{\mathbf{x}} \leftarrow \underset{\mathbf{x} \in \{\widehat{\mathbf{x}}_q\}}{\arg\min} \; \widehat{\sigma}_\tau^2(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}) + \lambda \|\mathbf{x}\|_2^2$.

4.6 Properties of the regularized τ estimator

In this section we study the properties of the regularized τ estimator. In particular, we focus on two aspects: the robustness of the estimator against outliers and the performance of the estimator at the nominal distribution, i.e. its efficiency. For the first task, we compute the SC of the estimator, we derive its IF, and we explore its breakdown point. For the second task, we study the asymptotic variance (ASV) and the bias of the estimator when the errors have a Gaussian distribution.

4.6.1   Sensitivity curve

Remember that the $\mathrm{SC}(\mathbf{a}_0, y_0)$ of the τ estimate $\widehat{\mathbf{x}}_\tau$ is given by

$$\mathrm{SC}(\mathbf{a}_0, y_0) = \frac{\widehat{\mathbf{x}}_\tau(\mathbf{A} \cup \{\mathbf{a}_0\}, \mathbf{y} \cup \{y_0\}) - \widehat{\mathbf{x}}_\tau(\mathbf{A}, \mathbf{y}))}{1/(m+1)}. \tag{4.44}$$

It measures the effect of one outlier in the estimate. To be able to visualize the SC we compute it for a one-dimensional dataset, i.e. where x is a scalar, and $\widehat{\mathbf{y}}, \mathbf{a}, \mathbf{e} \in \mathbb{R}^{1000 \times 1}$

$$\widehat{\mathbf{y}} = \mathbf{a}x + \mathbf{e}. \tag{4.45}$$

To generate a non-contaminated dataset, x is fixed to 1.5, and $\widehat{\mathbf{y}}$ and $\mathbf{a}$ are sampled from a Gaussian distribution with zero mean and variance 1. The added outliers a_0 and y_0 take values in the range $[-10, 10]$.

With this setup we compute, using IRLS algorithms, the SC for three different estimators:

1. **Non-regularized τ estimator** as in (4.20) with $\lambda = 0$. This SC is represented in the lower part of Figure 4.9

2. **τ estimator regularized with the ℓ_2 norm** as in (4.20) with $\lambda = 0.1$. This SC is represented in the lower part of Figure 4.10

3. **τ estimator regularized with the ℓ_1 norm** as in (4.21) with $\lambda = 0.1$. This SC is represented in the lower part of Figure 4.11

Observing the figures we realize that the three SC have a very similar shape. This suggests that the introduction of the regularization does not change the sensitivity of the estimate to one outlier.

If we look at the SC from the top, we realize that all the non-zero values are located in a line with a slope of 1.5, which is the value of x. This means that the outliers a_0, y_0 affect the estimator only if they have a proportion close to $y_0 = a_0 x$.

4.6.2   The Influence Function

As we said in Section 4.3, the IF considers the influence in the estimate of an infinitesimal contamination in the data when the number of measurements tends to infinity and the number of unknowns in $\mathbf{x}$ remains fixed and finite. This setup is usually known as *asymptotic* regime.

In this asymptotic regime where $m \to \infty$, usually sums converge to expectations, so the estimators can be redefined. For example the M-scale estimate $\hat{\sigma}_M(u)$ becomes [44]

$$\mathbb{E}_H \left[\rho_1 \left(\frac{u}{\widehat{\sigma}_{\mathrm{M}}(u)} \right) \right] = b. \tag{4.46}$$

The sub-index H in $\mathbb{E}_H[g(u)]$ indicates the distribution that we assume for the random variable u. If $f_H(u)$ is the p.d.f of the distribution H then

$$\mathbb{E}_H[g(u)] = \int g(u) f_H(u) du. \tag{4.47}$$

It is also shown in [76] that in the asymptotic regime the τ estimate becomes

$$\widehat{\mathbf{x}}_\tau = \arg\min_{\mathbf{x}} \widehat{\sigma}_{\mathrm{M}}^2(r(\mathbf{x})) \mathbb{E}_H \left[\rho_2 \left(\frac{r(\mathbf{x})}{\widehat{\sigma}_{\mathrm{M}}(r(\mathbf{x}))} \right) \right] + \lambda \|\mathbf{x}\|_2^2, \tag{4.48}$$

where $r(\mathbf{x}) = y - \mathbf{a}\mathbf{x}$.

To derive the IF, apart of redefining the estimators for the new regime, we need also to assume a distribution for our data. Let us model y as

$$y = \mathbf{a}\mathbf{x}_0 + e, \tag{4.49}$$

where $\mathbf{a} \in \mathbb{R}^{1 \times n}$ is a row vector of i.i.d. random entries independent of e, and $\mathbf{x}_0 \in \mathbb{R}^{n \times 1}$ is the deterministic vector of unknown true parameters. We assume $\mathbf{a}$ and y to have a joint distribution H. When the measurements and/or the model contain outliers, H takes the form

$$H_\epsilon = (1 - \epsilon)H_0 + \epsilon G, \tag{4.50}$$

where H_0 is the distribution of the *clean* data, G is *any* distribution different from H_0, and $\epsilon \in [0, 1]$ is the proportion of outliers in the data.

Now we have all the ingredients to talk about the definition of IF:

Definition 4.1

The IF of an estimate $\widehat{\mathbf{x}}$ is defined as the derivative of the estimate $\widehat{\mathbf{x}}$ w.r.t. the proportion of the contamination in the data ϵ assuming that

1. *The data $\mathbf{a}$ and y has a joint distribution H_ϵ such that*

$$H_\epsilon = (1 - \epsilon)H_0 + \epsilon G, \tag{4.51}$$

   where H_0 is any nominal distribution and $G = \delta_{\mathbf{a}_0, y_0}$ is the point mass at $(\mathbf{a}_0, y_0)$.

2. *The IF studies the influence of an infinitesimal contamination, so we assume $\epsilon \to 0$.*

Thus we can write the IF as

$$\mathrm{IF}(\mathbf{a}_0, y_0) = \left. \frac{\partial \widehat{\mathbf{x}}(H_\epsilon)}{\partial \epsilon} \right|_{\epsilon = 0}. \tag{4.52}$$

In this case the notation $\widehat{\mathbf{x}}(H_\epsilon)$ indicates that the underlying distribution of the data is H_ϵ.

Influence function for the τ estimator regularized by a twice-differentiable function

In Appendix 4.8.2 we derive the IF for the τ estimator when it is regularized with any twice differentiable function $J(\mathbf{x})$ such that

$$\widehat{\mathbf{x}}_\tau = \arg\min_{\mathbf{x}} \widehat{\sigma}_\tau^2(\widehat{\mathbf{y}} - \mathbf{A}\mathbf{x}) + \lambda J(\mathbf{x}). \tag{4.53}$$

Notice that the ℓ_2 norm is a particular case of $J(\mathbf{x})$. The result is given in the following theorem.

Theorem 4.1

Influence function for the τ estimator regularized with a twice differentiable function

Let $y = \mathbf{a}\mathbf{x}_0 + e$ be as given in (4.49), and $\widehat{\mathbf{x}}_\tau$ as given in (4.53). Let also $\rho_1, \rho_2, J(\mathbf{x})$ be twice differentiable functions. Assume the M scale estimate $\widehat{\sigma}_M(u)$ as given in (4.46). Define

$$
\begin{aligned}
r(\mathbf{v}) &= y - \mathbf{a}\mathbf{v}, \\
\tilde{r}(\mathbf{v}) &= \frac{r(\mathbf{v})}{\widehat{\sigma}_M(\mathbf{v})}, \\
\psi_k(u) &= \frac{d\rho_k(u)}{du}, \\
J'(\mathbf{v}) &= \left[\frac{\partial J(\mathbf{v})}{\partial v_1}, \ldots, \frac{\partial J(\mathbf{v})}{\partial v_n}\right]^T, \\
J''(\mathbf{v}) &= \left[\frac{\partial J'(\mathbf{v})}{\partial v_1}, \ldots, \frac{\partial J'(\mathbf{v})}{\partial v_n}\right]^T, \\
w(\mathbf{v}) &= \frac{\mathbb{E}_{H_\epsilon}[2\rho_2(\mathbf{v}) - \psi_2(\mathbf{v})\mathbf{v}]}{\mathbb{E}_{H_\epsilon}[\psi_1(\mathbf{v})\mathbf{v}]}, \\
\psi_\tau(\mathbf{v}) &= w(\mathbf{v})\psi_1(\mathbf{v}) + \psi_2(\mathbf{v}), \\
c(\mathbf{v}) &= \frac{\partial(\psi_\tau(\tilde{r}(\mathbf{v}))\widehat{\sigma}_M(r(\mathbf{v})))}{\partial \mathbf{v}}.
\end{aligned}
$$

With this, define

$$
\begin{aligned}
B &= -\mathbb{E}_{H_0}\left[\mathbf{a}^\top c(\widehat{\mathbf{x}}_\tau)\right] + \lambda J''(\widehat{\mathbf{x}}_\tau), \\
C &= \psi_\tau\left(\frac{y_0 - \mathbf{a}_0\widehat{\mathbf{x}}_\tau}{\widehat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))}\right)\widehat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))\mathbf{a}_0^\top, \\
D &= \mathbb{E}_{H_0}\left[\left(\psi_\tau(\tilde{r}(\widehat{\mathbf{x}}_\tau))\right)\widehat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))\mathbf{a}^\top\right].
\end{aligned}
$$

Then the influence function of the τ estimator regularized with $J(\mathbf{x})$ as in (4.53) is given by

$$
IF(\mathbf{a}_0, y_0) = B^{-1}(C - D). \tag{4.54}
$$

The detailed proof of the theorem is in Appendix 4.8.2.

Influence function for the τ estimator regularized by the ℓ_1 norm

The IF given in Theorem 4.1 is not valid for a non-differentiable regularization function like the ℓ_1 norm. To find the IF for this case we approximate the ℓ_1 function

$$
\|\mathbf{v}\|_1 = \sum_{i=1}^m |v_i| \tag{4.55}
$$

with a twice differentiable function $J_k(\mathbf{x})$ such that

$$
\lim_{k \to \infty} J_k(\mathbf{v}) = \|\mathbf{v}\|_1. \tag{4.56}
$$

In particular, we choose

$$J_k(\mathbf{v}) = \sum_{i=1}^{m} v_i \tanh(kv_i). \tag{4.57}$$

As J_k is twice differentiable, the IF of a τ estimator regularized with J_k, $\mathrm{IF}_k(\mathbf{a}_0, y_0)$, is given by Theorem 4.1. The limit of this IF_k when $k \to \infty$ is the IF for the τ estimator regularized with the ℓ_1 norm, IF when $k \to \infty$

$$\lim_{k \to \infty} \mathrm{IF}_k(\mathbf{a}_0, y_0) = \mathrm{IF}(\mathbf{a}_0, y_0). \tag{4.58}$$

This idea leads us to the IF of the τ estimator regularized with the ℓ_1 norm, which is given in the next theorem. The proof for this theorem is given in Appendix 4.8.3.

Theorem 4.2

Influence function for the τ estimator regularized with the ℓ_1 norm

Let $y = \mathbf{a}\mathbf{x}_0 + e$ be as given in (4.49). Let also ρ_1, ρ_2 be twice differentiable functions. Assume the M-scale $\hat{\sigma}_M(u)$ as given in (4.46). Assume $\widehat{\mathbf{x}}_\tau$ as in (4.21).

Without loss of generality, assume $\widehat{\mathbf{x}}_\tau$ has $t \leq n$ non-zero components such that

$$\widehat{\mathbf{x}}_\tau = [x_1, \ldots, x_t, 0, \ldots, 0]. \tag{4.59}$$

Define

$$
\begin{aligned}
r(\mathbf{v}) &= y - \mathbf{a}\mathbf{v}, \\
\tilde{r}(\mathbf{v}) &= \frac{r(\mathbf{v})}{\widehat{\sigma}_M(\mathbf{v})}, \\
\psi_k(u) &= \frac{d\rho_k(u)}{du}, \\
w(\mathbf{v}) &= \frac{\mathbb{E}_{H_\epsilon}[2\rho_2(\mathbf{v}) - \psi_2(\mathbf{v})\mathbf{v}]}{\mathbb{E}_{H_\epsilon}[\psi_1(\mathbf{v})\mathbf{v}]}, \\
\psi_\tau(\mathbf{v}) &= w(\mathbf{v})\psi_1(\mathbf{v}) + \psi_2(\mathbf{v}), \\
c(\mathbf{v}) &= \frac{\partial(\psi_\tau(\tilde{r}(\mathbf{v}))\widehat{\sigma}_M(r(\mathbf{v})))}{\partial\mathbf{v}}.
\end{aligned}
$$

The notation $\mathbf{v}_{1:t}$ indicates that

$$\mathbf{v}_{1:t} = [v_1, \ldots, v_t], \tag{4.60}$$

where $v_1, \ldots, v_t$ are the first t elements in $\mathbf{v}$.

Now define

$$
\begin{aligned}
B &= -\mathbb{E}_{H_0}[\mathbf{a}_{1:t}^\top c(\widehat{\mathbf{x}}_\tau)_{1:t}] \\
C &= \psi_\tau\left(\frac{y_0 - \mathbf{a}_0\widehat{\mathbf{x}}_\tau}{\hat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))}\right)_{1:t} \hat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))(\mathbf{a}_0^\top)_{1:t} \\
D &= \mathbb{E}_{H_0}\left[(\psi_\tau(\tilde{r}(\widehat{\mathbf{x}}_\tau)))_{1:t} \hat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))\mathbf{a}_{1:t}^\top\right]
\end{aligned}
$$

Then the IF for $\widehat{\mathbf{x}}_\tau$ is given by

$$IF(\mathbf{a}_0, y_0) = \begin{bmatrix} B^{-1}(C - D) \\ \mathbf{0}_{n-t} \end{bmatrix}, \tag{4.61}$$

where $\mathbf{0}_{n-t}$ indicates that the last $n - t$ rows of the matrix are filled with zeros.

For more details about the proof, see Appendix 4.8.3. Surprisingly, in this case the IF does not depend on the regularization parameter λ explicitly. Its dependency is implicit in $\widehat{\mathbf{x}}_\tau$.

The IFs given by Theorems 1 and 2 are not bounded in a_0, but they are bounded in y_0 if ψ_τ is bounded (i.e., if ψ_1 and ψ_2 are bounded). This also holds for the original non-regularized τ estimator. We assert then that the regularization does not change the robustness properties of the τ estimator, as measured by the IF.

Plotting the influence function

In the high-dimensional case, the IF cannot be plotted. So, to visualize the results that we obtained in Theorems 4.1 and 4.2, we create 1-dimensional examples. We model a and the additive errors e as Gaussian random variables with zero mean and variance equal to one. The ground truth x_0 is fixed to 1.5. We use ρ_1 and ρ_2 from the optimal family [60], with clipping parameters $c_1 = 1.21$ and $c_2 = 3.27$ (see Figure 4.8). The regularization parameter λ is set to 0.1.
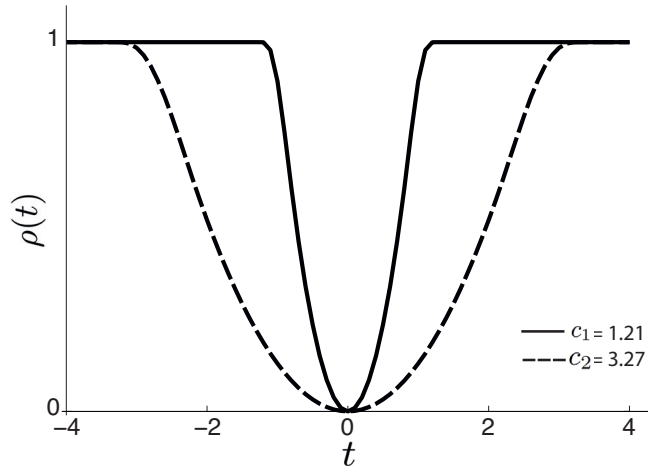
Figure 4.8: Example of two ρ-functions that can be used in the τ estimator. They belong to the optimal family [60], with clipping parameters $c_1 = 1.21$ and $c_2 = 3.27$.

In this setup, we compute the IF for the non-regularized, the ℓ_2-regularized, and the ℓ_1-regularized τ estimators. Figures 4.9, 4.10, and 4.11 show the resulting IFs. We can observe that in all cases the IF is bounded both in a_0 and y_0, so we can consider that in this example the estimators are robust against small fractions of contaminations in a and in e. Also, we notice that the amplitude of the IF is similar for the three different cases. This tells us that the introduction of the regularization does not affect the sensitivity of the estimate to small contaminations of the data. The IF and SC are very similar in all the cases. This may indicate that in these cases the SC converges to the IF.

4.6.3 Breakdown Point

So far we have studied the robustness of the regularized τ estimators when the proportion of outliers in the data is very small. Also, we have explored the behaviour of the estimators in low dimensional problems. The goal now is to investigate how the estimators empirically behave when there are more outliers in the data (up to 40 %) and when the dimension of the problem is higher. This will give us an estimate of the breakdown point (introduced in Chapter 1) of the different estimators.

