Euclid weak lensing : PSF field estimation

Morgan A. Schmitz

NNT : 2019SACLS359

Thèse de doctorat

Euclid weak lensing: PSF field estimation

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

Ecole doctorale n°127 Astronomie et Astrophysique d'Ile de France (A&A)
Spécialité de doctorat: Sciences de l'Univers

Thèse présentée et soutenue à Gif-sur-Yvette, le 22/10/2019, par

MORGAN A. SCHMITZ

Composition du Jury :

Nabila Aghanim Directrice de recherche CNRS, Université Paris-Sud	Présidente
James G. Bartlett Professeur, Université Paris-Diderot	Rapporteur
André Ferrari Professeur, Université Nice Sophia Antipolis	Rapporteur
Catherine Heymans Professeure, University of Edinburgh	Examinatrice
Frédéric Courbin Professeur, Ecole Polytechnique Fédérale de Lausanne	Examinateur
Jean-Luc Starck Directeur de recherche, CEA Paris-Saclay	Directeur de thèse
Martin Kilbinger Directeur de recherche, CEA Paris-Saclay	Co-directeur de thèse

To my Nuncle, mentor and shaman,
Though our journeys are so different, in many a ways we know we are on the same path

Acknowledgments

Ow-wee, what a ride. Many are the humans who made those last three years as enjoyable as they were. I am pretty sure I will forget several - so let me start with them. Thank you to all of you I have forgotten!

I would like to give my heartfelt thanks to Nabila Aghanim, Jim Bartlett, André Ferari, Catherine Heymans and Fred Courbin, and especially to my two referees, for their careful reading of this document. I am honored to have had you all on my jury.

Next, I would like to give my two advisors my thanks, and my apologies as well. There was so much more we could have done if only I could have worked faster. Working under Jean-Luc's guidance was more enriching than I could possibly express - and, knowing he is not a big fan of how excessively long and lyrical my writing sometimes gets, I won't even try. Let me simply express my gratitude to him then, for teaching me so much not just on the science side, but also on the human side.

I haven't had quite as much time as I would have liked to work with my co-advisor, Martin, but I will still leave with many fond memories of our work on CFIS and, especially, on our *Table Ronde* - may we have many more occasions to work together in the future!

I would like to thank every single good teacher I had throughout the whole journey so far. With a special mention to Mr Beguier, my very first - and certainly very best - physics teacher, who once wondered if I would remember him "if I ever became a physicist". I do, sir, and hope this is close enough!

I have learned so much during my time at the CosmoStat, not just from Martin and Jean-Luc, but from many of the other wonderful people there. To those who left before me - Ming, Arnau, Rémy, Julien (though he kind of came back) - and those who will remain when I am gone - Vanshi, Jérôme, Valeria, Sandrine, Florent, Sam, Isabella, Austin (though he will soon be leaving too), Kostas, Mel, Zacc, Hung, Jiaxin, Kamshat, Toby (good luck with all the PSFs) - thanks for all the good times!

Now, if there are quite a few names missing in those two lists, it is not because they are all included in the very first of the thanks I handed out on this page (though what a great safety net I weaved for myself there). Starting with more alumni, I would like to give *special* thanks to Fred, who guided me through the first leg of my CosmoStat journey - and what a guide he was!

A great many interns have come and bolstered the ranks of the CosmoStat during my stay. I give them all my thanks. Louis, Charles, and Jérôme of course deserve a special and nominative thank you for clearly being among the very best, as illustrated, for instance, by their surviving my mentorship.

Rebeca did both of the above, and a lot more. I don't think anybody's counting, or that there even is a limit, so I'm not going to just hand her a thousand thanks, I'm going to go ahead and hand her an *infinite amount*. Not just $\aleph_0$ thanks either, here's $\aleph_1$ thanks to Beca - for her (ridiculously) hard work, the passion and enthusiasm she never ceased to show, her willingness to dive into scary code, the swiftness with which she grasped hard concepts and toyed with them as if they were the simplest arithmetical operations... For Operation Unicorn, for getting as excited as I did with our absurd idea of a non-symmetric, "gravity-well" ground metric in optimal transport, and for all the hilarious, genuine comments I keep stumbling upon in the countless parts of the λRCA code she contributed to.

Then come two former CosmoStat PhD students who have had a great deal of influence on me, whether knowingly or not. The Robbs to my Bran - François and Linc. The former is also, in a way, the (early) Sansa to my Arya; the annoyingly perfect older sibling everyone dotes upon. Though in our case of starkness-with-a-c, instead of frustration, François' blazing trail was a source of motivation and desire to better myself as a young scientist. And, yes, it perhaps contributed the slightest in the feeling that I could have done a lot more in 3 years - but that is probably not altogether unhealthy (or untrue). Though we only met a few times, each of those occasions was also a massive source of inspiration to me. The actual talks of his I attended once again painted a clear example of the scientist I hope to one day grow into, while our offline conversations often had a profound impact on me and how I approach our work. I am similarly eternally grateful to Linc, for his wise counsel in professional matters, and stalwart friendship in all others.

A lifetime ago, an almost-beardless young man arrived in an office he shared with two exceptional people.

I would like to give special thanks to Joana (désolé pour les bruits de chaise et tout le reste) and Cécile (coin coin) for making me feel at home right from the beginning!

I'd like to extend my thanks to Atooooooooor, for being one of the most fun, funniest, most positive humans I ever crossed paths with. And a brilliant, brilliant mind besides. Thanks for the JC co-rule and all the rest!

To my little academic sister Virgi (and Philipino) - thanks for the Meta Jean Louis, the top-level conversations, and generally for being beautiful human beings. Don't work too hard, and sorry I got you in trouble that one time.

While I'm on the PhD-tree-as-genealogy slippery slope of an analogy, here's to Cricri. The Jaime to my Lancel, I suppose, but not in a weird way, in a *good* way; the successful cousin (though I guess I'm technically more of his academic uncle) your immediate family always uses half-jokingly to push you forward. "So, your cousin Christophe submitted his thesis; how's the writing, by the way?" I will miss our conversations, and I truly fear for the imminent day I am expected to do science without being able to just cross the corridor to ask him about maths, or signal processing, or mostly anything.

Then there's my Pexito, First Knight of the Order of #teammorgan, and last man flying our banners. I will never forget that memorable train ride - for which I am grateful - nor some of his darkest jokes - for which... yeah, perhaps just a little less so? Thanks for your kindness, good humour, and loyalty.

Working on this thesis the last three years brought me a great many things. I suppose the most obvious are the fancy title, myriad skills and knowledge, and people I got to meet. But there is, of course, a lot more beyond all that. As I am writing these lines, I would be hard pressed to rank all those things in order of importance, or of how much of an impact I expect them to have on the rest of my life - as I suppose one always is, upon the end of any arbitrary period of one's own life. If you look back and consider, say, the twenty-second year of your life on this Earth, what would you say was the single most important thing there? And if you could go back and ask yourself, on the eve of your twenty-third birthday, what you *then* believed it would be, would you be anywhere close to the mark? Granted, the use of such an exercise is extremely doubtful. When Morgan from the year 2047 looks back, just how important the skills and knowledge we have obtained these last three years will have proven to be will be highly dependent on a great many factors. Did we stay in academia? In a closely related field? Does society look anything like what it does today? Does cosmology? I suppose what I am trying to get at is that, of all the events that occurred and the people I met over the last three years, if I had to place bets on which would have the longer lasting, most beneficial effects on my existence...