One reasonable requirement for our estimators is to have simultaneously a small bias and a

Non-regularized tau estimator
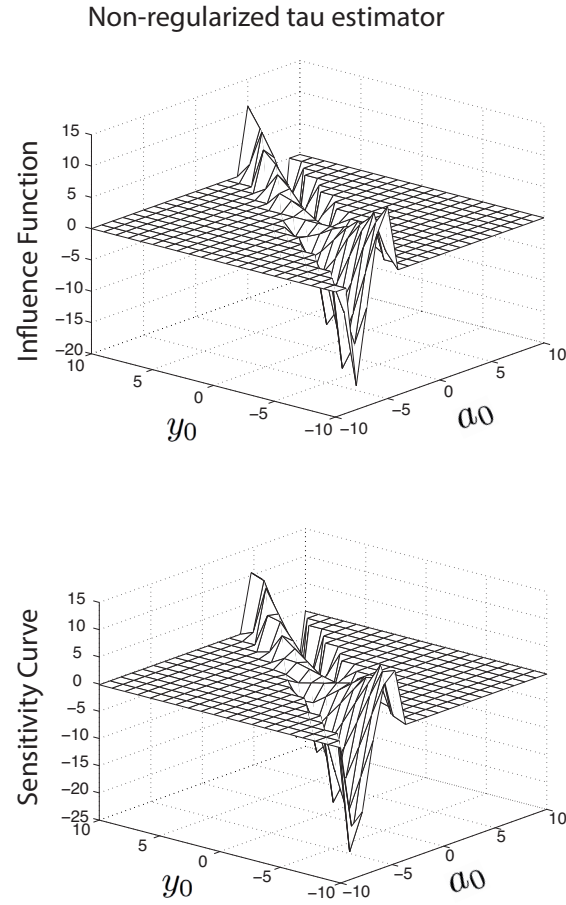
Figure 4.9: IF and SC of the non-regularized τ estimator for a 1-dimensional example.

small variance,

$$
\begin{aligned}
\text{Bias}(\widehat{\mathbf{x}}) &= \mathbb{E}_H[\widehat{\mathbf{x}}] - \mathbf{x}_0, & (4.62) \\
\text{Var}(\widehat{\mathbf{x}}, H) &= \mathbb{E}_H[(\widehat{\mathbf{x}} - \mathbb{E}_H[\widehat{\mathbf{x}}])(\widehat{\mathbf{x}} - \mathbb{E}_H[\widehat{\mathbf{x}}])^\top]. & (4.63)
\end{aligned}
$$

This can be measured by the mean square error. If $\text{Tr}[\mathbf{M}]$ indicates the trace of the matrix $\mathbf{M}$ then

$$
\text{MSE}(\widehat{\mathbf{x}}, H) \;=\; \mathbb{E}[\|\widehat{\mathbf{x}} - \mathbf{x}_0\|_2^2] = \text{Tr}\,[\text{Var}(\widehat{\mathbf{x}})] + \text{Bias}(\widehat{\mathbf{x}})^2. \tag{4.64}
$$

When the number of measurements in our dataset is finite, the MSE is defined as

$$
\text{MSE}(\widehat{\mathbf{x}}) = \frac{1}{m} \sum_{i=1}^{m} (\widehat{\mathbf{x}} - \mathbf{x}_0)^2. \tag{4.65}
$$

We use the MSE to summarize and compare the performance of different estimators, and to get an idea of their BP.

ℓ_2-regularized tau estimator
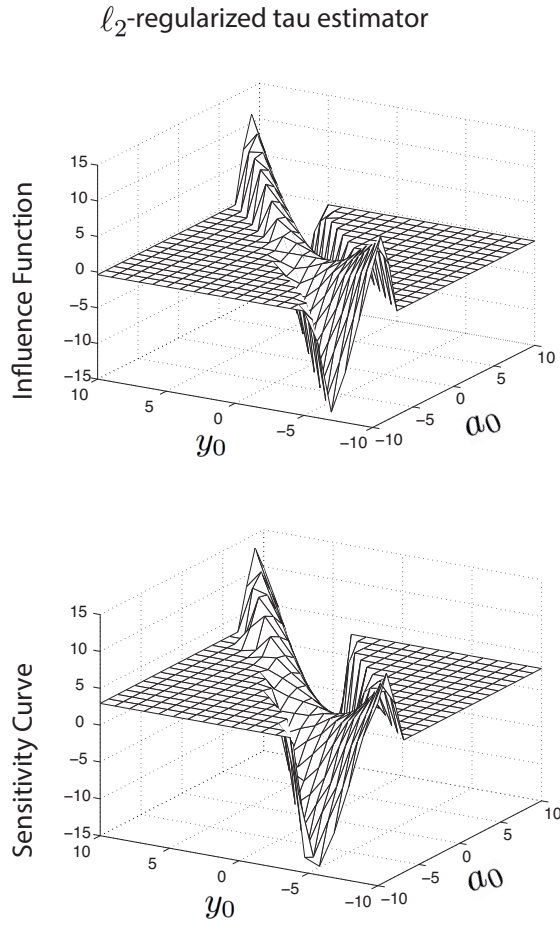
Figure 4.10: IF and SC of the ℓ_2 regularized τ estimator for a 1-dimensional example.

Non-regularized estimators

Here we compare the following estimators

1. **LS** as defined in (1.19).

2. **M estimator with given σ** as defined in 4.4. In this case we assume that the value of σ is known. The ρ function that we use is the Huber's function [41].

3. **M estimator with estimated σ** as defined in 4.4. In this case we assume the value of the scale is unknown and it has to be estimated. To compute $\widehat{\sigma}$ we use the LS estimate $\widehat{\mathbf{x}}_{LS}$ and the normalized median absolute deviation (MADN) [44] which is a robust estimator of the scale of a set of points $\mathbf{v}$

$$\text{MADN}(\mathbf{v}) = \frac{1}{0.6745}\text{median}(|\mathbf{v} - \text{median}(\mathbf{v})|). \tag{4.66}$$

Thus we estimate $\widehat{\sigma}$ as

$$\widehat{\sigma} = \text{MADN}(\mathbf{r}(\widehat{\mathbf{x}}_{\text{LS}})). \tag{4.67}$$

ℓ_1-regularized tau estimator
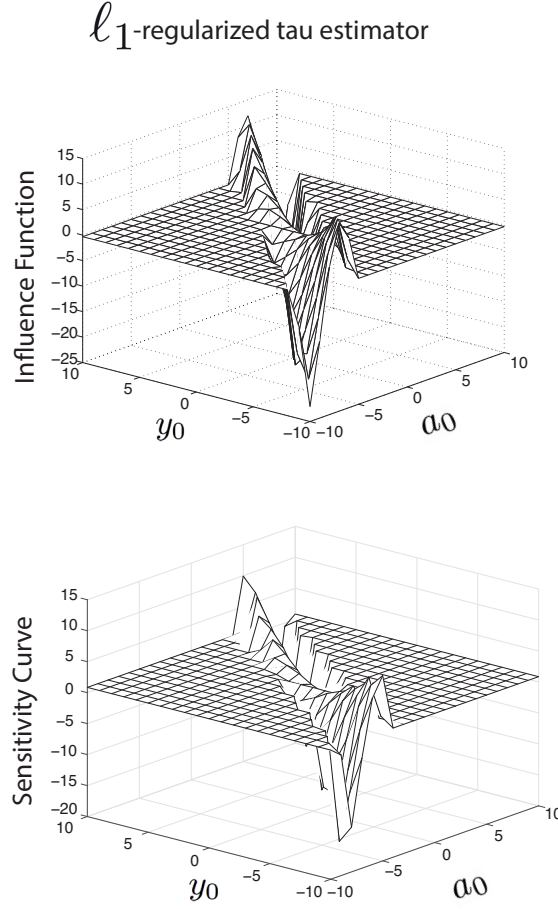
Figure 4.11: IF of the ℓ_1 regularized τ estimators for a 1-dimensional example.

The ρ function that we use is the Huber's function [41].

4. τ **estimator** as given in (4.20) with $\lambda = 0$. We choose ρ_1 and ρ_2 to be optimal weight functions [60], shown in Figure 4.8.

To perform the study, we run numerical simulations with the following setup: we generate a matrix $\mathbf{A} \in \mathbb{R}^{60 \times 20}$ with random i.i.d. Gaussian entries. The measurements $\mathbf{y}$ are generated using additive outliers and additive Gaussian noise

$$\mathbf{y} = \mathbf{Ax} + \mathbf{e}_G + \mathbf{e}_o, \tag{4.68}$$

where $\mathbf{e}_G \in \mathbb{R}^{60 \times 1}$ has random i.i.d. Gaussian entries. On the other hand, $\mathbf{e}_o \in \mathbb{R}^{60 \times 1}$ is a sparse vector. The few non-zero randomly selected entries have a random Gaussian value with zero mean and variance equal to 10 times the variance of the noiseless data. These entries represent the outliers. For each experiment, we perform 100 realizations. In the case of the regularized estimators, we manually select the optimum regularization parameter λ with respect to the MSE. We carry out experiments with 0%, 10%, 20%, 30%, 40% of outliers in the data. Results are given in the upper part of Figure 4.12.

All the compared non-regularized estimators are unbiased, so the MSE is equivalent to the variance of the estimators. With clean data (0% outliers) the variance of the τ estimator is slightly larger than the variance of the other estimators. However, the τ estimator performs better when there are more outliers in the data, it even outperforms the M estimator with a given scale when $30-40\%$ of the data is contaminated.

We notice that the performance of the M estimator is very sensitive to the value of the scale. Its best performance is given when σ is given. Unfortunately this is not often the case, and σ should be estimated. The quality of this estimate will define also the performance of the M estimator.

In this setting, the τ estimator and the M estimator with the given σ reach a BP $= 0.5$, which is the maximum possible value. However notice that the τ does not need any estimate of the scale σ.

Estimators regularized with the ℓ_2 norm

Here we compare the following estimators

1. **LS regularized with the ℓ_2 norm** as defined in (1.10).

2. **M estimator with given σ and regularized with the ℓ_2 norm** as

$$\widehat{\mathbf{x}}_{\mathrm{M}} = \arg\min_{\mathbf{x}} \sum_{i=1}^{m} \rho\left(\frac{r_i}{\sigma}\right) + \lambda\|\mathbf{x}\|_2^2. \tag{4.69}$$

   In this case we assume that the value of σ is known. The ρ function that we use is the Huber's function [41].

3. **M estimator with estimated σ and regularized with the ℓ_2 norm** as defined in (4.69). In this case we assume the value of the scale is unknown and it has to be estimated. To compute $\widehat{\sigma}$ we use (4.67).

4. **τ estimator regularized with the ℓ_2 norm** as given in (4.20).

The setup of the experiments is the same except that the condition number of the matrix $\mathbf{A}$ in this case is 1000. The parameter λ in all the estimators is adjusted to its optimal value, i.e. the value that minimizes the MSE in each case. Results are given in the middle plot of Figure 4.12.

Here we find a similar behaviour as in the first experiment: the τ estimator is slightly worse than the M and LS estimators when the data is not contaminated, but it surpases the performance of all the others when $30-40\%$ of the data are outliers. Since we select the optimal regularization λ, there is a bound on the error of all the regularized estimators. In other words, if we increase λ sufficiently, we will force the algorithm to return the $\widehat{\mathbf{x}} = 0$ solution. That is the worse case solution, corresponding to a breakdown, and we can see it clearly in the LS estimator.

Again, as in the previous experiment for the non-regularized case, the τ estimator and the M estimator with the given σ reach a BP $= 0.5$, which is the maximum possible value.

Estimators regularized with the ℓ_1 norm

In this part compare the following estimators

1. **LS regularized with the ℓ_1 norm** as defined in (1.17).

2. **M estimator with given σ and regularized with the ℓ_1 norm** as

$$\widehat{\mathbf{x}}_{\mathrm{M}} = \arg\min_{\mathbf{x}} \sum_{i=1}^{m} \rho\left(\frac{r_i(\mathbf{x})}{\sigma}\right) + \lambda\|\mathbf{x}\|_1. \tag{4.70}$$

   In this case we assume that the value of σ is known. The ρ function that we use is the Huber's function [41].

3. **M estimator with estimated σ and regularized with the ℓ_1 norm** as defined in (4.70). In this case we assume the value of the scale is unknown and it has to be estimated. To compute $\widehat{\sigma}$ we use (4.67).

4. **τ estimator regularized with the ℓ_1 norm** as given in (4.21).

The setup of the experiments is the same as in the ℓ_2 case except that here we use a sparse source $\mathbf{x}$ with 20% of non zero entries. Here as well the parameter λ in all the estimators is adjusted to its optimal value, i.e. the value that minimizes the MSE in each case. Results are given in the plot in the lower part of Figure 4.12.

The results are very similar to the last experiment, with the difference that, in the case, the M estimator with the estimated scale does not performs better than LS.

Again, as in the two previous experiments for the non-regularized case, the τ estimator and the M estimator with the given σ reach a BP = 0.5, which is the maximum possible value. Thus we can say that the regularizations does not change the BP of the τ estimate, and it keeps being the maximum possible also in the regularized cases.

4.6.4   Bias and Asymptotic Variance

In this section we study the behaviour of the regularized τ estimator when there are no outliers in the data, i.e. when the errors are Gaussian. In particular, we study how the introduction of the regularization affects the bias and the asymptotic variance of the estimator.

The asymptotic variance (ASV) is defined as

$$\mathrm{ASV}(\widehat{\mathbf{x}}, H) = m \lim_{m\to\infty} \mathrm{Var}(\widehat{\mathbf{x}}_m), \tag{4.71}$$

We explore it in a one-dimensional example with numerical simulations. In this case, the source x_0 is set to 1.5, the model a is a standard Gaussian random variable and the measurements y are generated adding standard Gaussian errors e. We use $m = 5000$ measurements.

The upper part of Figure 4.13 shows the ASV of the ℓ_2- regularized τ estimator for different values of the parameter λ. We can observe that, as λ grows, the ASV of this estimator decreases. However, in the ℓ_1 regularized case, the opposite happens. It is shown in the lower part of Figure 4.13: as the value of λ increases, the ASV of the estimator increases as well.

The non-regularized τ estimator has zero bias [76], but the introduction of the regularization biases the estimator. To study this bias in the ℓ_2 and ℓ_1 regularized cases, we again use the one-dimensional example described above. To compute the expectation in (4.62) we use Monte Carlo integration. Results are shown in Figure 4.14. With both regularizations the magnitude of the bias increases with λ.
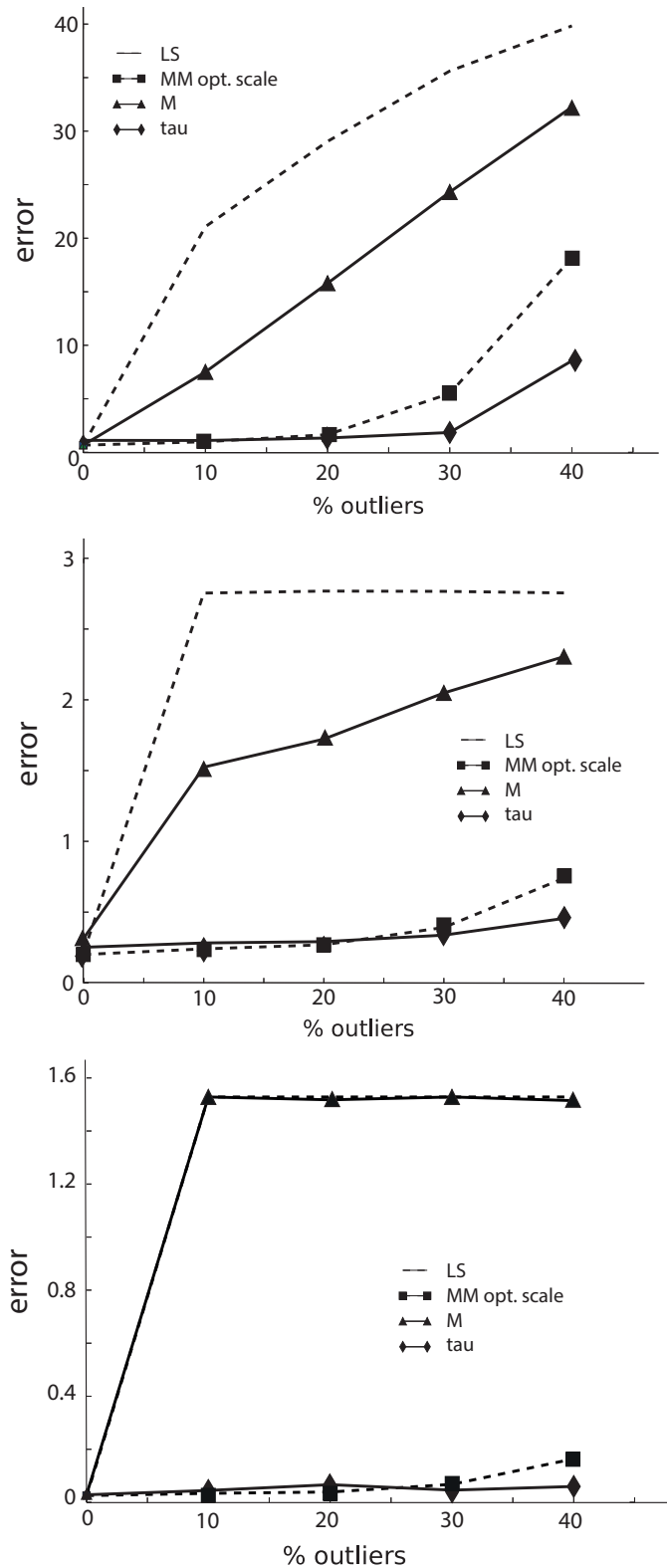
Figure 4.12: Results of the experiments to explore and compare the behaviour of the τ estimators in the presence of different proportions of additive outliers. The plot on the left corresponds to non-regularized estimators, the plot in the center corresponds to ℓ_2-regularized estimators and the plot in the right corresponds to ℓ_1-regularized estimators.
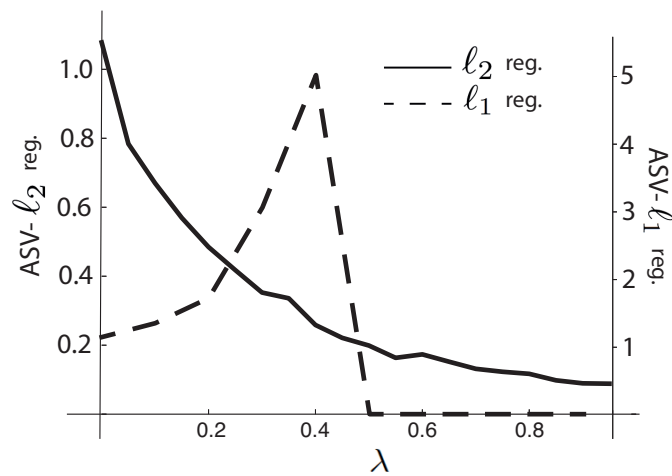
Figure 4.13: The figure represents the ASV for the ℓ_2 and ℓ_1 regularized estimator in a one-dimensional problem.
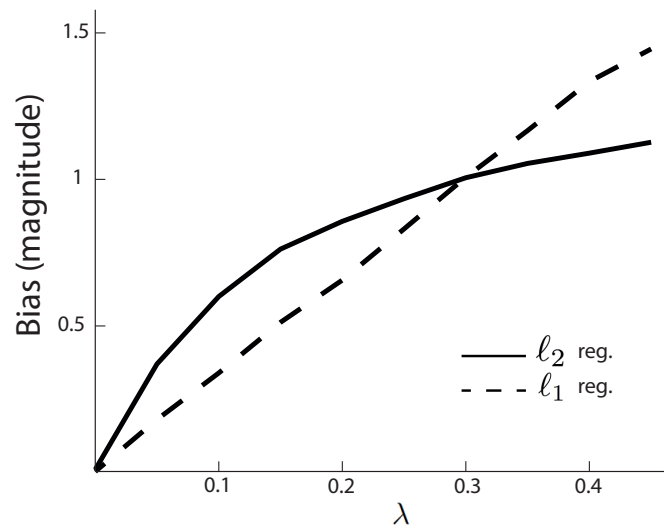
Figure 4.14: Bias of the regularized τ estimator in a one-dimensional problem. The solid line corresponds to an ℓ_2 regularized estimator. The dashed line corresponds to an ℓ_1 regularized estimator.