Thanks to Imane. For her perpetual support, unrivalled understanding, for her help with all the emails, for the laughter and the joy but also the sad bits, for *teaching* me so much - lessons that just might end up impacting me more than the purely scientific ones. Thanks for always being here for me, for opening my eyes to so much, for being an amazing sister - and considering how high the bar was set by the one life and biology had seen fit to give me, it's not a title I hand out easily - my excessive appreciation for the academic-tree-as-genealogy parallels notwithstanding.

Beyond CosmoStat, there are a great many people in the DAp I'd like to extend my thanks to: Delphine, Lisa, Max, Mansour (DEUX SEVRES), Bruno (thanks for your sharpeye arrowshots to the many administrative nooses that could have otherwise abruptly ended my renegade academic life), Floriane, Raphaël, Kyle, the other Kyle, Wenjie, Anna (I never got that contact for the Antarctic icebreaker, did I?), Henri, Koryo.

During a few weeks of my PhD, I had the chance to work on one specific aspect of *Euclid* that would later prove to be *pretty* important. This was some of the most enjoyable work I have carried out so far, in no small part because this is also when I got to work with Pierre-Antoine and (yet another) Jérôme the closest. I would like to thank them both: PA, thank you for your kindness, diligence in every task you undertake, for our countless long and fruitful discussions about all things PSF. Jérôme, thank you for never failing to take the time to help me out and answer all my questions when I know how precious a resource your time is these days. Thanks as well for being a true inspiration: I doubt there is a single other person that truly knows so much about our whole mission, from the hardware to the cosmology.

Thanks to Jean-Charles, for the beautiful, *beautiful* maps that the reader shall find further down these pages (and for the amazing course he gave at the Euclid Summer School)!

Thanks to Bobby, Sr., for being an oasis of easy-goingness throughout those three years - which included a few times I needed it more than I think he'll ever know.

Thanks to Vianney and Matthias for being part of my "comité de thèse," and to the latter for also being my thesis godfather (I could have translated that bit too and used quotation marks, but it's got more of a ring to it that way, doesn't it?). Thanks to Pascale for her infinite kindness while she was a part of DAp. Thanks to Philippe, though we only interacted a very few times, for - as much as I could judge from those few encounters - being well up to the job. In terms of mantles that are hard to pick up, I would be remiss if I did not thank Marie for being nothing short of an IT wizard. I'd like to extend my thanks to those who came after her, too. Thanks to (the other) Pascale and to Dominique for their efficiency and helpfulness regarding all administrative matters.

In my last year at the DAp, a big climate thing happened - I don't mean global warming itself, though I suppose it's a big climate thing, and it was certainly happening during my last year at the DAp - where a few colleagues realized we should talk and do more. I hope they can keep doing what is already happening - the seminars and small actions - and hopefully a lot more, but for now, I would just like to give them my thanks. Maybe a special, nominative one to Jérôme - a thanks and a blame, really, for making me the *second* best beard in the building.

Thanks to Luca, Fran and Theo for being amazing friends. I have been so verbose, this one sounds rather short, but it is no way indicative of how much I mean it.

Thanks to Christian - I feel like I've already given a great many thanks related to quality "chats" and "conversations," and all deserved. But those Christian brought, or contributed to, are choicest among all the quality I was exposed to. Thanks for ruining Jay-Z *and* pandas for me, man. Sigh.

Now, we've all heard it: we are nothing but *stardust*. Or, well, I suppose, we are *mostly* stardust. Or, well, I suppose, everything in us that is not hydrogen is, maybe, stardust. . . ? Or, I mean, surely life as we know it would have never led to you reading and me writing those lines, if not for the Sun, a *star*. . . ! Stars are important. Actually - and this is a bit of a spoiler for the rest of this here book, but stars are super important for building a *PSF model*. So, stars are cool (well...). And what are stars? How do they form?! I don't know (it's gas collapsing.), but Vicki does - or, well, I suppose she doesn't know either, or she'd not be researching it, would she? But even if not for the stars, I'd like to thank Vicki. For the friendship and conversations and all that, for the ink and the eternal - few things, if any, are truly eternal, but as eternal as anything directly related to my physical vessel will ever get - artwork we've built.

Now, beyond DAp is CEA, and I would like to thank CEA as a whole for making this thesis possible.

My thanks as well to CNES, for funding this thesis, and the rest: the first "JC" were incredible, allowed me to meet some of the below-mentioned wonderful people, and, if I had not happened to already be a bit of a space-nerd - which I most definitely was - it might very well have turned me into one. The second were even better.

I'd like to thank Olivier for being a great CNES "correspondant[1]" - for his support and words of encouragement during our more official meetings, and for his never-failing interest in my work when we chanced upon one another on other occasions. Thanks to Philippe for, from what I can tell from our few meetings[2], his similarly high levels of interest and benevolence.

Thanks to Myriana for taking care of us this whole time, and to her and the rest of the organizing committee for the - I cannot stress this enough - *inspiring* CNES events I was lucky enough to attend during my PhD.

On the topic of CNES events, thanks to Michel for his candor and approachableness during those same events.

[1] Not a typo, just French.

[2] This seems to be a recurring thing in my interaction with Philippes.

Thanks to my fellow PhD students I was lucky enough to meet there: Selma, Victor, Max, Jay-Bee, Matteo (why would you eat a tomato, Matteo? A TOMATO!!), Ismael, Blandine and all her pet bears, Cyril, and all the rest.

On the topic of CNES people, here's a great segue into the next unofficial subsection (at this point - and I'm nowhere near done - maybe I should have made them official) of acknowledgments: Rodolphe. I would like to give another infinite amount of thanks to Rodolphe, for being the mastermind behind so many of the venues that made my *Euclid* life - both scientific and human - incredible those last three years. Chief among these: the *Table Rondes* of the Euclid Summer Schools.

Thanks to Céline and Théo - and I am not going to say Martin again because, as the careful reader may have noticed by now, I am trying to name each person once and only once, which forces my hand into making both odd omissions and agile throwbacks - for the amazing times and good science.

Thanks to the rest of the Euclid Summer School organizing committee - Alain, of course, and I suppose I am going to be a little anachronous here, but this tends to happen when a mortal witnesses the beginning of things, to Bogna and Stéphane as well. Those three sessions[3] of the school taught me loads (and there were loads to be taught, trust me), but also allowed me to meet some of the funkiest members of the Youth Within Euclid[4].

Thanks to Michel-Andrès, Tuan, Paul, Paulo (hipster*er* than I will ever be, try as I might not), Florian, Jenny, Isaac, and la "team des Marseillais" (in Frenglish in the original text): Marie, Alex, Sylvain, Adam, Philippe, and Pierros. Thanks to Calum (Calum Calum Calum Calum), C-A, Loulou, Mehdi, Romain, Miguel and Raphaël.

Thanks to Manu, Flo, and Thibaut for cofounding the very best game that mankind ever came up with[5]. And for the brilliant times we had playing it.

Thanks to Vero for being incredible, as a person, in general, but also as our host in 2017. Thanks to the Winch and all its staff[6] for 2018.

Now, 2018 was the very best session in many regards - but this is not about the Euclid Summer School, it's about me being thankful for people that helped during my thesis. In 2017, five souls found themselves discussing life, the Universe, music, and *a lot* of other stuff. It was not all discussing either, we sometimes drew circles in the sand and all that. Those five people I'd like to give special thanks to - wait, no, *four* people, I'm one of the five. Tguy the philosopher. The *other* Thomas, who brought into my life what people further down this lengthy note of thankfulness and I later would later come to re-dub "Le Jeu Irlandais". Axel (t'as cru que je t'avais oublié, hein ?) and his love for flatearthers and UFOlogists alike. And of course, Sylvain, our very own Lemaître. We were mates before Euclid schools, and I daresay we shall remain mates after.