4.7 The τ estimator with real-world data

In this section we evaluate the performance of our regularized τ estimator using the ETEX dataset that was described in Chapter 3. Recall that this dataset is special because it contains also the ground truth $\mathbf{x}$, so we can measure the MSE of the different estimates. It is composed of the model matrix $\widehat{\mathbf{A}}$, the measurements $\widehat{\mathbf{y}}$, and the ground truth $\mathbf{x}$. The model matrix $\widehat{\mathbf{A}}$ has a large condition number and, in addition, as we showed in Section 3.2, it contains outliers. Thus, we expect that a regularized algorithm that is also robust against outliers, like the regularized

τ estimator, will surpass the performance of estimators such as the classic LS with Tikhonov regularization given in (1.10). Given that the problem is ill-posed, it is also reasonable to expect that the introduction of the regularization in the τ estimator will improve its performance. To confirm these expectations, we evaluate first how the introduction of the regularization affects the performance of the τ estimator. Next, we compare the overall performance of our regularized τ estimator against the benchmark of the classic LS with Tikhonov regularization given in (1.10). Finally, we also compare it to LS with Tikhonov regularization combined with TRANSAC, the preprocessing step to eliminate outliers introduced in Chapter 3.

4.7.1 Regularization effect

Given that the condition number of the matrix $\widehat{\mathbf{A}}$ in the ETEX dataset is large, we expect that the introduction of a regularization to the τ estimator stabilizes the solution, thus improving the performance of the estimator in terms of the MSE. To check this, we compute the estimate $\widehat{\mathbf{x}}_\lambda$ using the τ estimator regularized with the ℓ_2 norm given in (4.20). As we are interested in observing the effect of regularization, we fix the clipping parameters to $c_1 = 1.89$ and $c_2 = 3.55$, and compute $\widehat{\mathbf{x}}_\lambda$ for varying values of the regularization parameter λ. Let n be the number of elements of $\mathbf{x}$. Then, the MSE for each λ value is defined as

$$\text{MSE} = \frac{1}{n}\|\mathbf{x} - \widehat{\mathbf{x}}_\lambda\|_2^2. \tag{4.72}$$

The MSE corresponding to different values of λ is shown in Figure 4.15. It is important to note that $\lambda = 0$ would represent the non-regularized τ estimator. As λ increases, the MSE decreases, as expected. The relative improvement is significant: the MSE drops to half of its initial value. Also we can see that, if the value of λ is too large, the MSE increases again with respect to its minimum value. In fact, if λ keeps growing, the estimate $\widehat{\mathbf{x}}_\lambda$ will tend to zero. Thus, we can conclude that also in practical real-world problems, if the model matrix has a large condition number and we choose a proper value for λ, the introduction of a regularization improves the performance of the τ estimator. As we said in previous chapters, finding the optimal value for λ blindly is a challenge, and is still an open research problem. Methods for this exist, but none of them give theoretical guarantees of finding the optimal value of λ. In order to compare different algorithms fairly, in the rest of this chapter we will always use the optimal performance of each estimators, i.e. we find the optimal value of λ for that estimator. We do so by using a grid search and the ground truth $\mathbf{x}$.

4.7.2 Overall performance

Now we compare the optimal performance of several estimators using the ETEX dataset, i.e. we fix the parameters of all the estimators to the values that produce the minimum MSE. In particular, we compare the errors produced by:

- LS with Tikhonov regularization introduced in (1.10), which uses as loss function the ℓ_2 norm,
- the above estimator combined with TRANSAC, a preprocessing step introduced in Chapter 3 to localize and remove outliers, and
- τ estimator regularized by the ℓ_2 norm given in this chapter in (4.20).
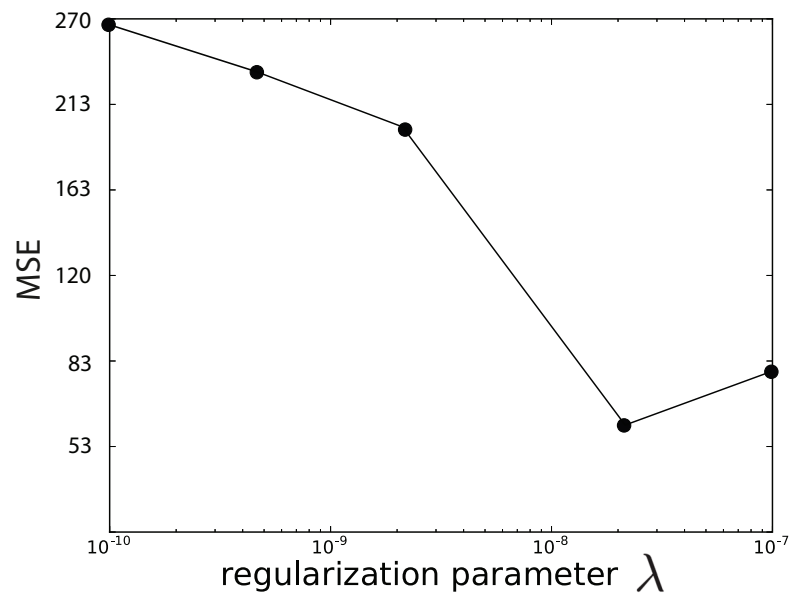
Figure 4.15: Regularization of the τ estimator improves its performance

It is important to note that all these estimators use the same regularization. Hence, the comparison will show the effect, in this particular real-world dataset, of the different techniques to fight outliers. The results of this comparison are shown below.

estimator	LS + Tikhonov	regularized τ	TRANSAC
MSE	23.58	21.88	15.09
improvement	baseline	7.2%	36.0%

The results are also given in Figure 4.16. The τ estimator outperformed the classic LS with Tikhonov regularization, as expected. Hence, introducing a robust loss function indeed reduced the effect of outliers. Similarly, TRANSAC also outperformed the classic LS with Tikhonov regularization. Hence, the introduction of a pre-processing step to remove outliers was very effective. In the ETEX dataset, the latter approach is clearly more efficient.
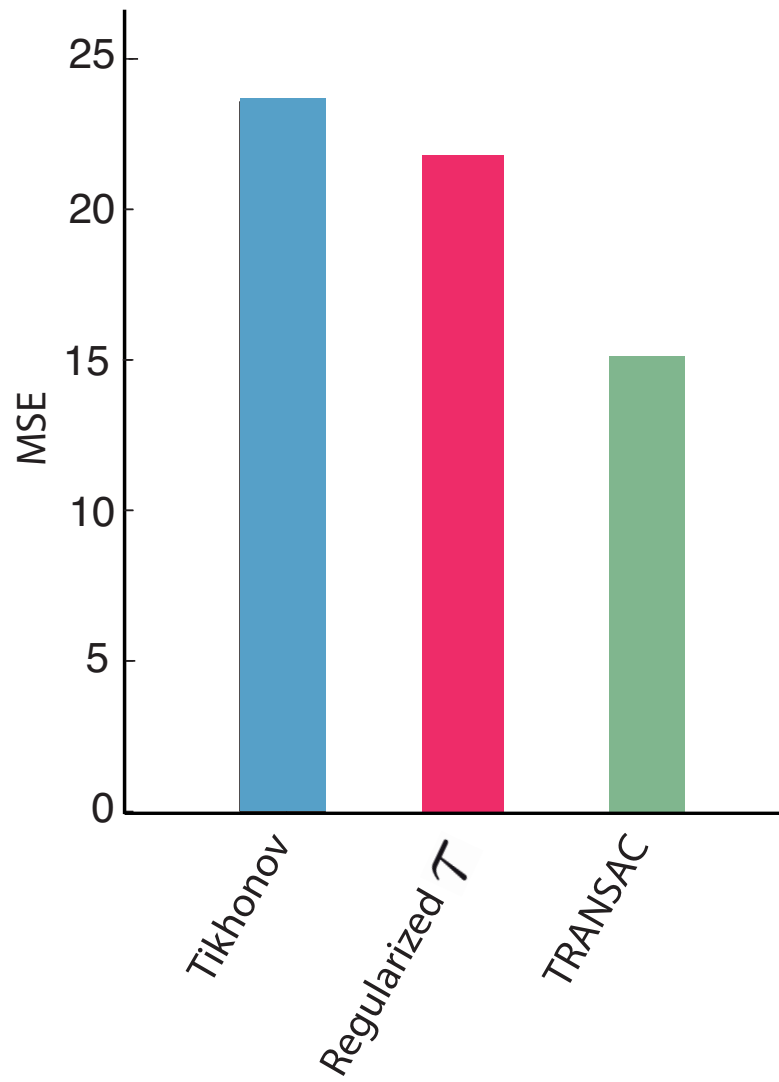
Figure 4.16: Performance in the ETEX dataset of the classical LS + Tikhonov estimator given in (1.10), the τ estimator regularized with the ℓ_2 norm given in (4.20), and the LS + Tikhonov estimator with TRANSAC, a preprocessing step to localize and remove outliers introduced in Chapter 3.

4.8   Appendix

4.8.1   Finding local minima

Define a twice differentiable regularization function $J(\mathbf{x})$. Then, every local minimum satisfies

$$\frac{\partial(\hat{\sigma}_\tau^2(\mathbf{r}(\mathbf{x})) + \lambda J(\mathbf{x}))}{\partial \mathbf{x}} = 0. \tag{4.73}$$

To have a more compact notation, define

$$\begin{aligned}
\tilde{\mathbf{r}} &= \frac{\mathbf{r}(\mathbf{x})}{\hat{\sigma}_M(\mathbf{r}(\mathbf{x}))}, \\
\psi_j(u) &= \frac{\partial \rho_j(u)}{\partial u}, \\
\sigma_M(\mathbf{r}) &= \hat{\sigma}_M(\mathbf{r}(\mathbf{x})), \\
\sigma_\tau(\mathbf{r}) &= \hat{\sigma}_\tau(\mathbf{r}(\mathbf{x})).
\end{aligned}$$

Using the definition of $\sigma_\tau(\mathbf{r})$ in Equation (4.18) and of $\sigma_M(\mathbf{r})$ in Equation (4.16), and defining

$$J'(\mathbf{v}) = \frac{\partial J(\mathbf{v})}{\partial \mathbf{v}} = \left[\frac{\partial J(\mathbf{v})}{\partial v_1}, \dots, \frac{\partial J(\mathbf{v})}{\partial v_n}\right]^\top, \tag{4.74}$$

$$\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_m]^\top, \tag{4.75}$$

we get

$$\frac{\partial(\sigma_\tau^2(\mathbf{r}) + \lambda J(\mathbf{x}))}{\partial \mathbf{x}} = \tag{4.76}$$

$$2\sigma_M(\mathbf{r})\frac{\partial \sigma_M(\mathbf{r})}{\partial \mathbf{x}}\frac{1}{m}\sum_{i=1}^m \rho_2(\tilde{r}_i) + \sigma_M^2(\mathbf{r})\frac{1}{m}\sum_{i=1}^m \psi_2(\tilde{r}_i)\left[\frac{-\mathbf{a}_i^\top \sigma_M(\mathbf{r}) - r_i\frac{\partial \sigma_M(\mathbf{r})}{\partial \mathbf{x}}}{\sigma_M^2(\mathbf{r})}\right] +$$

$$+\lambda J'(\mathbf{x}) = 0. \tag{4.77}$$

Taking the derivative of $\sigma_M(\mathbf{r})$ with respect to $\mathbf{x}$ we obtain

$$\frac{\partial \sigma_M(\mathbf{r})}{\partial \mathbf{x}} = -\sigma_M(\mathbf{r})\frac{\sum_{i=1}^m \psi_1(\tilde{r}_i)\mathbf{a}_i^\top}{\sum_{i=1}^m \psi_1(\tilde{r}_i)r_i}. \tag{4.78}$$

Replacing (4.78) in (4.77), and defining

$$w_m(\mathbf{x}) = \frac{\sum_{i=1}^m 2\rho_2(\tilde{r}_i) - \sum_{i=1}^m \psi_2(\tilde{r}_i)\tilde{r}_i}{\sum_{i=1}^m \psi_1(\tilde{r}_i)\tilde{r}_i}, \tag{4.79}$$

we obtain

$$-\frac{1}{m}\sum_{i=1}^m \Big(w_m(\mathbf{x})\psi_1(\tilde{r}_i) + \psi_2(\tilde{r}_i)\Big)\sigma_M(\mathbf{r})\mathbf{a}_i^\top + \lambda J'(\mathbf{x}) = 0, \tag{4.80}$$

where

$$w_m(\mathbf{x}) := \frac{\sum_{i=1}^m 2\rho_2(\tilde{r}_i(\mathbf{x})) - \sum_{i=1}^m \psi_2(\tilde{r}_i(\mathbf{x}))\tilde{r}_i(\mathbf{x})}{\sum_{i=1}^m \psi_1(\tilde{r}_i(\mathbf{x}))\tilde{r}_i(\mathbf{x})}. \tag{4.81}$$

If we reorganize the expressions given above, we can see that (4.80) is also the derivative of a penalized reweighted least squares function

$$f(\mathbf{x}) = \frac{1}{m} \sum_{i=1}^{m} z(\tilde{r}_i(\mathbf{x}))(y_i - \mathbf{a}_i\mathbf{x})^2 + \lambda J(\mathbf{x}) \tag{4.82}$$

with weights

$$z(u) = \begin{cases} \frac{\psi_\tau(u)}{2u} & \text{if } u \neq 0 \\ 1 & \text{otherwise,} \end{cases} \tag{4.83}$$

where

$$\psi_\tau(u) = w_m(\mathbf{x})\psi_1(u) + \psi_2(u). \tag{4.84}$$

ℓ_1 regularization

The ℓ_1 norm is not a differentiable function. To overcome this problem, we approximate $\|\mathbf{x}\|_1$ with a differentiable function. We choose

$$J_k(\mathbf{v}) = \sum_{i=1}^{m} v_i \tanh(kv_i), \tag{4.85}$$

where k is a non-negative scalar and

$$\lim_{k \to \infty} J_k(\mathbf{v}) = \|\mathbf{v}\|_1. \tag{4.86}$$

Then, we perform the same derivation as in the last section, using $J(\mathbf{v}) = J_k(\mathbf{v})$. In this case, Equation (4.82) becomes

$$f_k(\mathbf{x}) = \sum_{i=1}^{m} w_i(\mathbf{x})(y_i - \mathbf{a}_i\mathbf{x})^2 + \lambda J_k(\mathbf{x}). \tag{4.87}$$

Now, taking limits

$$f(\mathbf{x}) = \lim_{k \to \infty} f_k(\mathbf{x}) = \sum_{i=1}^{m} w_i(\mathbf{x})(y_i - \mathbf{a}_i\mathbf{x})^2 + \lambda J(\mathbf{x}). \tag{4.88}$$

So minimizing (4.19) with $J(\mathbf{x}) = \|\mathbf{x}\|_1$ is equivalent to minimizing

$$f_1(\mathbf{x}) = \sum_{i=1}^{m} w_i(\mathbf{x})(y_i - \mathbf{a}_i\mathbf{x})^2 + \lambda\|\mathbf{x}\|_1. \tag{4.89}$$

4.8.2 Influence function of the τ estimator with a twice differentiable regularization function

Proof.
To have a more compact notation define

$$
\tilde{r} = \frac{r(\widehat{\mathbf{x}}_\tau)}{\hat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))},
$$

$$
\psi_j(u) = \frac{\partial \rho_j(u)}{\partial u},
$$

$$
\sigma_M(r) = \hat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau)),
$$

$$
(4.90)
$$

where $\sigma_M(\mathbf{r})$ in given in Equation (4.46) as

$$
\mathbb{E}_{H_\epsilon}\left[\rho_1\left(\frac{r(\widehat{\mathbf{x}}_\tau)}{\hat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))}\right)\right] = b, \tag{4.91}
$$

where $b = E_{H_0}[\rho_1(u)]$. Define also

$$
J'(\mathbf{v}) = \frac{\partial J(\mathbf{v})}{\partial \mathbf{v}} = \left[\frac{\partial J(\mathbf{v})}{\partial v_1}, \ldots, \frac{\partial J(\mathbf{v})}{\partial v_n}\right]^\top, \tag{4.92}
$$

$$
\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_m]^\top, \tag{4.93}
$$

$$
w(\tilde{r}) = \frac{\mathbb{E}_{H_\epsilon}[2\rho_2(\tilde{r}) - \psi_2(\tilde{r})\tilde{r}]}{\mathbb{E}_{H_\epsilon}[\psi_1(\tilde{r})\tilde{r}]},
$$

$$
\psi_\tau(\tilde{r}) = w(\tilde{r})\psi_1(\tilde{r}) + \psi_2(\tilde{r}), \tag{4.94}
$$

$$
(4.95)
$$

From Equation (4.80), we know that, at the contaminated distribution H_ϵ, the estimate $\widehat{\mathbf{x}}_\tau$ given in (4.48) has to satisfy

$$
-\mathbb{E}_{H_\epsilon}[\psi_\tau(\tilde{r})\hat{\sigma}_M(r)\mathbf{a}^\top] + \lambda J'(\widehat{\mathbf{x}}_\tau) = 0. \tag{4.96}
$$

Defining the contaminated distribution as $H_\epsilon = (1 - \epsilon)H_0 + \epsilon\delta_{\mathbf{a}_0, y_0}$, we can rewrite (4.96) as

$$
-(1 - \epsilon)\mathbb{E}_{H_0}[(\psi_\tau(\tilde{r}))\sigma_M(r)\mathbf{a}^\top] - \epsilon\left(\psi_\tau\left(\frac{y_0 - \mathbf{a}_0\widehat{\mathbf{x}}_\tau}{\hat{\sigma}_M(r)}\right)\right)\sigma_M(r)\mathbf{a}_0^\top + \lambda J'(\widehat{\mathbf{x}}_\tau) = 0. \tag{4.97}
$$

Taking derivatives w.r.t. ϵ

$$
\mathbb{E}_{H_0}[(\psi_\tau(\tilde{r}))\hat{\sigma}_M(r)\mathbf{a}^\top] - (1 - \epsilon)\mathbb{E}_{H_0}\left[\mathbf{a}^\top\frac{\partial(\psi_\tau(\tilde{r})\hat{\sigma}_M(r))}{\partial\epsilon}\right] -
$$

$$
-\epsilon\left(\frac{\partial(\psi_\tau\left(\frac{y_0 - \mathbf{a}_0\widehat{\mathbf{x}}_\tau}{\hat{\sigma}_M(r)}\right)\hat{\sigma}_M(r)\mathbf{a}_0^\top)}{\partial\epsilon}\right) -
$$

$$
-\psi_\tau\left(\frac{y_0 - \mathbf{a}_0\widehat{\mathbf{x}}_\tau}{\hat{\sigma}_M(r)}\right)\hat{\sigma}_M(r)\mathbf{a}_0^\top + \lambda J''(\widehat{\mathbf{x}}_\tau)\frac{\partial\widehat{\mathbf{x}}_\tau}{\partial\epsilon} = 0, \tag{4.98}
$$

where $J''(\widehat{\mathbf{x}}^{\tau_R})$ represents the Jacobian matrix. Using the chain rule for derivation, we get

$$
\frac{\partial(\psi_\tau(\tilde{r})\hat{\sigma}_M(r))}{\partial\epsilon} = \frac{\partial(\psi_\tau(\tilde{r})\hat{\sigma}_M(r))}{\partial\widehat{\mathbf{x}}_\tau}\frac{\partial\widehat{\mathbf{x}}_\tau}{\partial\epsilon}, \tag{4.99}
$$

where

$$
\frac{\partial(\psi_\tau(\tilde{r})\hat{\sigma}_M(r))}{\partial\widehat{\mathbf{x}}_\tau} = \frac{\partial\psi_\tau(\tilde{r})}{\partial\widehat{\mathbf{x}}_\tau}\hat{\sigma}_M(r)\psi_\tau(\tilde{r})\frac{\partial\hat{\sigma}_M(r)}{\partial\widehat{\mathbf{x}}_\tau}, \tag{4.100}
$$

and

$$\frac{\partial \psi_\tau(\tilde{r})}{\partial \widehat{\mathbf{x}}_\tau} = \frac{\partial \psi_\tau(\tilde{r})}{\partial \tilde{r}} \frac{\partial \tilde{r}}{\partial \widehat{\mathbf{x}}_\tau}. \tag{4.101}$$

We already know from (4.78) that

$$\frac{\partial \hat{\sigma}_M(r)}{\partial \widehat{\mathbf{x}}_\tau} = -\hat{\sigma}_M(r) \frac{\mathbb{E}_{H_\epsilon}[\psi_1(\tilde{r})\mathbf{a}^\top]}{\mathbb{E}_{H_\epsilon}[\psi_1(\tilde{r})r]}. \tag{4.102}$$

We also need

$$\frac{\partial \psi_\tau(\tilde{r})}{\partial \tilde{r}} = \frac{\partial w(\tilde{r})}{\partial \tilde{r}} \psi_1(\tilde{r}) + w(\tilde{r})\psi_1'(\tilde{r}) + \psi_2'(\tilde{r}), \tag{4.103}$$

that uses

$$\begin{aligned}\frac{\partial w(\tilde{r})}{\partial \tilde{r}} &= \frac{\mathbb{E}_{H_\epsilon}[2\psi_2(\tilde{r} - 2\rho_2(\tilde{r})]}{\mathbb{E}_{H_\epsilon}[\psi_1(\tilde{r})\tilde{r}]^2} - \frac{\mathbb{E}_{H_\epsilon}[\psi_2'(\tilde{r})\tilde{r}]}{\mathbb{E}_{H_\epsilon}[\psi_1(\tilde{r}(\widehat{\mathbf{x}}_\tau))\tilde{r}(\widehat{\mathbf{x}}_\tau)]^2} - \\ &\quad - \frac{\mathbb{E}_{H_\epsilon}[\psi_2(\tilde{r}))\psi_1(\tilde{r})\tilde{r}]}{\mathbb{E}_{H_\epsilon}[\psi_1(\tilde{r})\tilde{r}]^2} - \frac{\mathbb{E}_{H_\epsilon}[-\psi_2(\tilde{r})\tilde{r}(\psi_1'(\tilde{r})\tilde{r} + \psi_1(\tilde{r}))]}{\mathbb{E}_{H_\epsilon}[\psi_1(\tilde{r})\tilde{r}]^2}.\end{aligned} \tag{4.104}$$

In Equation (4.101), we also need

$$\frac{\partial \tilde{r}}{\partial \widehat{\mathbf{x}}_\tau} = \frac{-\mathbf{a}^\top \hat{\sigma}_M(r) - r\frac{\partial \hat{\sigma}_M(r)}{\partial \widehat{\mathbf{x}}_\tau}}{\hat{\sigma}_M^2(r)} \widehat{\mathbf{x}}_\tau. \tag{4.105}$$

We have now all the elements of Equation (4.98). The next step, in order to find the IF, is to particularize Equation (4.98) for $\epsilon = 0$. Since

$$\left.\frac{\partial \widehat{\mathbf{x}}(H_\epsilon)}{\partial \epsilon}\right|_{\epsilon=0} = \text{IF}(\mathbf{a}_0, y_0), \tag{4.106}$$

we can write

$$\begin{aligned}\mathbb{E}_{H_0}[\psi_\tau(\tilde{r})\hat{\sigma}_M(r)\mathbf{a}^\top] \quad &-\mathbb{E}_{H_0}\left[\mathbf{a}^\top \frac{\partial \psi_\tau(\tilde{r})\hat{\sigma}_M(r)}{\partial r} \frac{\partial r}{\partial \widehat{\mathbf{x}}_\tau}\right] \text{IF}(\mathbf{a}_0, y_0) - \\ &- \left(\psi_\tau\left(\frac{y_0 - \mathbf{a}_0\widehat{\mathbf{x}}_\tau}{\hat{\sigma}_M(r)}\right)\right)\hat{\sigma}_M(r)\mathbf{a}_0^\top + \lambda J''(\widehat{\mathbf{x}}_\tau)\text{IF}(\mathbf{a}_0, y_0) = 0. \quad (4.107)\end{aligned}$$

Solving the last equation for $\text{IF}(\mathbf{a}_0, y_0)$, we get

$$\begin{aligned}\text{IF}(\mathbf{a}_0, y_0) \quad &= \left(-\mathbb{E}_{H_0}[\mathbf{a}^\top \frac{\partial \psi_\tau(\tilde{r})\hat{\sigma}_M(r)}{\partial \widehat{\mathbf{x}}_\tau}] + \lambda J''(\widehat{\mathbf{x}}_\tau)\right)^{-1} \\ &\left(\psi_\tau\left(\frac{y_0 - \mathbf{a}_0\widehat{\mathbf{x}}_\tau}{\hat{\sigma}_M(r)}\right)\hat{\sigma}_M(r)\mathbf{a}_0^\top - \mathbb{E}_{H_0}\left[(\psi_\tau(\tilde{r}))\hat{\sigma}_M(r)\mathbf{a}^\top\right]\right).\end{aligned} \tag{4.108}$$

$\square$

4.8.3   Influence function for the τ estimator with ℓ_1 regularization

Proof.

We approximate $J(x) = |x|$ with a twice differentiable function

$$
\begin{aligned}
J_k(x) &= x \tanh(kx) \\
\lim_{k \to \infty} J_k(x) &= |x|,
\end{aligned}
\tag{4.109}
$$

where k is a non-negative scalar. The second derivative of $J_k(x)$ is

$$
J_k''(x) = 2k(1 - \tanh^2(kx)) - 2k^2 x \tanh(kx)(1 - \tanh^2(kx)).
\tag{4.110}
$$

In particular,

$$
J_k''(0) = 2k.
\tag{4.111}
$$

For $J(x)$ twice differentiable, we know that the IF of a regularized τ estimator is (4.108). Without loss of generality, we can assume that the estimate $\widehat{\mathbf{x}}_\tau \in \mathbb{R}^{n \times 1}$ is t-sparse, i.e. only t out of the n entries of the vector are non-zero. Remember that $\mathbf{a} \in \mathbb{R}^{1 \times n}$. Then,

$$
\begin{aligned}
\mathrm{IF}(\mathbf{a}_0, y_0) = \quad & \left(-\mathbb{E}_{H_0} \left[\mathbf{a}^\top \frac{\partial(\psi_\tau(\tilde{r})\hat{\sigma}_M(r))}{\partial \widehat{\mathbf{x}}_\tau} \right] + \right. \\
& + 2\lambda \mathrm{diag}((J''(\widehat{\mathbf{x}}_\tau))_1, \ldots (J''(\widehat{\mathbf{x}}_\tau))_t, 2k, \ldots, 2k))^{-1} + \\
& + \left(\psi_\tau \left(\frac{y_0 - \mathbf{a}_0^\top \widehat{\mathbf{x}}_\tau}{\hat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))} \right) \right) \hat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))\mathbf{a}_0^\top - \\
& - \mathbb{E}_{H_0} \left[(\psi_\tau(\tilde{r}(\widehat{\mathbf{x}}_\tau))) \, \hat{\sigma}_M(r(\widehat{\mathbf{x}}_\tau))\mathbf{a}^\top \right],
\end{aligned}
$$

where $\mathrm{diag}(\mathbf{u})$ is a diagonal matrix whose diagonal is $\mathbf{u}$, $(\mathbf{u})_i$ is the i-th element of the vector $\mathbf{u}$ and $J''(\mathbf{u})$ is the vector whose i-th element is $J''((\mathbf{u})_i)$.

Now, we focus on the inversion of the matrix. For that, let us work with block matrices. If we call

$$
-\mathbb{E}_{H_0} \left[\mathbf{a}^\top \frac{\partial(\psi_\tau(\tilde{r})\hat{\sigma}_M(r))}{\partial \widehat{\mathbf{x}}_\tau} \right] = \mathbf{E} = \left(\begin{array}{c|c} \mathbf{E}_{11} & \mathbf{E}_{12} \\ \hline \mathbf{E}_{21} & \mathbf{E}_{22} \end{array} \right),
\tag{4.112}
$$

where $\mathbf{E} \in \mathbb{R}^{n \times n}$, $\mathbf{E}_{11} \in \mathbb{R}^{t \times t}$, $\mathbf{E}_{12} \in \mathbb{R}^{t \times (n-t)}$, $\mathbf{E}_{21} \in \mathbb{R}^{(n-t) \times t}$ and $\mathbf{E}_{22} \in \mathbb{R}^{(n-t) \times (n-t)}$. Then the matrix that we have to invert becomes

$$
\left(\begin{array}{c|c} \mathbf{E}_{11} + 2\lambda \mathrm{diag}(J_k''(\widehat{\mathbf{x}}_\tau)_{1:t})) & \mathbf{E}_{12} \\ \hline \mathbf{E}_{21} & \mathbf{E}_{22} + 2\lambda 2k\mathbf{I}_{n-t} \end{array} \right)^{-1},
\tag{4.113}
$$

where $\mathbf{I}_{n-t}$ is the $(n-t) \times (n-t)$ identity matrix and $(\mathbf{u})_{1:t}$ are the first t elements of the vector $\mathbf{u}$. The inverse of a block matrix can be written as

$$
\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array} \right)^{-1} =
\tag{4.114}
$$

$$
\left(\begin{array}{c|c} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ \hline (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{array} \right),
\tag{4.115}
$$

where $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and $\mathbf{D}$ are matrix blocks of arbitrary size. This inverse exists if $\mathbf{A}$ and $(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})$ are square and nonsingular matrices, so their inverse exists. In our case,

$$\begin{aligned}
\mathbf{A} &= \mathbf{E}_{11} + 2\lambda\mathrm{diag}(J_k''(\widehat{\mathbf{x}}_\tau)_{1:t}))), & \text{(4.116)} \\
\mathbf{B} &= \mathbf{E}_{12}, \\
\mathbf{C} &= \mathbf{E}_{21}, \\
\mathbf{D} &= \mathbf{E}_{22} + 2\lambda 2k\mathbf{I}_{n-t}.
\end{aligned}$$

Let us call

$$\mathbf{Z} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} = \mathbf{E}_{22} + 2\lambda 2k\mathbf{I}_{n-t} - \mathbf{E}_{21}(\mathbf{E}_{11} + 2\lambda\mathrm{diag}(J_k''(\widehat{\mathbf{x}}_\tau)_{1:t})))^{-1}\mathbf{E}_{12}, \tag{4.117}$$

where $\mathbf{Z} \in \mathbb{R}^{(n-t)\times(n-t)}$. We want to know what happens with $\mathbf{Z}$ when $k \to \infty$. In the first place, we know that

$$\lim_{k\to\infty} J_k''(x) = 0 \text{ for } x \neq 0. \tag{4.118}$$

So

$$\mathbf{Z} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} = \mathbf{E}_{22} + 4\lambda k\mathbf{I}_{n-k} - \mathbf{E}_{21}\mathbf{E}_{11}^{-1}\mathbf{E}_{12}. \tag{4.119}$$

Let us also call

$$\mathbf{Y} = \mathbf{E}_{22} - \mathbf{E}_{21}\mathbf{E}_{11}^{-1}\mathbf{E}_{12}. \tag{4.120}$$

If the eigenvalues of $\mathbf{Y}$ are $\nu_1, \ldots, \nu_n$, then the eigenvalues of $\mathbf{Z}$ are $\nu_1 + 4\lambda k, \ldots, \nu_n + 4\lambda k$. If $\mathbf{Z}$ is a square matrix with eigenvalues different from zero, then it can be factorized as $\mathbf{Z} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}$, where $\mathbf{Q}$ is the square matrix whose $i-$th column is the eigenvector $\mathbf{q}_i$ of $\mathbf{Z}$, and $\boldsymbol{\Lambda}$ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues $\nu_i + 4\lambda k$. The inverse of this matrix can be written as $\mathbf{Z}^{-1} = \mathbf{Q}\boldsymbol{\Lambda}^{-1}\mathbf{Q}^{-1}$, so $\mathbf{Z}^{-1} \to 0$ as $k \to \infty$. $\mathbf{Z}^{-1}$ appears in all the components of (4.115)

$$\left(\begin{array}{c|c} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{Z}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{Z}^{-1} \\ \hline \mathbf{Z}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{Z}^{-1} \end{array}\right), \tag{4.121}$$

so in the end we have

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)^{-1} = \left(\begin{array}{c|c} \mathbf{A}^{-1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{E}_{11}^{-1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right). \tag{4.122}$$

From here, it is straight forward to arrive to (4.61). $\qquad\square$

Chapter 5

To code or not to code

Talk is cheap. Show me the code.
- Linus Torvalds

In the previous chapters we have seen many different estimators and discussed their characteristics. Once any estimator is defined, the estimate $\hat{\mathbf{x}}$ still needs to be actually computed. This computation can be achieved by a direct analytical solution, an iterative algorithm that asymptotically converges to the exact estimate, or a heuristic. We have discussed such algorithms for the computation of the regularized τ estimator in Section 4.5.