Now, as Chapter 4 will show you (I mean, if you're here because you plan on reading this cover to cover, at least), a great good part of my thesis work was more applied maths than applied cosmology. In this time, I came to meet, interact, collaborate, and eventually co-build a huge thing with several brilliant people.

I would like to give especially warm thanks to Matthieu, for our countless email conversations, then Slack conversations, then, *phone* conversations[7].

Thanks to David, and especially to Nicolas, who started the whole game before Fred and I even knew it was Wasserstein Dictionary Learning we *wanted* to do.

[3] And only three - and here would go a sadface emoji, but I promised myself there would be *none* in this thesis, not even in the acknowledgments.

[4] *Not* an official *Euclid* term.

[5] Meta Jean-Louis. Please contact the corresponding author for details.

[6] *All* its staff.

[7] And some people could tell you, especially among those yet-to-be-named in those here acknowledgments, that a Morgan does hate a phonecall.

Thanks to Marco for - well, for Cuturi (2013), for starters. And then, for being a wonderful collaborator - not just in the science, but even in the *writing* thereof.

Thanks to Gabriel for being ridiculous - this is the third (I think?) time I use this adjective in a very flattering way, but I mean it. Gabriel is a *ridiculous* researcher, in all the best ways possible. Here, since I already referenced a paper, take a look at equations (4.11-4.13). I bothered Gabriel *so much* with these ones[8]. Gabriel was once described to me - by someone I won't name - but I *will* say I might have thanked him.her.them already, maybe - as "probably the greatest contemporary applied mathematician." Despite that, and the fact you know that that one thing he's working on with you is but one of the seven hundred projects he's actively involved in *today*, he will still take the time - however long you need - to discuss that extra term in (4.11-4.13). So, thanks, Gabriel!

On a completely unrelated note, but before I move on, I would like to thank all members of the Cosmostatistics Initiative, especially fellow CRP#6 participants, and even more especially Emille, Alberto, and Raphael.

Thanks to the Kingscrew: Badri, Isaac, Adarsh, Ismaël, Agus, François, and of course Gilles, for making my very first conference memorable.

I do believe this concludes the "professional" part. Now, a thesis - or any research work for that matter - is not just about the people you interact with professionally. In fact, among the people named beyond, some would gladly point out how a lot of "great" astronomers and physicists would not have been quite so great without the help of other humans history may not remember quite as vividly. Great or not, I would have most definitely not made it here if not for all my wonderful friends.

Thanks to Albert for Albert's and to Joce and Juan for la Bodega.

This is as good as any another spot, so here's to him.her: I'd like to thank Marianne, for his.her everlasting friendship. I'd like to thank him.her for the art though ah, I suppose I'm not the only slacker around here, eh?

Le Club Social might be dead - longue vie au Club Social ! - but it certainly hit its golden age during the time I was working on this thesis. You all are special to me - and most of you were before this whole thing - but, as I hand you my thanks, I hope you know the thing[9], dead as it may be, really meant a lot to me.

I therefore would like to thank Seigneur Lapinier of House Faucheuse, Leo Longthorn, Lady Rose and Ser Cavalier of House Brouzesse.

I would also like to thank our wonderful pledges: Nat, who was to be the second of the order of Z, Seg the Capitalistus (who taught me how to buy $TSLA - right this instant, this was a good thing), and Marion, soon-to-be (in Club Social Terms, no pressure otherwise) of the House Faucheuse, who helped us push the boundaries of le Club Social, literally and figuratively.

First Lady of the Second Sun, aka Dame Lao - earlier I was rambling about climate. Dame Lao has certainly been another staple in my recent awaking. Thanks to her for that, apologies also for that one time the padawan thought he knew better than the master - and see you in America!

I'd like to thank Milady of Memoria, aka Tétée Matinale[10], always kind of implicitely number 3 of le Club Social - though there was of course never any *numbers* since there was never any hierarchy. Our friendship started very early in my Parisian life, something about makeup and collecting certain kinds of bottles, long before there was a Club Social, and never ceased to grow since then.

Thanks to the Countess of House Faucheuse as well - though it is perhaps best not to dwindle on the early days of our friendship, I can similarly say it's only grown stronger over the years.

[8] Long story short: if you ever find yourself thinking there's a discrepancy between these and, e.g., those of Benamou et al. (2015)... There isn't. We're solving the same problem.

[9] "Wh.. wh... wha... what thing?" "The thing, the *THING*! It... It's got like buttons on, lights on, it beeps."

[10] The best thing I thing I *ever* came up with, and I will stand by that statement, even though we're not quite yet at page-1-out-of-over-200 of this here document.

Some of the foundations on which le Club Social was built were laid out years before it was ever founded, When a few young humans converged upon a certain prépa, where they thought they would unlock the secrets of maths and science. Did they? Eh.

What they did unlock, undoubtedly, was the key to each other's friendship. Wait, one does not unlock a key. What they did *find* was the key to each other's friendship.

Thanks to Gibergues, Davy, Coco.

I10 was born then, isn't dead, and I hope never will. Wait, I guess we do not want to impose anything upon our (more-or-less, depending on which of us three you ask) hypothetical children, so maybe it should, someday. In the meantime, I'd like to thank Thomas for being a third of I10, for being my meme dealer when I needed it back then, for being the VP of my Club Fusée, and most importantly, for being a true, everlasting, friend. Allô. Allez ? Allô, allez ? Allez, allô, allez ?

I'd like to thank Steven. Steven and I have been through this over-the-top thank you game before. I'm pretty happy with where we left it last time, so I won't try to outdo it this time. Thanks, Steph.

Also, thanks to Carbal. For the freshness you brought - and keep bringing - to us all. Carbalou, there's this time, when you were doing the thing with the models and the paint and all, and I wasn't there. I don't even remember, so I couldn't tell you why now, but I suppose it was something important. You never blamed me for that. This is supposed to be thanks, not apologies, but I've already done the thing where I hand out both at the same time, so Carballou - thanks for being amazing, and sorry for not being there that time.

It's been a few paragraphs since we had a healthy (hah!) Song of Ice and Fire reference, hasn't it? I'd like to thank Martina and Cash, the Ygritte and Ghost to my Jon Snow.

I started (and finished) my thesis in the very, very far suburbs of Paris (ask me about my commute[11]), while living inside the walls of the City of Love. Before then, I came from a better place.

From there I keep some of my oldest and staunchest friends, that as I went through those 3 years of science, went alongside me through 3 years of *life* - such different lives all, yet we never lost one another.

My thanks to Rosé (of the Jeu de Joséphine et Morgan fame), VITHOR (lead guitarist of the Filaments de Banane), Clothilde, Clément, Anouk, Manu and Nono.

I thank Q for all the ONE, TWO, WOOP-WOOPs and for being central in holding us all together through the years.

Especially though, I'd like to thank *les Parisévriens*[12], for their frequent and crucial support these last few years.

Thanks to Lady Manue (and the Lord Cedex), and DC (another role model of mine).

Thanks, Maxou - you've certainly been one of those I could have named most often, for you've been a part of so many different aspects of my life, as far back as our TSB days. Through each of those we shared amazing moments - thanks.

Speaking of TSB!

I'd like to thank, of course, my bro Maury. Our story started long before this thesis - our story started long before almost anything that is still around in our lives today, really. Our story went through some really good phases and some just-decent phases. Through it all, we've remained TSBs, and we've remained brothers.

Jo (tout ça parce que le chien est noir). We haven't met face-to-face in far too long. But still, I know you're there. Thanks for your random showing up to that one new year's eve, and for our very recent online conversation. TSB for life!

I'd like to thank Lionel as well, and Virgile. And here's another great segue! Okay, maybe not great, but it's a segue!

Eléonore, I'd like to thank you, not because it is easy, but because it is hard[13]. This is three years of my life,

[11] Don't ask me about my commute.

[12] A name so brilliant that I wish I came up with it, though it was none of my doing.

[13] Please read this with JFK's voice.

and you were a big part of some of them, and also, you never spilled the beans on what TSB means. Thanks!

I'd like to thank Billy for being Billy. "Is this a Billy and Z thing?"

Here's some neither Z, nor Billy people I'd like to thank for all the Billly-and-Z-things: Meg (Day 799. They still don't know I'm scrambled eggs.), Brian, SydSyd, Oddmund, Lady.

Most people end their acknowledgments with something along the lines of "of course, thanks to my family". I am not most people, though I'd definitely be nowhere without my family.

My beloved siblings. I know you'd make drama out of who came first, so you both get the same paragraph, okay? The Morgan that wrote this only came to being through kid Morgan, that you both had an incredible influence over. And you still do, of course. Thank you both, for everything. I love you, and I love all my wonderful nieces and nephews. Thanks for them all! And thanks to my amazing sister- and brother-in-law as well!

Thanks to Manou for being the best grandmother there ever was.

Dad. Thanks for all our father-and-son times, and for all the rest!

Alex. Thanks for all our Final Fantasy summers, and for all the crazy unexplainable stuff that happened to us back then. Let's do it again soon!

Thanks to my kiwi family for their neverending oversea support.

And lastly, here come a few full specials.

My two C0110C20UF[14], mon Pellos et ma Comtesse. You two are amazing and I had some of the very best times of my life *so far* in our amazing house. So far in italics, because house or no, as long we have each other, I'm sure we'll have many more. Thank you so much.

Mon Nas, second oldest of all my friends, in this realm and a few others. Thanks for everything.

Valou, "mon ex-meilleur-ami," you know how much I admire you and how big an influence you have on me. Thanks for fighting the good fight(s)!

Milady Hand, the Bittersteel to my Blackfyre, the Orys to my Aegon, "my shield, my stalwart, my strong right hand," my Best Friend. You make life great. Thank you.

And, last but not least, thanks to the best, most important person in my life, my beloved mother.

[14]This used to be our Wi-Fi password - this is true.

Summary in French

Depuis la découverte de l'accélération de l'expansion de l'Univers, à la fin du millénaire dernier, un "modèle standard de la cosmologie" a émergé. Ancré dans la théorie de la relativité générale d'Einstein, il repose sur l'existence de *matière noire* - qui interagit gravitationnellement mais n'émet pas de lumière - et d'une mystérieuse *énergie sombre*. Avec seulement sept paramètres libres, ce modèle explique remarquablement bien les observations que nous pouvons faire de notre Univers aujourd'hui.

De nombreuses questions demeurent cependant ouvertes. On ignore toujours la nature précise de la matière noire. L'énergie sombre - le nom que l'on donne à la cause l'accélération de l'expansion - est plus mystérieuse encore. L'hypothèse la plus simple est qu'il s'agirait simplement d'une constante cosmologique. Cette possibilité est considérée par beaucoup comme peu satisfaisante d'un point de vue théorique. De plus, de récentes observations semblent indiquer l'existence de potentielles *tensions* : selon la manière dont on mesure la valeur prise, dans notre Univers, par les quelques paramètres cosmologiques du modèle standard, nous obtenons des résultats qui pourraient être incompatibles. Si ces tensions sont confirmées, elles pourraient indiquer le besoin de nouvelle physique, au delà du modèle standard.

L'un des objectifs principaux de la cosmologie observationelle moderne est donc de mesurer, avec une très grande précision et de différentes manières, la valeur des paramètres cosmologiques. Chacune de ces manières constitue une sonde cosmologique. Parmi ces sondes figure le *lentillage gravitationnel faible*, un moyen particulièrement prometteur de relier des observations aux paramètres cosmologiques.

Lentillage gravitationnel

La théorie de la relativité générale prédit que les objets massifs déforment l'espace temps. Lorsque la lumière se propage, elle suit ces déformations. Les objets massifs modifient donc, tels des lentilles, le chemin de la lumière. Cet effet, dit de *lentillage gravitationnel*, a été confirmé par de nombreuses observations, notamment lorsque notre système solaire, un objet d'avant-plan (une galaxie ou un amas de galaxies) et une galaxie d'arrière-plan sont presque alignés. Plutôt que de regarder individuellement de tels systèmes, on peut s'intéresser à l'effet de lentillage causé par toute la structure à grande échelle de notre Univers. Cela revient à mesurer la forme d'un très grand nombre de galaxies. Prises seules, ces mesures ne nous permettent pas de tirer d'information cosmologique, puisque nous ne connaissons pas la forme que possédait la galaxie avant que sa lumière ne se propage jusqu'à nous : c'est le régime du lentillage gravitationnel *faible*. En en mesurant un très grand nombre, et en regardant les propriétés *statistiques* de la distribution de formes de galaxies, on peut cependant extraire de l'information sur la distribution de matière dans l'Univers.

De telles mesures cosmologiques, basées sur le lentillage gravitationnel, ont déjà été effectuées depuis le sol et avec le télescope Hubble. C'est également l'un des deux objectifs scientifiques principaux du télescope spatial *Euclid*, de l'Agence Spatiale Européenne, dont le lancement est prévu en 2022. *Euclid* observera la presque-totalité du ciel extra-galactique, et nous permettra ainsi d'obtenir des contraintes extrêmement fortes sur les valeurs des paramètres cosmologiques... A condition que l'on arrive à mesurer précisément la forme de plusieurs *milliards* de galaxies.

Cette mesure s'avère extrêmement difficile en pratique, principalement à cause de la fonction d'étalement du point (PSF, pour Point Spread Function en anglais). Toute observation faite à travers un instrument optique est légèrement floutée. Ainsi, si un objet devait apparaître comme un point isolé, on l'observerait plutôt comme une tâche légèrement étalée. C'est cet effet qu'on appelle PSF. Ses origines sont diverses : diffraction, imperfections dans l'optique de l'instrument, turbulences atmosphériques (pour les télscopes au sol), etc. L'impact de la PSF sur la forme des galaxies est beaucoup plus important que celui causé par le lentillage gravitationnel. Il est donc crucial de corriger très précisément cet effet. Pour ce faire, il faut d'abord connaître la PSF elle-même. Dans le ciel, les étoiles non-résolues devraient apparaître comme des points ; aussi donnent-elles une mesure de la PSF.

Estimation du champ de PSF

Cette thèse porte sur le problème d'estimation du *champ de PSF* : à partir d'images dégradées d'étoiles, comment reconstruire un modèle de PSF qui pourra ensuite être utilisé pour la mesure de forme de galaxies ? Nous nous intéressons en particulier à la résolution de ce problème dans le cas d'*Euclid*. La variation *spatiale* de la PSF, c'est-à-dire le fait que la PSF varie en fonction de la position de l'objet dans l'image, était déjà présente dans toutes les études passées de lentillage gravitationnel. En raison des spécificités d'*Euclid*, la modélisation de sa PSF demande en plus de résoudre de nombreux défis qui n'avaient pas été rencontrés auparavant. En particulier, la PSF sera *sous-échantillonnée* et subira de fortes variations *chromatiques*.