Given an algorithm, the next step in the process is to create a software implementation of it. Unfortunately, this step is often poorly documented and generally underestimated. Nevertheless, it has two important functions: it enables us to carry out computational experiments, and provides a very effective way to share our research results with the rest of the scientific community. Precisely such software implementations for the algorithms in this thesis are the subject of this chapter.
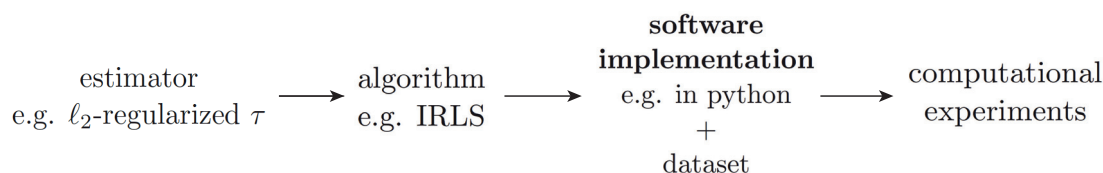
Figure 5.1: Diagram that shows the different steps of a research process.

Computational experiments were used to generate many of the results and figures that we have seen in the previous chapters. Until now these experiments have been described just with words and pseudocode. Also, details such as the exact content of the datasets, parameter values, and initial points have been largely omitted for simplicity. However, these details are necessary to obtain the exact results reported in this thesis. For this reason, in this chapter we provide the software implementations and datasets that fully specify all the results explained in the last chapters. In addition, we also present our efforts to bring our findings to a broader community by making our algorithms open source, reliable, easy to install on any computer, and easy to use.

5.1   Reproducible results

The lack of reproducibility of scientific results generates much skepticism [71]. In recent years, this skepticism transcended the academic world and is now also part of the general public opinion [7]. Since Galileo's times, reproducibility is one of the fundamental pillars of the scientific method. It is the key trait that makes science objective and lends it credibility in the public eye. If reproducibility disappears, it will not be long until science loses the important place it has in today's society.

As we explained above, to make our work easily reproducible we need to provide, with every publication, the software implementation and dataset that we used to generate the results and figures. All such implementations are made freely available via the Audiovisual Communications Lab (LCAV) online reproducible results (RR) platform at `http://rr.epfl.ch`. On this platform the PDF version of the article is available together with the associated software implementation and dataset. Instructions on how to use the software are also provided, together with posters and presentations to facilitate the communication of results. This thesis is no exception: it is also available in the same format on the RR platform. Figure 5.2 shows a screen shot of the RR platform.

Figure 5.2: Screen shot of the RR platform of the LCAV.

In addition to the distribution of results, the RR platform also collects feedback from users. A voting system lets the users give their opinion on the usability and clarity of the code. Another feature prompts the users to report bugs in the code. This feedback allows us to keep continuously improving our code. The RR associated to this thesis are provided at:

– Chapter 2: `http://rr.epfl.ch/paper/MDR2013`
– Chapter 3: `http://rr.epfl.ch/paper/MSV2014`
– Chapter 4: `http://rr.epfl.ch/paper/MMZ2015`

5.2 LinvPy

In addition to making our results reproducible, we also aim to make them as useful as possible to the community. With this goal in mind, we created LinvPy [13], a software implementation that contains the regularized τ estimator that we proposed in Chapter 4 and other robust estimators to solve linear inverse problems. LinvPy is an effort to make the algorithms designed in this thesis reliable, accessible, easy to install, and easy to use for everybody. LinvPy was implemented by Guillaume Beaud, a master student in Computer Science at EPFL who did his optional project at the LCAV.

5.2.1 Specifications

To define the specifications of LinvPy, we performed an informal survey of a dozen researchers who also develop software. We asked them about the features they look for in an implementation of an algorithm, how they use it, what they use it for, and the properties that differentiate the most useful implementations from the rest. This survey lead us to the following list of requirements:

1. **Well tested code**. All the methods should work correctly, i.e. they should not contain bugs. To achieve this, a thorough testing strategy must be in place.

2. **Easy to obtain, easy to install, easy to use**. If the users need too much time to find the code, to install it, and to make it work on their computers, they will not use it. In addition, the users are not ready to invest too much time into learning how to use the software, so it needs to be user-friendly.

3. **Excellent documentation**. The software needs to be well documented: we need to explain what the software does, and how to use it. This documentation should be easily accessible, easy to navigate, and *efficient*, in the sense that it should explain the essential information in just a few lines.

4. **Open source**. By making LinvPy free in the gratis sense (i.e. the users do not have to pay for it), we made it accessible to the widest possible audience. We also made it free in the libre sense (i.e. the users have access to the source code), thus increasing transparency, and enabling contributions by other developers.

5. **Easy to contribute**. We want to encourage the users to contribute to the project with new features and ideas. An open source project allows the users to look into the project, modify it, and contribute with new versions of the software.

5.2.2 Solution: Python package

To fulfil all the specifications we proposed above, we implemented our algorithms in a Python package. Python is a widely-used, high-level, general-purpose, interpreted, dynamic, and open-source programming language. Its design philosophy emphasizes code readability, and its syntax allows programmers to express concepts in fewer lines of code than what is possible in languages such as C++ or Java [4]. A Python package is simply a directory of Python modules. A Python module is a file consisting of Python code. A module can define functions, classes and variables. It can also include runnable code. Python packages are the standard way in Python of distributing methods and classes implemented by other users.

5.2.3 Architecture

LinvPy's architecture is based on two principles: modularity and the master-slave paradigm.

1. **Modularity**

   Modularity is crucial when implementing complex methods. It allows the developer to track errors easily, produce methods that are easier to use, avoids redundancy, and simplifies the process of adding new methods to the package. To achieve this modularity, LinvPy is based on the object-oriented programming paradigm.

   LinvPy is composed of several classes, as we can see in the UML diagram in Figure 5.3. The core of LinvPy is the Estimator class. It aggregates exactly one LossFunction and one Regularization. These three classes contain the attributes and methods that every estimator, regularization, and loss function object will have. Particular estimators inherit from the Estimator class all its methods and attributes. The same is true for particular regularizations, and loss functions. With this structure, we force all the objects of the same type to have the same basic *footprint*. For example all the regularizations, in our case Tikhonov and Lasso, contain the same method *regularize*, that is defined in the class Regularization.
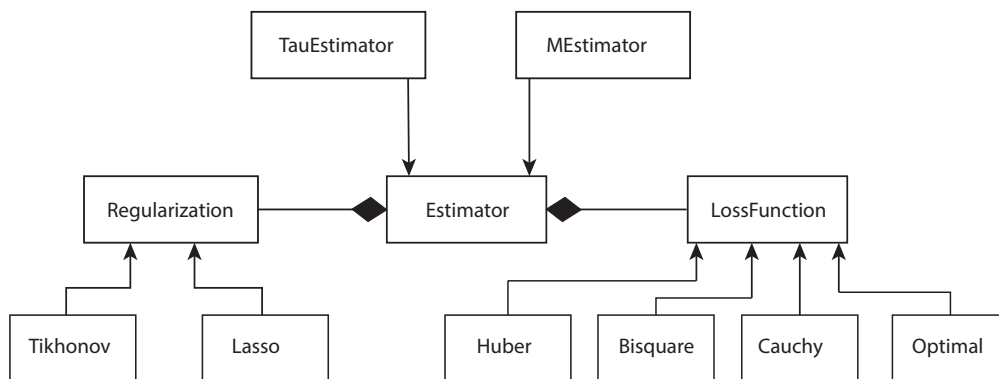
Figure 5.3: UML diagram for LinvPy

This architecture facilitates the implementation and integration of new features. Simply by inheriting from the three super-classes, newly introduced classes will automatically have the necessary structure to work properly with the rest of the methods in the package.

2. **Master-slave paradigm**

   This paradigm defines how the different functions interact to work together. In LinvPy, functions are organized in a hierarchy with different levels. Functions in the upper levels call the functions in the lower levels (see Figure 5.4). Each function does its job when it is called, but it does not know why it is called: low-level functions (slaves) do not know what happens at the upper level (master). This principle avoids interferences between different low-level functions: the master controls everything. In our case, the methods in the class Estimator are masters, whereas the methods in the classes Regularization and LossFunction are slaves.
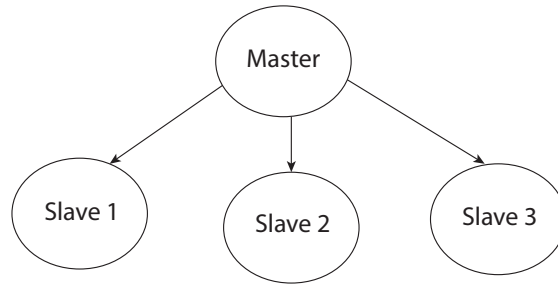
Figure 5.4: Master-slave paradigm.

5.2.4 Distribution

LinvPy and its documentation are hosted in a public repository at `https://github.com/LCAV/linvpy`. This repository can be cloned easily and downloaded by anyone. There are two ways users can install LinvPy on their local computer:

1. Downloading the code directly from the repository and then running the setup file.

2. Automatic download and installation from the Python Package Index, PyPI (`https://pypi.python.org/pypi/linvpy`). This is the easiest and most popular way to install and update python packages. On a system where the installer pip is available, running the single command given below will download and install the latest version of LinvPy:

   ```
   > pip install linvpy
   ```

5.2.5 Documentation

The documentation of LinvPy can be accessed online in the ReadTheDocs platform (`http://linpy.readthedocs.io/en/latest/`). This documentation contains:

1. Instructions for obtaining and installing LinvPy.

2. *Quick start* guide, with examples to learn how to use LinvPy.

3. Intuitive tutorial on robust statistics and regularizations.

4. Summary of all the functions contained in the package.

5. Exhaustive documentation of all the classes and methods contained in LinvPy.

6. Instructions for implementing custom functions by the user.

The dynamic HTML version of the documentation hosted in ReadTheDocs is generated using Sphinx [18]. Sphinx is a documentation generator written and extensively used by the Python community. Sphinx converts reStructuredText files [5] into HTML websites, as well as other formats including PDF. Among other functionalities, Sphinx automatically generates documentation from the source code, writes mathematical notation, and highlights source code.

5.2.6 Software development process: Scrum

The process of designing and implementing LinvPy was based on the Scrum methodology [67]. Scrum is an iterative and incremental software development framework for managing product development. It challenges the conventional *water fall* method, a sequential (non-iterative)

Figure 5.5: Screenshot of LinvPy documentation in the ReadTheDocs platform.

process traditionally used in the industry (see Figure 5.6 and 5.7). Scrum introduces flexibility in the product development, allowing goals and objectives to evolve along the way, absorbing what the team learns during the process. It also enables teams to be self-organized and to take their own decisions, and encourages close collaboration among all the team members.
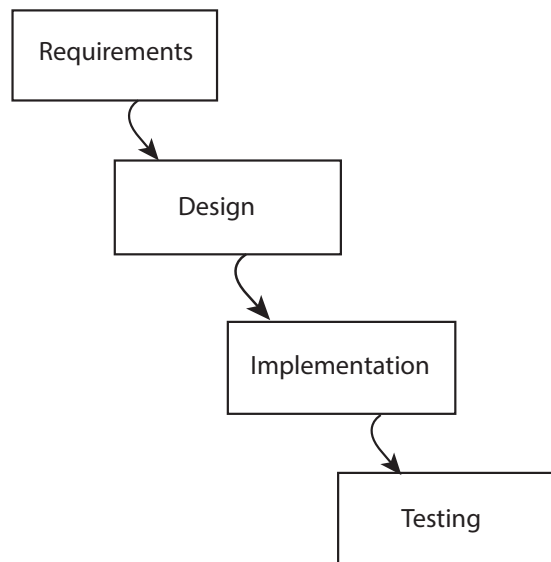

Figure 5.6: Water fall process.

As is usual under Scrum, our main rule was that at every iteration we should have a working prototype of LinvPy. This idea is represented in Figure 5.7: we started building a very simple package with just one method. However this *primitive* package contained all the pieces of the final version: it could be downloaded, installed, it was documented and it was tested. It could already be *used*, though with only limited functionality. In the following iterations, all these pieces evolved together, until we reached the final version. This process allowed us to:

– Learn in a hands-on way. We start the project by learning and implementing the simplest algorithms and the simplest features a package should have. Theory is learned and put into practice immediately and incrementally.

– Be able to deliver immediately a working prototype at any point in the project.

– Obtain and react to feedback from users very early in the process.
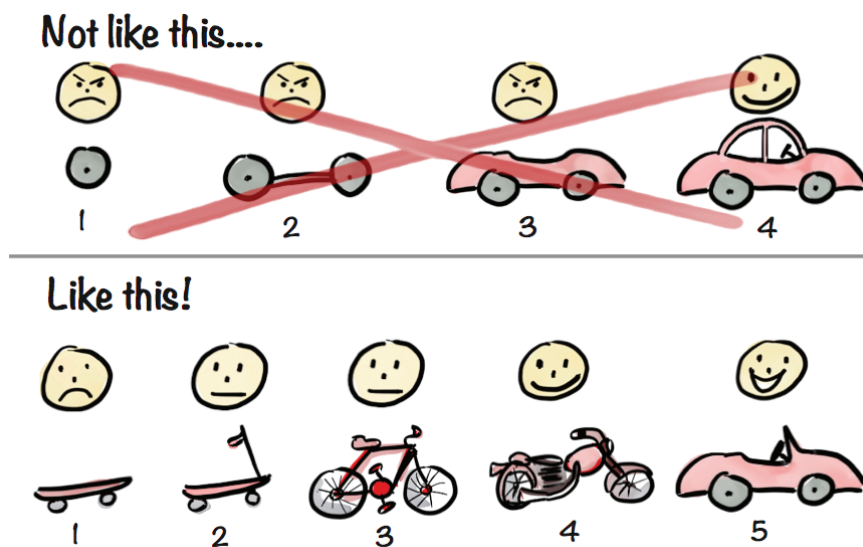
Figure 5.7: Iterative process. At every iteration there is a fully working prototype, from the very beginning. Drawing by Henrik Kniberg.

5.2.7 LinvPy users

LinvPy is already in use both in academia and industry. Colleges from academia informed us that they are using LinvPy successfully in their own projects. In addition, colleages working in industry informed us that they are testing LinvPy on their datasets. However, LinvPy is still too young to have meaningful statistics of usage. Nevertheless, the testimonies from our colleagues make us believe that all the effort invested into building LinvPy is starting to bring rewards beyond just this thesis. Our dream of bringing robust statistics and regularizations closer to people who need them is starting to be a reality.

Chapter 6

What's next: new challenges

As we said in **Chapter 1**, every day tons of pollutants are emitted into the atmosphere all around the world. These pollutants are changing the climate of our planet, and there are cities with concentrations of solid particles in the air so high, that its inhabitants cannot go outdoors without a mask. Knowing how to estimate these emissions is critical to developing and deploying new public policies to limit them. In this thesis we contributed new methods for the estimation of temporal emissions of substances into the atmosphere.

In **Chapter 2** we proposed a new method to estimate the emissions of Xenon (Xe-133) during the accident of the nuclear power plant in Fukushima, Japan, in March 2011. This method is composed of two steps: the cleaning matrix step and the non-negative sparse regularization step. The first step reduces the condition number of the model matrix by discarding the parts of the model matrix that do not add information to the problem. The non-negative regularization introduces a priori information about the source that makes the problem well-posed. The temporal emissions estimated with this method were evaluated using side information: the timeline of the events that took place during the accident, and the inventory of Xe-133 existent in the plant before the accident. The timing of the estimated emissions closely matches the events in the timeline of the accident, and the total estimated emissions are very close to the inventory. These two facts suggest that the estimate that our method produced is reasonable. However, given that the ground truth in this case is unknown, we cannot really assert with precision how close to reality our estimate is.

This takes us to **Chapter 3**, where we worked with real-world data obtained with ETEX, a large-scale controlled experiment. This dataset is one of the very few available controlled experiments of this kind. The ETEX dataset played a key role in this thesis: it allowed us to study the errors in the problem and to evaluate the performance of our methods. Using ETEX we demonstrated the presence of outliers in the data. This made us realize that we need estimators that can handle ill-conditioned problems and at the same time are robust to outliers in the data. Existing classical estimators such as least squares can handle ill-posed problems using regularizations, but are far too sensitive to outliers. To make these estimators robust to outliers, we added a preprocessing step. This step localizes outliers blindly and remove them from the dataset. TRANSAC is the best version of this preprocessing steps, and was developed in this thesis. We showed that adding TRANSAC to existing estimators improves their performance by up to 38%.