Nous proposons l'utilisation d'outils avancés, issus des mathématiques appliquées et du traitement du signal, pour relever ces défis. Pour contrecarrer le sous-échantillonnage, c'est-à-dire le fait que les pixels du détecteur de l'instrument visible d'*Euclid* sont "trop gros" pour capturer tous les détails de la PSF, et les variations spatiales, nous étendons la méthode RCA (Resolved Components Analysis, en anglais), développée récemment au CEA. En utilisant les notions de *représentation parcimonieuse* et de *théorie des graphes*, nous construisons un modèle de PSF qui permet de procéder à la super-résolution (c'est-à-dire de retrouver de l'information contenue dans des pixels plus fins) d'images bruitées d'étoiles, et qui capture les variations spatiales de la PSF se produisant à différentes échelles.

Pour étudier la qualité de ce modèle et le comparer à l'état de l'art, nous réalisons un grand nombre de simulations d'images, non seulement d'étoiles, mais également de galaxies. Cela nous permet d'étudier l'impact des erreurs commises dans l'estimation de la PSF sur notre estimation de la forme des galaxies. Nous montrons ainsi que cette propagation d'erreur s'avère, dans le cas d'*Euclid*, plus complexe qu'anticipé. En particulier, le formalisme largement utilisé dans la communauté jusqu'ici pour propager les erreurs de PSF ne sera alors plus valable.

L'instrument visible d'*Euclid* couvre une très large bande du spectre électromagnétique : d'environ 550 à 900nm en longueur d'onde. Contrairement aux études précédentes, où les observations utilisées pour le lentillage gravitationnel étaient effectuées à travers des filtres beaucoup plus fins, il devient donc capital de prendre en compte les variations chromatiques de la PSF. L'objectif est alors de construire un modèle à trois dimensions : les deux dimensions spatiales habituelles, ainsi que la longueur d'onde. C'est un problème extrêmement difficile à résoudre, puisque nous ne disposons pour chaque étoile que d'une seule image qui, plutôt que d'être prise à une longueur d'onde particulière, est *intégrée* avec un spectre, propre à l'étoile !

Transport optimal

Pour faire face à cette très grande dégradation de l'information, nous proposons d'utiliser le *transport optimal numérique*. Le transport optimal est une branche des mathématiques qui présente un grand intérêt théorique depuis plusieurs siècles. On attribue souvent son origine à Monge, qui en 1781 proposait d'étudier les problèmes de transportation de masse : étant données deux configurations de masse (par exemple, un tas de sable qu'on désirerait utiliser pour remplir un trou de même volume), et connaissant l'effort associé au déplacement de masse, quelle est la manière *optimale* (c'est-à-dire associée au moindre effort) pour passer d'une configuration à l'autre ?

Jusqu'à récemment, le calcul des quantités associées aux problèmes de transport optimal était trop coûteux pour qu'elles puissent être utilisées dans des problèmes pratiques. Depuis, des méthodes pour calculer efficacement des approximations de ces quantités, comme celle proposée par Cuturi en 2013, ont rendu possible l'avènement du transport optimal numérique. On peut ainsi, par exemple, calculer très rapidement le *barycentre de Wasserstein* de plusieurs images. On peut songer au barycentre de Wasserstein comme les différentes étapes intermédiaires par lesquelles passent notre tas de sable au cours de son transport. C'est un opérateur qui permet ainsi de capturer les déformations géométriques séparant plusieurs objets.

Nous proposons d'utiliser ces notions pour capturer les variations chromatiques de la PSF d'*Euclid*. En combinant les outils, basés sur le transport optimal et développés dans cette thèse, avec les fondements de la méthode RCA, nous créons le premier modèle permettant de modéliser de manière polychromatique la PSF

Euclid. Cette approche, nommée λRCA, permet ainsi de construire un modèle à trois dimensions de la PSF en n'utilisant rien d'autre que des observations sous-échantillonnées d'étoiles, intégrées avec leurs spectres (supposés connus). En particulier, elle ne nécessite aucune connaissance a priori sur l'instrument.

En plus de l'application aux PSFs polychromatiques, les outils développés dans cette thèse nous ont permis d'introduire l'apprentissage de dictionnaire de Wasserstein (Wasserstein Dictionary Learning, en anglais), une méthode de représentation des données qui utilise pleinement les propriétés du transport optimal. Ses applications potentielles sont très vastes, comme le montrent par exemple les travaux de Xu et collaborateurs, qui ont utilisé (et étendu) dès 2018 notre approche pour du traitement de texte, avec une application aux formulaires d'admission de patients à l'hôpital.

Contents

Chapter 1

Introduction

In 1915, Einstein introduced the theory of General Relativity (GR) that would come to have a huge impact on our understanding of physics and the Universe. This year, just over a century after Einstein's presentation at the Prussian Academy of Sciences, the team of the Event Horizon Telescope revealed the first direct image of a black hole's shadow. This came very shortly after another validation of GR, right before I started the PhD work that will be presented in this thesis: in February of 2016, the LIGO collaboration announced the first detection of gravitational waves.

Long before these observations, the application of GR to the Universe as a whole had already laid the foundations for what would become modern cosmology. A picture of the history of our Universe, from its beginnings - the Big Bang - to the present day started to emerge. From these early days, we eventually reached an era of *observational cosmology*. Much like astronomers and astrophysicists had already been doing, the different models and predictions of cosmology could now be put to the test with actual observations of our own Universe - notably, in the 1990s, the COsmic Background Explorer (COBE) gave us a wealth of cosmological insight from measurements of the Cosmic Microwave Background (CMB).

It is also thanks to observational evidence that two key components of our current understanding of the Universe came to our consideration. First, several different observations indicated that ordinary matter alone could not explain all of the gravitational effects we observed. This led to the addition of *dark matter*, which interacts gravitationally but emits no light, as one of the main components of our Universe. Dark matter is in fact much more prominent than ordinary matter, by about a factor of five in its contribution to the total energy content. Second, Hubble noticed that the farther away galaxies are, the faster they appear to move away from us: not only is our Universe expanding, but this expansion is accelerating. *Dark Energy* is the name we gave to the force responsible for this acceleration.

These three components - GR, dark matter and Dark Energy - are the pillars of the Standard Model of Cosmology. This model is remarkably well supported by observational evidence, yet several issues still need to be addressed. The nature of dark matter and dark energy remains a mystery. Is the latter truly nothing but a cosmological constant? While this possibility is not ruled out by present observations, it would come with questions of its own. Why this particular value, conspicuously close to 0? It also is, to many, a deeply unsatisfying answer from a theoretical standpoint. Moreover, as our cosmological probes reach ever higher levels of accuracy, potential *tensions* start to emerge - that is, measurements of the same quantity, from two independent sources, that would be in contradiction within the Standard Model. At the time of the writing of this thesis, the most notable is that related to H_0, the *rate* of the expansion of the Universe. Measurements, made either directly from supernovae or indirectly from the latest CMB data, appear strongly at odds with one another. Another potential tension is related to dark matter and its distribution. Its significance is less certain, but the sources of the potentially conflicting measurements are particularly relevant to this work: one comes from the *Planck* mission and its CMB measurements, while the others are made from the study of *weak gravitational lensing* - the heart of this thesis.

Gravitational lensing, another effect predicted by GR, can lead to spectacular effects. For instance, in an Einstein Ring, what should be a simple background object appears as a ring-like structure around a massive foreground object. In the weak (gravitational) lensing (WL) regime, these effects are of a much smaller amplitude, and cannot be discerned by eye. In this case, no information can be gleaned on an object-by-object basis, but a statistical study of the (observed) shapes of a large number of objects can reveal information about the late-time Universe, and especially about the distribution of matter (dark or otherwise). This makes WL a very powerful cosmological probe.

For our observational efforts to address some of today's most pressing questions in cosmology, whether it is the nature of Dark Energy or the possible tensions between different probes, we must achieve exquisite measurement accuracy over very large portions of the sky and/or for very large numbers of objects. These ambitious goals make up what is often referred to as *precision cosmology*. After *Planck* and the CMB, we must now bring WL to this new level of sky coverage and precision. That is one of the main goals of the *Euclid* space mission, expected to launch in 2022.

Until then, we can turn our efforts to some of the many *data processing* challenges that come with any WL survey, as well as some of the new ones that will arise specifically in the case of *Euclid*. Precision cosmology is not just about converting "observables" to cosmological information - care must also be given to how we go from the actual observations (raw images from our telescopes) to the quantities of interest (galaxy shapes that contain the WL signal, in our case). Throughout these various steps, many challenging problems arise, some of which we can only hope to solve by deploying advanced signal processing methods.

Amongst the most major of these challenges is that of the Point Spread Function (PSF). Every astronomical image is distorted because of various effects: diffraction, imperfect optics, atmospheric effects (for ground-based telescopes) or the slight jitter of the instrument (when in space). In the case of WL, where the signal of interest lies in the *shapes* of galaxies, it is naturally of paramount importance to correctly account for the distortions caused by the PSF. This, in turn, can only be achieved if our knowledge of the PSF itself is sufficient. The PSF *estimation* problem is the main focus of this thesis, with a special emphasis on the Euclid Visible instrument (VIS).

Finding an adequate PSF model from the objects found in *Euclid* exposures amounts to solving an ill-posed *inverse problem*. Within this framework, we will use advanced mathematical signal processing tools to try and achieve the best possible PSF model that is able to account for all of *Euclid*'s specificities. Super-resolution, spatial variations, and chromatic variations of the PSF will be addressed in this thesis. Each of these sub-problems associated with the Euclid PSF will lead us to delve into advanced and exciting methodological tools. Sparsity, (a small amount of) graph theory, and especially Optimal Transport (OT) will all come together as we get closer to a non-parametric PSF model applicable to *Euclid*.

This thesis is organized as follows. Starting from the very general, in Chapter 2, we will give some cosmological context, then narrow our attention first to lensing as a cosmological probe, then specifically to PSF modelling and *Euclid*. We will then start our PSF modelling efforts in Chapter 3, in a simplified context where the chromatic variations of the PSF will be left aside. Before turning to these, we will first take a look at the recent field of numerical optimal transport in Chapter 4. These tools will then be incorporated in our PSF model in Chapter 5, where we will add an OT component to our monochromatic model to make it the first non-parametric PSF model able to capture chromatic variations.

Chapter 2

Cosmological and practical context

Contents

In this chapter, we present the cosmological context of this thesis. Starting with a general presentation of the standard cosmological model as it stands at the time of this writing, we then quickly narrow our attention to the particular cosmological probe of interest for this work by giving the basic lensing formalism on which WL experiments and results are based. By the end of this chapter, we will have established the role of the PSF in WL studies, the importance of accurately modelling it, and the challenges encountered in the particular case of *Euclid*. It should be noted that, while we use the standard model and the estimation of its parameters as the cosmological motivation for this task, WL can absolutely also be used to constrain exotic Dark Energy or Modified Gravity models. Similarly, in the present chapter, we will introduce and focus on *cosmic shear* as the way to exploit WL observations. They can however be used in numerous other ways, both for cosmological inference, and to achieve different science objectives. In all cases, the PSF model used is crucial, and can lead to severe systematic errors.

2.1 The standard cosmological model

We start by giving a brief overview of our current understanding of our Universe and its history, as summarized in Figure 2.1. It begins about 13.8 billion years ago with a singularity in space-time, famously known as the *Big Bang*. Shortly thereafter, a period of exponential expansion, called *inflation*, occurs. It ends when our Universe is but 10^{-32} seconds old, yet that is enough to expand small quantum fluctuations into the footprints of what will ultimately allow for the creation and growth of structure. Following the inflation, all of the Universe's components live in a hot plasma. In particular, this means photons are coupled to baryonic matter, and are not free to escape for about $380,000$ years. During this time, as the Universe expands (albeit at a much slower rate than during inflation), it cools down. $380,000$ years marks the time when the Universe's expansion causes it to fall below a temperature of around 4000K, which allows for *recombination*: electrons and protons combine to form (hydrogen) atoms, and the photons are free to escape for the first time. Some of these photons can be observed today in the microwave domain, and constitute the CMB. Observations of the CMB contain a wealth of cosmological information, making it one of the main *cosmological probes*.

Following the emission of the CMB photons, as the Universe continues slowly expanding, the small anisotropies imprinted by inflation lead to the creation of *structure*. These structures are mostly governed by the gravitational action of dark matter particles. After several hundred million years, enough ordinary matter falls into these dark matter(-dominated) structures to start forming stars, and eventually galaxies. This marks the end of the period ominously called the *Dark ages*, as these new objects start emitting photons. These, in turn, cause some electrons to be stripped from atoms that had been neutral since recombination, an event called *reionization*. While these first stars have not yet been observed by mankind, as time goes on and more of them form, we eventually start seeing more and more of these galaxies, which then act as tracers allowing us to indirectly observe the Large Scale Structure (LSS) of the recent Universe. In the present day, this LSS appears as the beautiful *cosmic web* that can be seen toward the right-hand side of Figure 2.1: central, dark matter *halos* are connected by filamentary structures of dark matter. Within these, structure can now be discerned even when observing ordinary matter, as galaxies themselves gather in gravitationally-bound *clusters*.

One last cosmological landmark worth mentioning here occurred only about 4 billion years ago[1]. The Universe had been in the matter-dominated era until then (and ever since it was only a few 10^4 years old, *before* which it was radiation-dominated). Starting then, our Universe entered the present era of Dark Energy domination, in which its expansion is accelerating.

We give in the following subsections some very brief insight into how the picture we have just presented is built, starting with some necessary basic notions in the underlying theory it is built upon: GR (Einstein, 1916). Much like measuring the CMB makes it possible to impose constraints on our Universe and its history, the LSS of the recent Universe makes for a rich vein of cosmological information. In Section 2.2, we will dive into WL, one of the main ways to probe this LSS.

2.1.1 General Relativity

Classical physics consider a 3-dimensional (3D) space, wherein systems may evolve with time. In GR, a 4-dimensional *spacetime* is considered instead, the local geometry of which is then entirely defined through

$$\mathrm{d}s^2 = \sum_{\mu,\nu=0}^{3} g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu := g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu, \tag{2.1}$$

where $g_{\mu\nu}$ is the *metric tensor* and the index of coordinates x^μ varies from 0 to 3 (with $\mu = 0$ corresponding to the time coordinate, and $1 \leq \mu \leq 3$ to space). Through the second equality, we implicitly define the Einstein summation convention, wherein repeated indices are summed over. We shall use this convention for the

[1] Which, interestingly though irrelevantly, happens to be in the same order-of-magnitude timescale as that of recent estimates for the appearance of life on Earth (Dodd et al., 2017).