As we explained in Chapter 1, usually there is a lack of real-world measurements to solve

the problems we are working on. Thus, what we did in Chapter 3, to discard measurements to improve the estimate, seems counter-intuitive. This is what takes us to **Chapter 4**, where we developed estimators that are robust against outliers. The problem with existing robust estimators such as the τ estimator is that they cannot handle ill-conditioned problems. To make the τ estimator suitable for ill-conditioned problems, we created the regularized τ estimator. To compute this estimator we designed two different algorithms. We also derived the analytic expression of the influence function of our proposed estimator. We also computed other metrics of robustness such as the sensitivity curve and the break-down point. The regularized τ estimator can handle up to 40% of additive outliers in the data even if the problem is ill-posed. In addition, it is also efficient, i.e. when there are no outliers in the data and the errors are Gaussian, it has a performance close to that of least squares. Figure 6.1 shows a summary of the properties of different estimators. The estimators developed in this thesis are in the rows with a grey background.

Linvpy, the Python package described in **Chapter 5**, contains an implementation of the regularized τ estimator, as well as other more traditional robust estimators such as the M. The goal of Linvpy is to make our estimators available and easy to use for other researchers and engineers. Linvpy is open source, easy to use, well documented, and modular, making it easy for other developers to contribute to the package with new estimators.

	Challenge 1 Ill-conditioned	Challenge 2 Outliers	Efficiency (Gaussian errors)
LS	☹	☹	☺
Regularized LS	☺	☹	☺
TRANSAC	☺	☺	☺
M	☹	☺	☹
Regularized M	☺	☺	☹
Tau	☹	☺	☺
Regularized Tau	☺	☺	☺

Figure 6.1: Summary of the properties of different estimators. The rows with a grey background indicate the new estimators we designed in the thesis.

6.1 Looking to the future

Although important progress has been made, still there are important challenges that need to be addressed. Here we give a list of open problems that we hope will help researchers find new paths and directions of work.

Modelling errors in the model

In this thesis, we did not assume anything about the errors in the model. The reason why we did that is that, in fact, we do not know anything about the nature of these errors. However, it would be a great advantage to know how to model them. Then we could design estimators that use this information to produce a better estimate of the emissions.

Discovering the nature of the errors in the model in a large scale problem that uses real-world data is quite a challenge: Lagrangian dispersion models need hundreds of parameters as inputs. Most of these parameters come from forecast models, which also contain their own (unknown) errors. Instead of attacking directly the large-scale real-world problem, we can simplify it, and explore this simpler problem first. The idea is to gain intuition and to understand the core of the problem. Emily Hentgen started to work on this idea during her master semester project. She developed a simple Lagrangian dispersion model in which all the particles move with the same random velocities (see Figure 6.2). If we fix the location of the source and of the sensors, the matrix model corresponding with this system can be computed easily. Then we can study what happens when some of the velocities of these particles are wrong. How do these errors translate into the model matrix? Is there any particular characteristic in the errors of the matrix? Do they have any structure? Do they depend on the source/sensors locations? How could we model them?
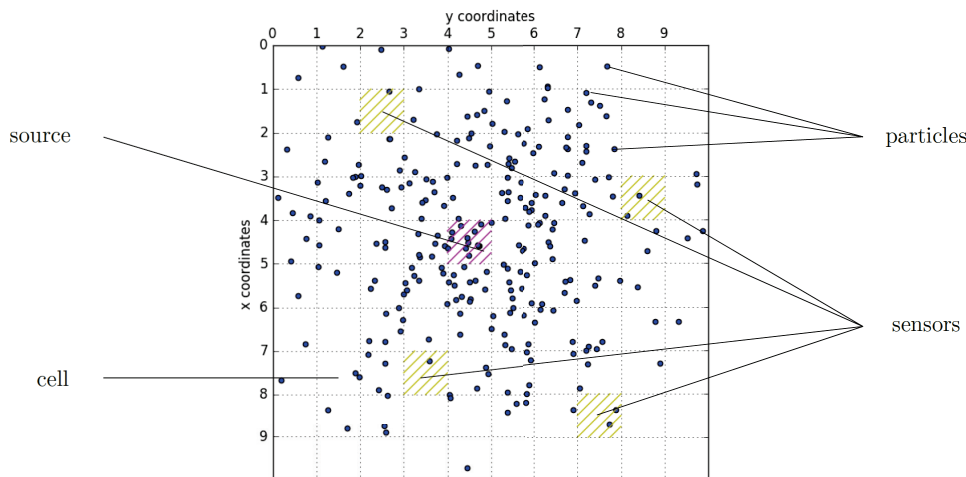

Figure 6.2: Simplified dispersion model designed and implemented by Emily Hentgen. In this system, all the particles have the same pseudo-random velocities.

Regularized τ estimator extended

Robust estimators, including our regularized τ estimator, are robust against additive errors. However they are sensitive to errors in the model, above all if these errors do not have any particular structure. One way of improving the performance of the regularized τ estimator when errors are present in the model is to pre-process the model matrix. The goal of this pre-processing step is to remove the outliers from the model. If the noiseless matrix has a low-dimensional structure, principal components analysis (PCA) can be used for removing the outliers. We assume the outliers do not live in a low-dimensional space, so they will be eliminated after applying PCA to the model matrix. PCA is one of the most basic pre-processing steps we can apply, but others can be explored. Of course for this pre-processing step would be useful to know how to model the errors in the model, as we proposed in the previous point of this list.

More data

Large-scale controlled experiments are expensive and not at all environmentally friendly. Yet these experiments are necessary to learn more about the modelling of the problem and to test the estimators we design. To compensate for the lack of large-scale experiments, well-designed small-scale experiments could be performed in controlled environments such as wind tunnels. Besides controlled experiments, real-world datasets should be openly access to the scientific community. Currently, access to many of the existing datasets is strongly restricted. The more data is available, the faster new research results will become available.

New applications

In this thesis we developed methods for a particular application, but these methods can be applied to other problems which are also ill-conditioned and/or contain outliers in the data. It would be very interesting to observe how these algorithms work in a different framework, and which value they can bring to other fields of research. We hope that the Linvpy package will encourage researchers from different fields to tests our estimators on the problems they are working on. This will help us discover new useful applications for our methods.

Extending the problem

The main theme of this thesis was the estimation of emissions of substances into the atmosphere. We assumed that the emissions are located at a single point, and the location of this point is known. One variation on this theme is to consider that the emissions are located at a single point, but its location is unknown. This variation is relevant for incidents like the one that took place in Europe in January 2017 [1], in which Iodine-131, a radionuclide of anthropogenic origin, was detected in the ground-level atmosphere, but its origin is still unknown. Figure 6.3 shows the levels of Iodine detected by different stations. One way of solving this new variation is to introduce more unknowns into the problem. These unknowns represent the temporal variation of the emissions at each discretized point in the space. Another approach is to develop a two-step method: first the source location is estimated, and then the temporal emissions are estimated. For the last step the methods developed in this thesis can be applied.

Another variation of our main theme is to assume that the emissions are localized at several points, and in addition the locations of these points are unknown. The estimation of CO_2 emissions fits into this category.
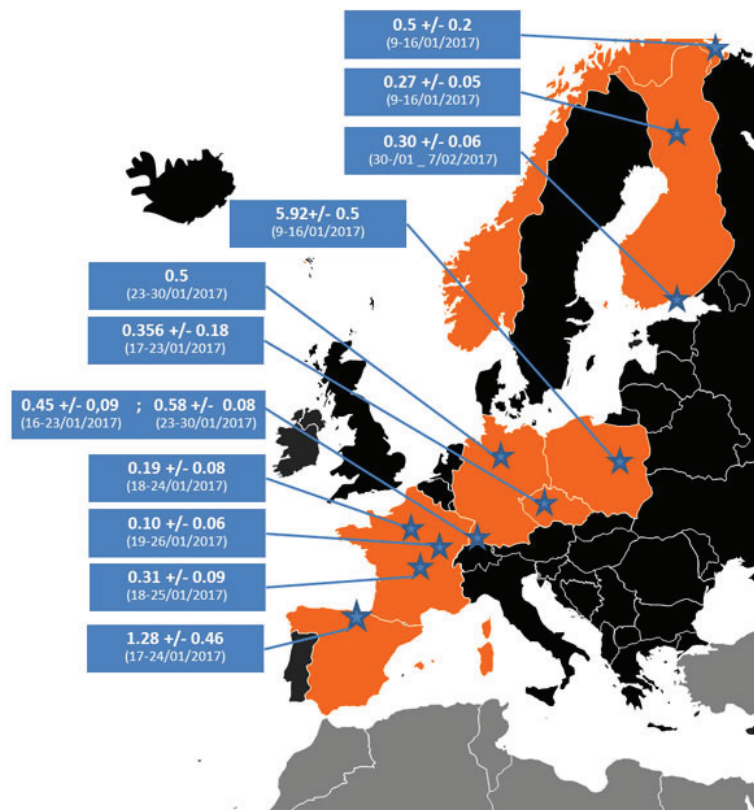
Figure 6.3: Levels of Iodine-131 detected in Europe in January 2017 [1].

Appendix A

Meeting the real world

In this chapter we review my six-month internship in industry. I cannot disclose all the details of the project I worked on because of confidentiality issues. Hence, along this chapter I will provide only as much detail of the business context and my technical contributions as I am allowed to.

Universities and companies are two very different classes of organizations, but they share a common goal: to create value for society. They pursue this goal in different ways: the former provides higher education and research (it does not necessarily create profit), whereas the latter in general develops useful goods and services with focus on generating profit. This difference results in dissimilar dynamics inside these organizations. There were two main reasons for me to do an internship in a company. The first one is to experience these contrasting dynamics, and the second one is to learn how research brings value to a profit-driven company. This internship did not produce any new results directly continuing on the topics presented in the previous chapters of this thesis. Nevertheless, I did participate in the development of a new product that is currently in production and that the clients of the company currently use. In this chapter we give an overview of this development process and my contribution.

A.1 Developing a new product

The goal of my project was to design and implement a system that raises an alarm if a household electrical heating device is not working correctly. The system has access to the instant power consumption of all the heating devices in the system, their locations, and the local weather conditions. The resulting system needs to comply with the following specifications:

1. All failures should be detected less than six hours after they happen. Otherwise, the product is not practically useful, because the house will cool down sufficiently for the customer to notice by him/herself that the heating system is not working properly.

2. Less than 1% of the total heating devices in the system are allowed to produce one or more false alarms in one week. False alarms cause calls to the call center. False alarms also lead to the entire system being perceived as unreliable by the customer. Thus, each false alarm has a cost associated to it. This cost and the consequent maximum number of false alarms that our system is allowed to produce are defined by the business plan of the company.

3. The system should work with more than 6000 different heating devices. These devices vary by manufacturer, installation setup, and of course thermal properties of the house they are installed in.

4. The system should work in all the seasons when the heating systems are in use. In general, this means autumn, winter, and spring.

5. The system should be scalable up to 100'000 devices.

A.1.1   Performance evaluation

To know if our system is currently fulfilling the specifications, we need a method to evaluate its performance. The output of this evaluation system should be easily understandable not only by engineers, but also by the management team.

Key Performance Indicators

To measure the performance of our system, we need to define metrics. These metrics should be designed such that they link the performance of the system with the business plan of the company. The business plan uses key performance indicators (KPIs). A KPI is a measurable value that demonstrates how effectively a company is achieving key business objectives. In our case we need to define KPIs that

1. represent the impact the product has on the business plan,

2. can easily be understood by the engineering and management teams, and

3. are easily measurable.

Thus, we define the following KPIs to measure the performance of our system:

1. **Devices with one or more false alarms over the last one week period (%):** Percentage of all the heating devices where the system produced one false alarm or more in the period of one week.

2. **Average delay of detection (hours):** Averaged time over all the heating devices that the system took to detect failures.

3. **Devices with false negatives (%):** Percentage of all the heating devices in which an actual failure was not detected.

4. **Undefined outputs (%):** Percentage of the total outputs of the detection system that are undefined, i.e. the system cannot give any specific answer. This is mostly due to missing data.

Test bench

The KPIs should be computed in a systematic and fast way. Every time we make a modification in the algorithm or its implementation, we must be able to evaluate the impact of that modification on the overall performance of the system. Hence, we built a test bench. Its structure is represented in Figure A.1. The test bench works as follows:

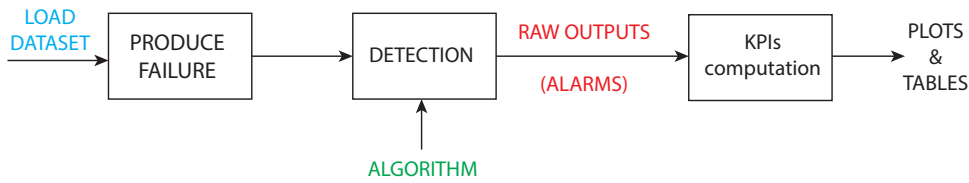1. Load the consumption data of 100 devices during 2 weeks.

Figure A.1: Block diagram of the test bench.

2. Produce an artificial failure in the consumption data of every device. We do that by setting the consumption of all the devices to zero during 24 hours.

3. Run the algorithm under test using the above dataset with artificial failures.

4. Get the raw output from the algorithm, i.e. alarms over time for all the devices.

5. Compute the KPIs defined above.

A.1.2 Algorithms

The test bench helped us compare the different algorithms we created, in terms of the relevant KPIs, as well as select the most promising one. What follows is an overview of all the algorithms we created, and the advantages and challenges of each one.

Threshold algorithm

This is the simplest algorithm to detect failures that we could think of. Very often the simplest solution is the one that performs best with real data. In addition, it is easy to implement and easy to scale. This is why the threshold algorithm is a good point of reference.

Description: For every heating device, we measure the time since the last power consumption. We call this time t_{off}. As soon as t_{off} is greater than a certain threshold T, the system raises an alarm. There are several strategies to compute the threshold T:

1. Static threshold: the threshold does not change over a few days. For example, T is recomputed every day using data from the last few days.

2. Dynamic threshold: the threshold continuously changes over time. For example, T is recomputed every second using a moving window.

Challenges: The difficulty with using this algorithm is to find a proper threshold T that produces a good trade off between the average delay of detection and the false alarm KPI. Using just consumption data from the last few days to compute the threshold is not enough to achieve satisfactory performance.

House modelling algorithm

The power consumption of a heating device is tightly linked to the interior temperature of the house where it is installed. If the heating device fails, the house gets cold. Unfortunately, we do not have access to the interior temperature of each house. Thus, instead, we model this

---

Algorithm A.1 Threshold algorithm

$t_{\text{off}} = 0$
for each second s **do**
    Get consumption $c[s]$
    # Update t_{off}
    if $c[s] > 0$ **then**
        $t_{\text{off}} = 0$
    else
        $t_{\text{off}} = t_{\text{off}} + 1$
    end if
    # Check if the device was off for too long
    if $t_{\text{off}} > T$ **then**
        Raise an alarm
    end if
end for

---

interior temperature using the power consumption of the heating device and the external weather conditions of the house.

Description: Using thermodynamics, external weather conditions and power consumption of the heating device, we set up a simple model for the interior temperature of a house. If the heating system is working correctly, we assume the interior temperature in the house should be over 20 degrees. When the modelled temperature drops one degree under this threshold, we raise an alarm.

Algorithm A.2 House modelling algorithm

$T = 20$
for each second s **do**
    Update weather conditions $w[s]$
    Update consumption $c[s]$
    prediction model $(w[s], c[s]) \rightarrow t_{\text{int}}[s]$
    if $t_{\text{int}}[s] > (T - 1)$ **then**
        Raise an alarm
    end if
end for

Challenges: We use the same model for all the houses. This model captures the general behaviour of an average house, but not the particularities of each one. The thermodynamic behaviour of a house depends also on factors such as the number of rooms, number of windows, thermal isolation, or orientations towards the sun. We do not have access to this data and, thus, we can only use a very generic model.

Signal processing algorithm

The power consumption of a heating device has a 24 hour periodicity. This periodicity appears because the consumption of a device depends on the external temperature and sun radiation on

the house. Both of these factors have a daily cycle (see Figure A.2). If the consumption signal is filtered with a low-pass filter, its periodicity becomes evident (see Figure A.3). However, if there is a failure in the heating device, the periodicity of its consumption signal will be destroyed. We use this *disruption* of the periodicity to raise an alarm.
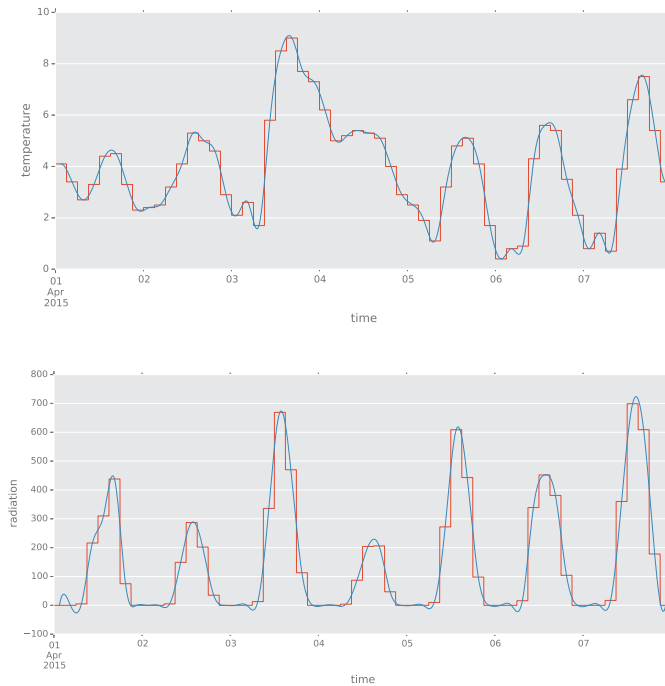

Figure A.2: Forecasted temperature (upper plot) and radiation (lower plot) over 7 days. The red line shows the original data, which is given every 3 hours. The blue line shows the signal smoothed with a polynomial interpolation of order 3.