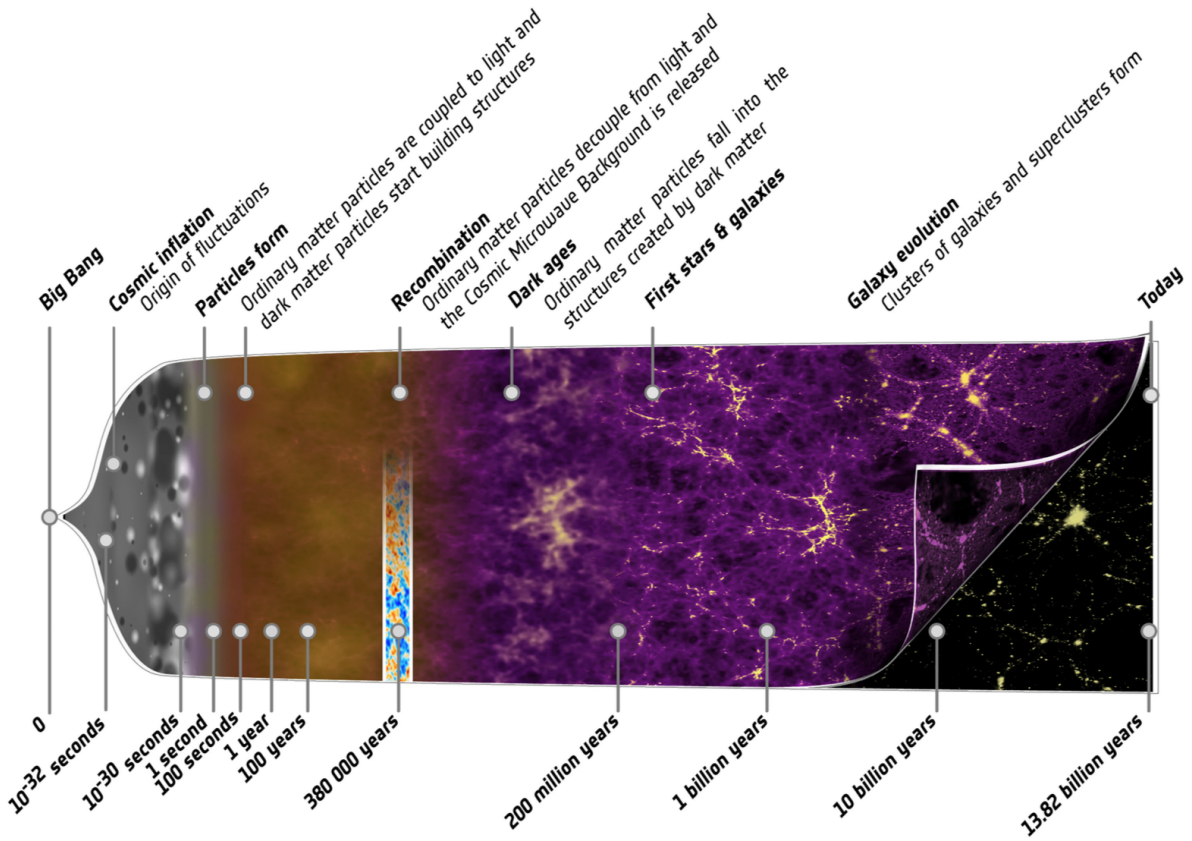


Figure 2.1: An overview of the cosmic history of our Universe. Credit: European Space Agency (ESA), C. Carreau

remainder of this chapter only. The := symbol, on the other hand, will be used for definitions throughout this thesis.

In classical physics, and in the absence of influence of any external force, the movement of particles through time occurs along straight lines. By analogy, in GR, the motion of particles is described via the *geodesic equation*,

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = 0. \tag{2.2}$$

Since we now consider time to be one of the dimensions of spacetime, we use an increasing scalar λ (rather than time t) to parametrize the evolution of our particle across all 4 dimensions. $\Gamma^\mu_{\alpha\beta}$ are the *Christoffel symbols* and depend only the chosen metric:

$$\Gamma^\mu_{\alpha\beta} = \frac{g^{\mu\nu}}{2} \left(\frac{\partial g_{\alpha\nu}}{\partial x^\beta} + \frac{\partial g_{\beta\nu}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^\nu} \right). \tag{2.3}$$

Equation (2.2) gives us the GR equivalent of the equation of motion: instead of moving across space in a straight line, our particles will now move along the *geodesics* of the curved spacetime. The last element we need is the origin of this curvature when we add content to our system; this is described by the famous Einstein field equations:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{2.4}$$

where $G_{\mu\nu}$ is the Einstein tensor which, once again, is only a function of metric $g_{\mu\nu}$, G is Newton's Gravitational Constant, c is the speed of light, and $T_{\mu\nu}$ is the energy-momentum tensor that describes said contents.

2.1.2 The isotropic, homogeneous Universe

GR-based cosmology amounts to applying the theory to the Universe as a whole. If we are to find solutions to (2.4) in this context, we need further assumptions. We will thus assume the *Cosmological Principle* to hold true. It states that the Universe is isotropic and homogeneous - a reasonable assumption at large scales. Actual observations of the CMB, for instance, confirm the Universe we live in is indeed very homogeneous - almost alarmingly so, as we shall discuss shortly. Of course, this is clearly no longer the case at the scales we experience daily. We will address the issue of structure formation in Section 2.1.4.

For now, let us simply assume the cosmological principle to hold true. It can then be shown that the metric is unique, called the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, and allows for (2.1) to be written as:

$$\mathrm{d}s^2 = -c^2 \mathrm{d}t^2 + a^2(t) \mathrm{d}l^2, \tag{2.5}$$

where $a(t)$ is the *scale factor*. As first suggested by Lemaître (1927), there is very strong evidence that our Universe is expanding. A typical, yet useful illustration of this expansion is that of a sheet of graph paper: as shown in Figure 2.2, suppose two galaxies MW and A lie (motionless) at the intersections of two consecutive vertical lines. Their distance, as measured in units of the graph paper's squares, would be $\chi = 1$. If the graph paper itself was to expand with time, the *physical* distance (red arrow) between our galaxies would increase with time; yet, since they are motionless, they would remain on the same nodes of the graph paper, therefore still be separated by a "graph paper unit" distance of $\chi = 1$. For this reason, we call χ the *comoving* distance. The physical distance between MW and A would then be some function of time, $a(t)$, describing how the expansion of our graph paper takes place - this function is none other than the scale factor we came across in (2.5).

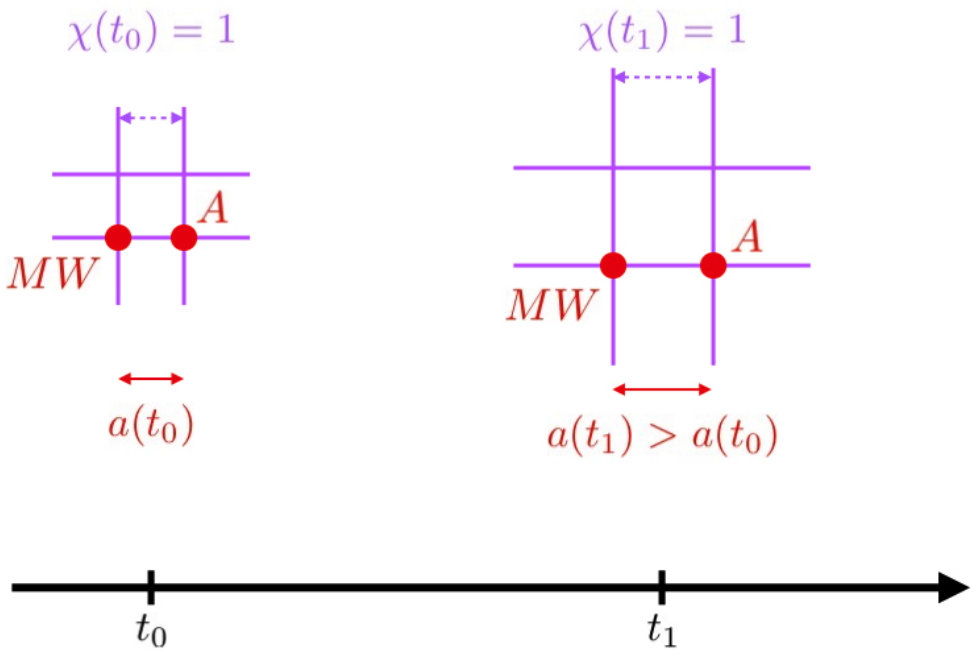


Figure 2.2: Schematic representation of the expansion of the Universe.