Description: We would like to observe the frequency content of different signals over time. For that we use the Discrete Fourier Transform (DFT). In order to obtain sufficient frequency resolution to observe the lowest frequency corresponding to the daily cycle, we need to use a total of around 70 hours of real data (this was found experimentally). However, we also need to detect a failure in a heating device in less than six hours: we cannot wait 70 hours to detect a failure. To solve this problem, we compute the DFT of the consumption signal every hour, using the last 70 hours of data. This is in fact a Short Time Fourier Transform (STFT) that uses a 70-hour rectangular window and an overlap of 98.5%.

We compute the STFT of a consumption signal without a failure and compare it with the STFT of a signal with a failure. We can see the results in Figure A.4. The upper part of the figure shows the spectrogram of a signal without a failure, the bottom part of the figure shows the spectrogram of a signal with a failure at the end. The difference between both spectrograms is clear. The failure leaves a footprint in the STFT: the magnitude of low frequencies increases when there is a failure. We should exploit this effect to detect the failure as soon as possible. One idea is the following: if the low frequencies of the STFT raise over a threshold, we raise an
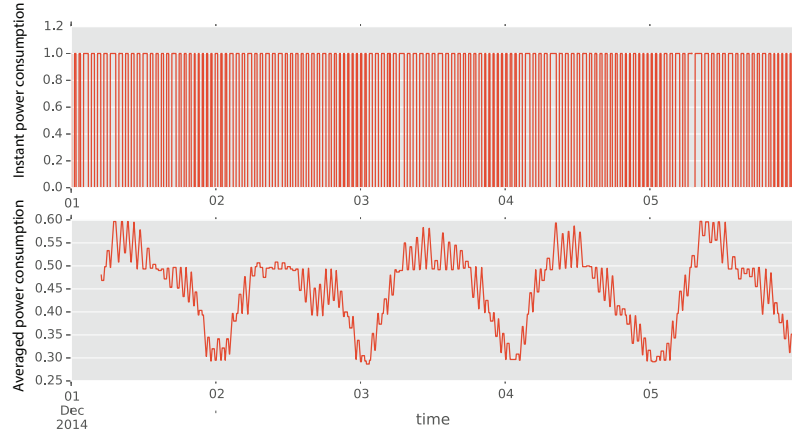
Figure A.3: The plot in the upper part of the figure shows an example of instant power consumption of a device over 5 days. The plot in the bottom part of the figure shows the same signal, but filtered with a moving average filter.

alarm. Figure A.5 shows this idea. It shows the magnitude of the frequency bin corresponding to low frequencies in the STFT. The red area indicates that a failure is produced. Soon after the beginning of the failure, the magnitude of this frequency bin rises.

To understand why the low frequencies rise when there is failure, we create a simplified model of a consumption signal. We assume that a normal consumption signal $s[n]$, without a failure, is a sinusoid with a 24-hour period with an offset. A signal $x[n]$ that contains a failure is modelled as the multiplication of the signal $s[n]$ and a box function $b[n]$. If N is the length of the DFT, and M is the number of samples during which the device is working properly, $b[n]$ is defined as

$$b[n] = \begin{cases} 1 & \text{for } n \leq M - 1 \\ 0 & \text{for } M \leq n \leq N - 1. \end{cases} \tag{A.1}$$

The points where $b[n]$ is 0 represent the failure. Thus, we can write

$$x[n] = s[n] \cdot b[n]. \tag{A.2}$$

The DFT of $s[n]$, $S[k]$, is a set of deltas located at the zero frequency and at frequency of the sinusoid and its multiples. The DFT of $b[n]$, $B[k]$, is the *sinc* function

$$B[k] = \begin{cases} M & \text{if } k = 0, \\ \frac{\sin(\frac{\pi}{N}Mk)}{\sin(\frac{\pi}{N}k)} e^{-j\frac{\pi}{N}(M-1)k} & \text{otherwise.} \end{cases} \tag{A.3}$$

Given this, we can compute $X[k]$, the DFT of $x[n]$, as

$$X[k] = S[k] * B[k]. \tag{A.4}$$

If we observe the DFT of a real consumption signal without a failure like the one plotted in the lowest part of Figure A.4, we notice that the delta located at zero has an amplitude several times larger than the rest. The side lobes of the sinc convolved with this delta will also dominate the
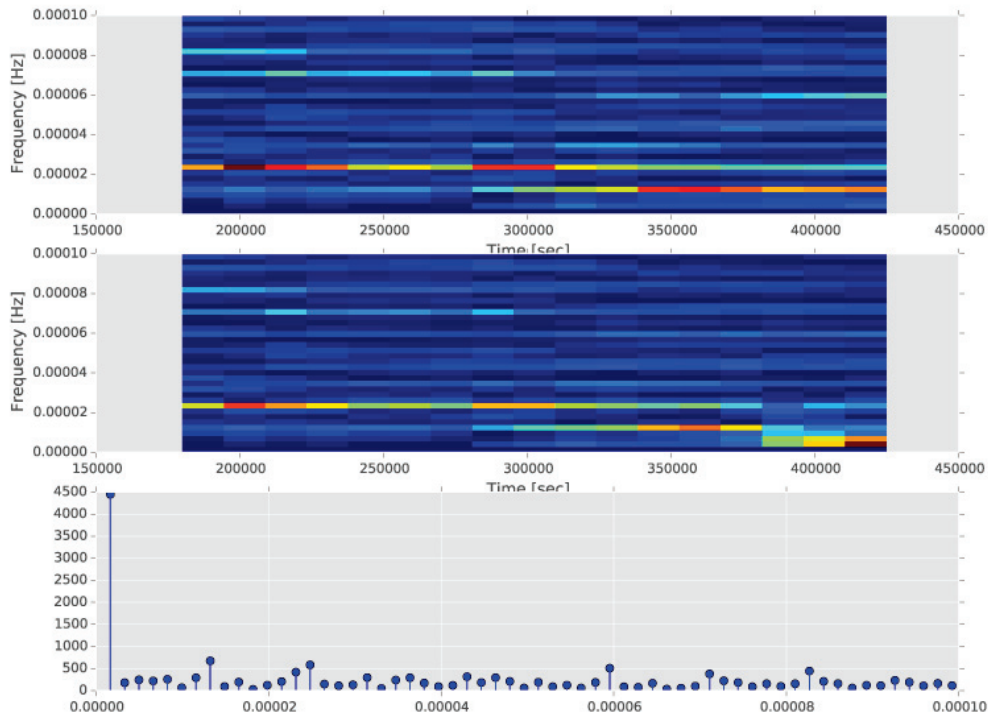
Figure A.4: Spectrogram of a normal signal, spectrogram of a signal with a failure at the end, and the DFT of a normal signal (without failure).
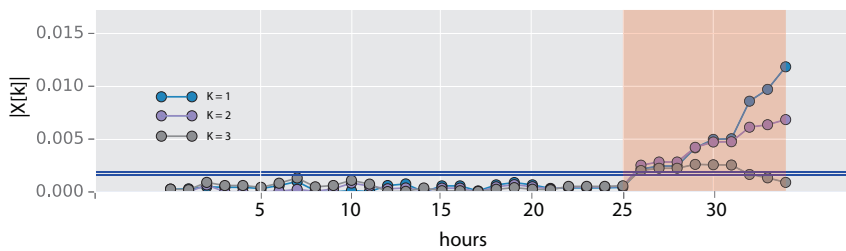
Figure A.5: Evolution of the lowest three samples of the DFT over time. Every hour the DFT is recomputed. The red area indicates when the failure occurs. When the failure happens, these lowest frequencies rise. When they go over a certain threshold, we raise an alarm. The thresholds are computed using data from the past few days.

rest at low frequencies. Thus we can approximate the DFT for these low frequencies using only this dominating sinc and ignoring the rest. When we observe that the magnitude of the STFT

rises if there is a failure, in fact what we are seeing are the lobes of this sinc. With this simplified model, we simulate the low frequencies of the DFT of a consumption signal with a failure and compare it with real data with the same failure. The results are shown in Figure A.6. The shadowed part in the plots indicates the failure. We can appreciate that, despite its simplicity, this model describes reality quite well.

We used these side-lobes of the sinc located at zero to detect a failure. We designed an algorithm with the following steps:

1. Every hour, compute the DFT of the consumption signal over the last 70 hours.

2. Track the magnitude the DFT associated with low frequencies. If the heating device is working correctly, these magnitudes should be under a certain threshold. If two or more of these magnitudes go over the threshold, then raise an alarm.

Algorithm A.3 Signal processing algorithm

Set T_1, T_2, T_3
for each hour h **do**
    Compute DFT $X[k]$
    if $|X[1]| > T_1$ and $|X[2]| > T_2$ and $|X[3]| > T_3$ **then**
        Raise an alarm
    end if
end for

Difficulties: There are cases where our model does not represent reality very well. Consumption signals are not perfectly periodic over time. The magnitude of the DFT corresponding to low frequencies do not only rise because of failures. Abrupt changes in the external conditions also cause changes in the periodicity of the power consumption signal. To discriminate between a real failure and an unrelated change of behaviour in the system, we need to set higher thresholds, which means to increase the delay of detection. By doing this, we reduce the number of false alarms, but the detection of failures is too slow for our requirements.

Machine learning algorithm

The main idea behind this algorithm is to predict the consumption signal using external conditions and assuming that the device is working correctly. For that we use a parametric model. Next, we compare the real consumption signal with our predicted one. If the real consumption is significantly smaller than the predicted one, we assume the heating device is not working properly and we raise an alarm.

Difficulties: Predicting consumption is not always easy. The consumption of a heating device depends on many external factors, such as the number of people inside the house, because people produce heat. If a house is full of people, the internal temperature of the house rises and at some point the heating device stops working altogether, though it is not faulty. Unfortunately, our system cannot predict when someone is having a party at home! In fact, any exceptional behaviour, different from the routine, is almost impossible to predict. This fact means our machine learning algorithm may produce many false alarms.

The performance of the algorithms proposed above was compared using the test bench. The only algorithm that satisfied all the requirements was the machine learning algorithm. After developing a first prototype, we rolled out into production a complete detection system that uses
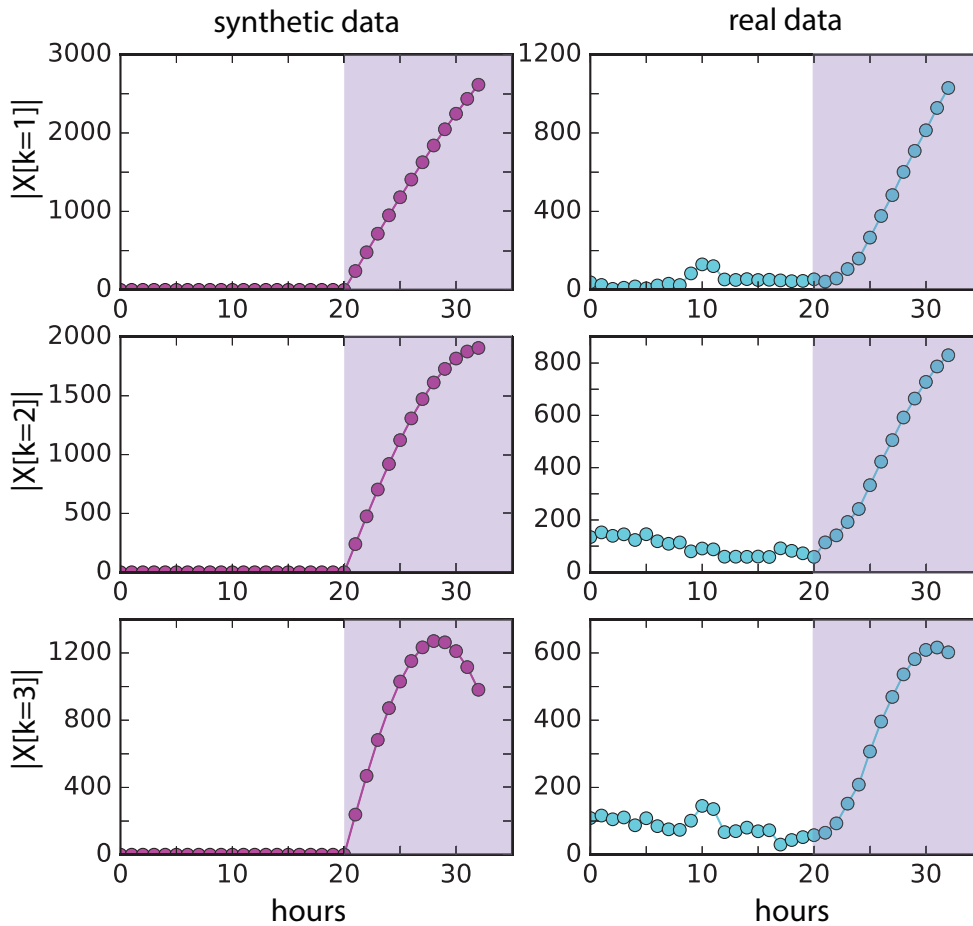
Figure A.6: Evolution over time of the absolute value of the first three components of the DFT of a consumption signal. A failure occurs at hours = 20. The shadowed area represents the failure in the device. The results with synthetic data are shown on the left, while the results with real data are shown on the right.

this algorithm. Currently, the system we designed and implemented is running in the production server of the company. Accordingly to the feedback from the management team, it is one of the company's most innovative products.

A.2   New skills

Participation in the development of a product broadened my set of skills. These are the main lessons I learned:

1. **Playing jazz.** Priorities in a company and at the university are different: research in a company needs to create specific value for its clients, whereas research at the university creates fundamental knowledge whose concrete value may not necessarily be immediately obvious. Hence, in a company, the top priority is that systems designed and implemented must *work* as expected at all times. *How* a system is implemented and *why* it works have only secondary importance. At the university, on the other hand, these priorities are inverted.

   A similar dyad appears also in music. Two different styles, jazz and classical, require different ways of playing from a musician. Making this change of mindset requires open mindedness, hard-work, and humbleness, but it brings new richness to our repertoire and augments our capacity to understand a larger set of people. Linvpy (see Chapter 5) was born from applying this new industrial perspective to my university work.

2. **Team work.** Usually in a small company like a start-up, people are organized in highly interdisciplinary teams. The goal of these teams is to solve problems in a collaborative manner. It does not matter where the ideas come from, or who does what, or who made a mistake. The different skills from the team members complement each other, and the output of the team tends to be larger than the sum of the outputs of the individual members. Several management techniques have been developed for the sole purpose of setting up this constructive collaboration environment. These techniques give general guidance on how a team should be organized to work in a constructive manner. Scrum, which we used in Chapter 5, is one of them.

3. **Software development.** In a company, software implementations are usually written collaboratively by many developers. The final goal of these implementations is to form a part of a product and in some cases to run on the company's production server. There exist processes to assure good code quality, as well as code writing in a collaborative way without conflicts. One such process is the concept of code reviews. These are systematic examinations of the newly produced code. They are done by one or more team members that are not the author(s) of the code. These reviews are intended to find mistakes in the initial development phase. In addition, they have the following effects:

   (a) **A different way of writing code.** The author of the code changes his/her way of writing code if he/she knows someone other than a computer will read it. This produces code that is clearer, well-commented, and simpler.

   (b) **Teaching.** Code reviews are very useful to teach someone to code relatively quickly.

   (c) **Knowledge exchange.** A code review is also an exchange of ideas and tips between team members.

   (d) **Shared responsibilities.** Code reviews distribute responsibility among the team members. If, even after the review, there is still a mistake in the code, this mistake is not just the responsibility of the original author(s) of that code: it is the responsibility of the whole team. This shared responsibility decreases the fear of making mistakes among the team members, and boosts creativity.

   Code reviews were a key part of the development process of LinvPy.

Bibliography

[1] Detection of radioactive iodine at trace levels in Europe in January 2017 . [Online]. Available: http://www.irsn.fr/EN/newsroom/News/Pages/20170213_Detection-of-radioactive-iodine-at-trace-levels-in-Europe-in-January-2017.aspx.

[2] ECMWF public datasets. [Online]. Available: http://apps.ecmwf.int/datasets/.

[3] ECMWF re-analysis. [Online]. Available: https://en.wikipedia.org/wiki/ECMWF_re-analysis.

[4] Python. [Online]. Available: http://en.wikipedia.org/wiki/Python%20%28programming%20language%29.

[5] reStructuredText. [Online]. Available: http://docutils.sourceforge.net/rst.html.

[6] Safecast. [Online]. Available: http://www.safecast.org.

[7] "How Science Goes Wrong," *The Economist*, Oct. 2013.

[8] "Trump and the Truth: Climate-Change Denial ," *The New Yorker*, Oct. 2016.

[9] A. Allisy, "From the Curie to the Becquerel," *Metrologia*, vol. 31, no. 6, pp. 467–479, Feb. 2005.

[10] I. Andreev, M. Hittenberger, P. Hofer, H. Kromp-Kolb, W. Kromp, P. Seibert, and G. Wotawa, "Risks Due to Beyond Design Base Accidents of Nuclear Power Plants in Europe—the Methodology of Riskmap," *Journal of Hazardous Materials*, vol. 61, no. 1-3, pp. 257–262, Aug. 1998.

[11] D. F. Andrews, F. R. Hampel, P. J. Bickel, P. J. Huber, W. H. Rogers, and J. W. Tukey, *Robust Estimates of Location: Survey and Advances.* Princeton: Princeton University Press, 1972.

[12] G. E. Backus and J. F. Gilbert, "Numerical Applications of a Formalism for Geophysical Inverse Problems," *Geophysical Journal International*, vol. 13, no. 1-3, pp. 247–276, July 1967.

[13] G. Beaud and M. Martinez-Camara. LinvPy. [Online]. Available: http://linvpy.readthedocs.io/en/latest/.

[14] A. Beck and M. Teboulle, "Gradient-Based Algorithms with Applications to Signal Recovery," in *Convex Optimization in Signal Processing and Communications*, D. P. Palomar and Y. C. Eldar, eds., 2009.

[15] A. Becker, G. Wotawa, A. Ringbom, and P. R. J. Saey, "Backtracking of Noble Gas Measurements Taken in the Aftermath of the Announced October 2006 Event in North Korea by Means of PTS Methods in Nuclear Source Estimation and Reconstruction," *Pure and Applied Geophysics*, vol. 167, no. 4-5, pp. 581–599, 2010.

[16] G. Box, "Robustness in the Strategy of Scientific Model Building," *Robustness in statistics*, 1979.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Mar. 2004.

[18] G. Brandl. Sphinx. [Online]. Available: http://www.sphinx-doc.org/en/1.5.1/.

[19] D. R. Brillinger, "John W. Tukey: His Life and Professional Contributions," *Annals of Statistics*, vol. 30, no. 6, pp. 1535–1575, 2002.

[20] J. C. Brown, "Overview of Topical Issue on Inverse Problems in Astronomy," *Inverse Problems*, vol. 11, no. 4, pp. 635–638, Jan. 1999.

[21] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images." *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.

[22] G. C. Calafiore and L. El Ghaoui, *Optimization Models*. Cambridge University Press, 2014.

[23] A. Cazenave, "How Fast Are the Ice Sheets Melting?" *Science*, vol. 314, no. 5803, pp. 1250–1252, Nov. 2006.

[24] Y. Chen, A. Ebenstein, M. Greenstone, and H. Li, "Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy," *Proceedings of the National Academy of Sciences*, vol. 110, no. 32, pp. 12 936–12 941, Aug. 2013.

[25] M. Chino, H. Nakayama, H. Nagai, H. Terada, G. Katata, and H. Yamazawa, "Preliminary Estimation of Release Amounts of 131I and 137Cs Accidentally Discharged from the Fukushima Daiichi Nuclear Power Plant into the Atmosphere," *J. Nucl. Sci. Technol.*, vol. 48, no. 7, pp. 1129–1134, 2011.

[26] R. D. Cook, "Detection of Influential Observation in Linear Regression," *Technometrics*, vol. 19, pp. 15–18, 1977.

[27] CTBTO, "Preparatory Commision for the Comprehensive Nuclear-Test-Ban Treaty Organization," Tech. Rep., 2014.

[28] CTBTO Preparatory Commision, *The 11 March Japan Disaster*. http://www.ctbto.org/verification-regime/the-11-march-japan-disaster/, 2011.

[29] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Communications on Pure and Applied Mathematics*, vol. 63, no. 1, pp. 1–38, Jan. 2010.

[30] D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. Van De Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. Mcnally, B. M. Monge Sanz, J. J. Morcrette, B. K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J. N. Thépaut, and F. Vitart, "The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System," *Quarterly Journal of the Royal Meteorological Society*, vol. 137, no. 656, pp. 553–597, Apr. 2011.

[31] D. L. Donoho and J. Tanner, "Sparse Non-Negative Solution of Underdetermined Linear Equations by Linear Programming," *Proceedings of the National Academy of Sciences*, vol. 102, no. 27, pp. 9446–9451, July 2005.

[32] D. L. Donoho and P. J. Huber, "The Notion of Breakdown Point ," in *A Festschrift For Erich L. Lehmann*, P. J. Bickel, K. Doksum, and J. L. Hodges, eds., California, 1983.

[33] G. J. Ferber, J. L. Heffter, and R. R. Draxler, "Cross-Appalachian Tracer Experiment (CAPTEX 83) Final Report," *NOAA Tech Memo*, 1986.

[34] L. Goering. Data App Pushes Chinese Factories to Cut Pollution. [Online]. Available: http://www.reuters.com/article/us-china-pollution-apps-idUSKBN0N81MB20150417.

[35] M. Grant and S. Boyd. (2014) Matlab Software for Disciplined Convex Programming, version 2.1 . [Online]. Available: http://cvxr.com/cvx.

[36] J. Hadamard, "Sur les Problemes aux Derives Partielles et leur Signification Physique," *Princeton University Bulleting*, vol. 13, pp. 49–52, 1902.

[37] F. R. Hampel, "The Influence Curve and its Role in Robust Estimation," *Journal of the American statistical association*, vol. 69, no. 346, pp. 383–393, June 1974.

[38] J. Hansen, "Earth's Energy Imbalance: Confirmation and Implications," *Science*, vol. 308, no. 5727, pp. 1431–1435, June 2005.

[39] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least-squares," *Communications in Statistics - Theory and Methods*, vol. 6, no. 9, pp. 813–827, June 2007.

[40] P. J. Huber, "John W. Tukey contributions to robust statistics," vol. 30, no. 6, pp. 1640–1648, 2002.

[41] P. J. Huber and E. M. Rochetti, *Robust Statistics*. John Willey and Sons, 2009, vol. 2.

[42] International Atomic Energy Agency, *The International Nuclear and Radiological Event Scale*. Vienna (Austria): International Atomic Energy Agency, 2008.

[43] H. Li and Z. Lin, "Accelerated Proximal Gradient Methods for Non-Convex Programming," in Proc. *NIPS'15: Proceedings of the 28th International Conference on Neural Information Processing Systems*, Peking University. MIT Press, Dec. 2015.

[44] R. A. Maronna, R. D. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods*. John Wiley and Sons, 2006.

[45] M. Martinez-Camara, B. Béjar Haro, A. Stohl, and M. Vetterli, "A robust method for inverse transport modeling of atmospheric emissions using blind outlier detection," *Geoscientific Model Development*, vol. 7, no. 5, pp. 2303–2311, 2014.

[46] M. Martinez-Camara, I. Dokmanic, J. Ranieri, R. Scheibler, M. Vetterli, and A. Stohl, "The Fukushima inverse problem," in Proc. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4330–4334, Vancouver, 2013.

[47] M. Martinez-Camara, M. Muma, B. Bejar, A. M. Zoubir, and M. Vetterli, "The regularized tau estimator: A robust and efficient solution to ill-posed linear inverse problems," *arXiv.org*, May 2016.

[48] M. Martinez-Camara, M. Muma, A. M. Zoubir, and M. Vetterli, "A New Robust and Efficient Estimator for Ill-Conditioned Linear Inverse Problems with Outliers," in Proc. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3422–3426. Brisbane: IEEE, 2015.

[49] M. Martinez-Camara, M. Vetterli, and A. Stohl, "Outlier removal for improved source estimation in atmospheric inverse problems," in Proc. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6820–6824. Florence: IEEE, 2014.

[50] Y. Nesterov, "A Method of Solving a Convex Programming Problem with Convergence Rate O (1 / k2)," *Soviet Mathematics Doklady*, vol. 27, pp. 372–376, 1983.

[51] K. Nodop, R. Connolly, and F. Giraldi, "The Field Campaigns of the European Tracer Experiment (ETEX): Overview and Results," *Atmospheric Environment*, vol. 32, pp. 4095–4108, 1998.

[52] V. Öllerer, C. Croux, and A. Alfons, "The Influence Function of Penalized Regression Estimators," *Journal of Theoretical and Applied Statistics*, vol. 49, no. 4, pp. 741–765, 2015.

[53] N. Parikh and S. Boyd, "Proximal Algorithms." *Foundations and Trends in optimization*, vol. 1, no. 3, pp. 127–239, 2014.

[54] K. Pearson, "Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 187, no. 0, pp. 253–318, Jan. 1896.

[55] J. Pudykiewicz, "Simulation of the Chernobyl Dispersion with a 3-D Hemispheric Tracer Model," *Tellus*, vol. 41B, no. 4, pp. 391–412, Sept. 1989.

[56] V. Ramanathan and Y. Feng, "On Avoiding Dangerous Anthropogenic Interference with the Climate System: Formidable Challenges Ahead," *Proceedings of the National Academy of Sciences*, vol. 105, no. 38, pp. 14 245–14 250, Sept. 2008.

[57] J. Ranieri, I. Dokmanic, A. Chebira, and M. Vetterli, "Sampling and reconstruction of time-varying atmospheric emissions," in Proc. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3673–3676. Kyoto: IEEE, 2012.

[58] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. New York, NY, USA: John Wiley and Sons, 1987.

[59] R. E. Rowland, *Radium in Humans: a Review of US Studies.* Argonne National Lab., IL (United States), 1994.

[60] M. Salibian-Barrera, G. Willems, and R. H. Zamar, "The Fast-Tau Estimator for Regression," *Journal of Computational and Graphical Statistics*, vol. 17, pp. 659–682, 2008.

[61] P. Seibert, "Inverse Modelling with a Lagrangian Particle Dispersion Model: Application to Point Releases over Limited Time Intervals," *Air Pollution Modeling and Its Application XI*, pp. 381–389, 2001.

[62] C. V. Stewart, "Robust Parameter Estimation in Computer Vision," *SIAM Review*, vol. 41, no. 3, pp. 513–537, 1999.

[63] A. Stohl, C. Forster, A. Frank, P. Seibert, and G. Wotawa, "Technical Note: The Lagrangian Particle Dispersion Model FLEXPART Version 6.2," *Atmospheric Chemistry and Physics*, vol. 5, no. 9, pp. 2461–2474, Sept. 2005.

[64] A. Stohl, M. Hittenberger, and G. Wotawa, "Validation of the Lagrangian Particle Dispersion Model FLEXPART Against Large-Scale Tracer Experiment Data," *Atmospheric Environment*, vol. 32, pp. 4245–4264, 1998.

[65] A. Stohl, P. Seibert, G. Wotawa, D. Arnold, J. F. Burkhart, S. Eckhardt, C. Tapia, A. Vargas, and T. J. Yasunari, "Xenon-133 and Caesium-137 Releases into the Atmosphere from the Fukushima Dai-ichi Nuclear Power Plant: Determination of the Source Term, Atmospheric Dispersion, and Deposition," *Atmospheric Chemistry and Physics*, vol. 12, no. 5, pp. 2313–2343, 2012.

[66] G. Strang, *Introduction to Linear Algebra.* Wellesley-Cambridge Press, 2011.

[67] J. Sutherland and J. J. Sutherland, *Scrum,* The Art of Doing Twice the Work in Half the Time. Crown Business, Sept. 2014.

[68] K. Tharmaratnam, G. Claeskens, C. Croux, and M. Salibian-Barrera, "S-Estimation for Penalised Regression Splines," *Journal of Computational and Graphical Statistics*, vol. 19, pp. 609–625, 2010.

[69] J. W. Tukey, "A Survey of Sampling from Contaminated Distributions," in *Contributions to Probability and Statistics*, I. Olkin, ed., Stanford, CA, 1960, pp. 448–485.

[70] S. M. Uppala, P. W. KÅllberg, A. J. Simmons, U. Andrae, V. D. C. Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. V. D. Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. Mcnally, J. F. Mahfouf, J. J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen, "The ERA-40 Re-Analysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 612, pp. 2961–3012, 2006.

[71] P. Vandewalle, J. Kovacevic, and M. Vetterli, "Reproducible Research in Signal Processing," *IEEE Signal Processing Magazine*, vol. 26, no. 3, pp. 37–47, May 2009.

[72] M. Vetterli, J. Kovacevic, and V. K. Goyal, *Foundations of Signal Processing*. Cambridge University Press, Sept. 2014.

[73] B. Wernsperger and C. Schlosser, "Noble Gas Monitoring within the International Monitoring System of the Comprehensive Nuclear Test-Ban Treaty," *Radiation Physics and Chemistry*, vol. 71, no. 3-4, pp. 775–779, Oct. 2004.

[74] G. Wotawa, A. Becker, M. Kalinowski, P. Saey, M. Tuma, and M. Zähringer, "Computation and Analysis of the Global Distribution of the Radioxenon Isotope 133Xe based on Emissions from Nuclear Power Plants and Radioisotope Production Facilities and its Relevance for the Verification of the Nuclear-Test-Ban Treaty," *Pure and Applied Geophysics*, vol. 167, no. 4-5, pp. 541–557, 2010.

[75] C. Wunsch, "The Steady Ocean Circulation Inverse Problem," in *The Ocean Circulation Inverse Problem*. Cambridge: Cambridge University Press, 2009, pp. 212–296.

[76] V. J. Yohai and R. H. Zamar, "High Breakdown-Point Estimates of Regression by Means of the Minimization of an Efficient Scale," *Journal of the American statistical association*, vol. 83, no. 402, pp. 406–413, 1988.

[77] P. Zannetti, "Eulerian Dispersion Models," in *Air Pollution Modelling*. Boston, MA: Springer US, 1990, pp. 107–139.

[78] ——, "Lagrangian Dispersion Models," in *Air Pollution Modelling*. Boston, MA: Springer US, 1990, pp. 185–222.

Marta Martinez-Camara

📞 +41789096581

http://lcav.epfl.ch/people/
marta_martinez

✉ *marta.martinez-camara@epfl.ch*

Personal information

Sex	female
Born	04 April 1986 in Burgos, Spain
Citizenship	Spanish
Languages	Spanish (native), English (fluent), French (medium), German (basic)

Education

2011–present — **Computer, Communication and Information Sciences (Ph.D.)**, *École Polytecnique Fédérale de Lausanne*, Switzerland, Advisor: Prof. Martin Vetterli.
Applying signal processing to inverse problems in environmental sciences. Developing robust methods to estimate emissions into the atmosphere: Fukushima accident, emissions of CO_2 and other greenhouse gases.

2010–2011 — **Communications Technologies and Systems (M.Sc.)**, *Universidad Politécnica de Madrid*, Spain.
Mathematical methods for signal processing, optimization techniques, computer vision, artificial intelligence.

2004–2010 — **Telecommunication Engineering (B.E./M.Sc.)**, *Universidad Politécnica de Madrid*, Spain.
Design and implementation of a high-quality music synthesizer based on SoundFonts.

2004-2010 — **Profesional Degree**, *Conservatorio Profesional de Música Amaniel, Madrid*, Spain.
Specialization in saxophone with Prof. Ángel Luis de la Rosa.

Experience

2011–present — **Research Assistant**, *École Polytechnique Fédérale de Lausanne*, Switzerland.
Leader of the Fukushima project. Establishing and maintaining a collaboration with the Norwegian Institute for Air Research (NILU) and with the Signal processing group at Technische Universität Darmstadt. Teaching assistant in undergrad courses *Audio Signal Processing and Virtual Acoustics*, *Signal Processing for Communications*, and *Statistical Neurosciences*.

Oct.2015– Mar.2016 — **Data Analyst**, *Swisscom Energy Solutions*, Switzerland.
Six-month internship, working on demand shaping in smart grids. Focus on big data analysis and anomaly detection algorithms. Rolling out the algorithms into production. Using Python (Pandas, Numpy, scikit-learn), PostgreSQL, Git, Phabricator.

2010–2011 — **Researcher**, *Universidad Politécnica de Madrid*, Spain.
Parallel implementations of various video codecs on a Tilera 64-core processor. European Commission projects WHERE2 (FP7) and SIMPLE (Artemis). Research interests in distributed systems and consensus algorithms.

2009-2010 — **Junior Researcher**, *Universidad Politécnica de Madrid*, Spain.
Building a universal transcoder, to convert between various audio codecs for speech and music.

Selected Prizes and scholarships

2012-2013	EPFL Doctoral fellowship.
2011	UPM co-financed fellowship in the Applied Signal Processing group.
2008	First prize, Hazen national chamber music competition.
2004	Extraordinary academic performance scholarship by the Government of Madrid.

Publications

M. Martinez-Camara, M.Muma, A. M. Zoubir and M. Vetterli. "A new robust and efficient estimator for ill-conditioned linear inverse problems with outliers", ICASSP 2015, Brisbane, Australia, 2015.

M. Martinez-Camara, B. Bejar Haro, A. Stohl and M. Vetterli. "A robust method for inverse transport modelling of atmospheric emissions using blind outlier detection", in Geoscientific Model Development, vol. 7, num. 5, p. 2303–2311, 2014.

M. Martinez-Camara, A.Stohl and M. Vetterli. "Outlier removal for improved source estimation in atmospheric inverse problems", ICASSP 2014, Florence, Italy, 2014.

M. Martinez-Camara, I. Dokmanic, J. Ranieri, R. Scheibler, and M. Vetterli et al. "The Fukushima Inverse Problem", ICASSP 2013, Vancouver, Canada, 2013.