Back to equation (2.5), the line element $\mathrm{d}l^2$ can be further decomposed by making use of the comoving distance:

$$\mathrm{d}l^2 := \mathrm{d}\chi^2 + f_K(\chi)^2 \mathrm{d}\omega. \tag{2.6}$$

f_K is the comoving *transverse* distance, related to the geometry of the Universe, parametrized by K,

$$f_K(\chi) := \begin{cases} K^{-1/2} \sin\left(K^{-1/2}\chi\right) & \text{for } K > 0 \text{ (spherical)} \\ \chi & \text{for } K = 0 \text{ (flat)} \\ |K^{-1/2}| \sinh\left(|K|^{-1/2}\chi\right) & \text{for } K < 0 \text{ (hyperbolic).} \end{cases} \tag{2.7}$$

The combination of some of our latest cosmological observations (Planck Collaboration et al., 2018) seem to indicate we live in a flat Universe, and place strong constraints around the value of K (or rather, Ω_K, which we shall define shortly). In the particular case where we consider the Universe to be exactly flat, (2.5) gives us this explicit expression for the metric tensor:

$$g_{\mu\nu} = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & a^2(t) & 0 & 0 \\ 0 & 0 & a^2(t) & 0 \\ 0 & 0 & 0 & a^2(t) \end{pmatrix}. \tag{2.8}$$

So far, we have made use of the Cosmological Principle to simplify the geometry of the Universe, i.e., the left-hand side of (2.4); another of its consequences is that we can treat the contents of the Universe as a perfect fluid, entirely characterized by

$$T^{\mu}{}_{\nu} = \begin{pmatrix} \rho(t) & 0 & 0 & 0 \\ 0 & p(t) & 0 & 0 \\ 0 & 0 & p(t) & 0 \\ 0 & 0 & 0 & p(t) \end{pmatrix}, \tag{2.9}$$

where ρ and p are, respectively, the *energy density* and *pressure* of said fluid.

Making use of both the FLRW metric and this energy-momentum tensor, the Einstein field equations (2.4) can be simplified, through the 00-component and the trace respectively, to yield the *Friedmann equations*:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2}, \tag{2.10}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right). \tag{2.11}$$

In the above expression, we have omitted the time dependency for readability, and introduced the usual $\dot{x} := \frac{\mathrm{d}x}{\mathrm{d}t}$ notation for time derivatives.

Let us now define two important quantities. First, the *Hubble parameter* (or Hubble rate), as

$$H(t) := \frac{\dot{a}(t)}{a(t)}. \tag{2.12}$$

By convention, we set the value of the scale factor today at $a(t_0) = 1$. This, in turn, allows us to define the *Hubble constant* as $H_0 := H(t_0) = \dot{a}(t_0)$, often parametrized *via* the *reduced Hubble constant*[2], h, as follows:

$$H_0 := 100h\,\mathrm{km\,sec^{-1}\,Mpc^{-1}}. \tag{2.13}$$

The Hubble parameter (resp. Hubble constant) gives a measure of the *rate* (resp. speed) of the expansion of the Universe at a given time t (resp. today).

Second, we can see from the first Friedmann equation (2.10) that there exists a specific value of the density for which the term related to curvature K vanishes; let us define this quantity, at present time, as the *critical density*

$$\rho_c := \frac{3H_0^2}{8\pi G}. \tag{2.14}$$

Combining (the derivative of) (2.10) and (2.11), we get, omitting time dependencies again, the following conservation law:

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0. \tag{2.15}$$

Considering (2.15), it shall behoove us to define the Equation of State (EOS), relating the energy density and pressure through a parameter w, as

$$p = c^2 w\rho. \tag{2.16}$$

[2]Note that despite similar notations, and usage that could lead to confusion, h is not in any direct way related to Planck's constant $\bar{h}$.

While we have, once again, omitted the time dependency in (2.16), w could in principle vary with time, though that is not the case for the components of our Universe we are most familiar with. On the topic of components, now is a good time to remark upon how, starting from the energy-momentum tensor in (2.4), we have, so far in this section, considered the constituents of the Universe to make up a single, perfect fluid. As can immediately be established simply by gazing around us, that is, of course, not true of our own Universe, composed of varied and different constituents - matter and light, at the very least. Much like we shall go beyond the smooth Universe we are currently describing in Section 2.1.4, we shall return to the issue of these different constituents in Section 2.1.5.

For now, let us simply consider a certain number of them, each component α having its own energy density ρ_α, pressure p_α, and EOS parameter w_α. It shall prove useful to scale the present-day[3] density by that of the critical density defined in (2.14),

$$\Omega_\alpha := \frac{\rho_\alpha(t_0)}{\rho_c}. \tag{2.17}$$

Let us now divide the first Friedmann equation (2.10) by H_0^2 and rewrite it when considering several components:

$$\frac{H^2(t)}{H_0^2} = \sum_\alpha \Omega_\alpha - \frac{Kc^2}{H_0^2 a(t)^2}, \tag{2.18}$$

which, evaluated at $t = t_0$, yields $\sum_\alpha \Omega_\alpha + \Omega_K = 1$, where we have defined

$$\Omega_K := \frac{Kc^2}{H_0^2}. \tag{2.19}$$

This allows for the curvature to be treated, after a fashion, as an extra constituent of the Universe. For a flat Universe, both K and Ω_K are equal to 0, and the Ω_α's are direct indicators of the relative contribution of each component to the overall energy density of the Universe today.

2.1.3 Redshift

Much like the sound of an ambulance speeding away from the observer is altered by the Doppler effect, the expansion of the Universe causes an alteration to the wavelength of light that reaches us. We call this alteration *redshift*: as distant objects are moving away from us, the wavelength of light in the visible domain is shifted towards the red. Let us denote λ_{obs} the wavelength at which we observe light originally emitted at wavelength λ_{emit}. Formally, we define redshift as

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}}. \tag{2.20}$$

For light emitted at time t, we have:

$$\frac{\lambda_{\text{obs}}}{\lambda_{\text{emit}}} = \frac{a(t_0)}{a(t)}. \tag{2.21}$$

[3]Much like ρ_c can be defined at any time t by replacing present-day values in (2.14) with their value at time t, the Ω_α's can be computed at any arbitrary time. In the present thesis, the ρ_c and Ω_α notations shall always refer to present-day values, despite their lacking the usual 0 subscript.

Recall that by convention, $a(t_0) = 1$; combining (2.20) and (2.21) allows us to give a simple relation between scale factor and redshift:

$$a = \frac{1}{1+z}. \tag{2.22}$$

Beneath these simple relations lies one of the reasons redshift has come to be such a central concept in cosmology. As our illustration in Figure 2.2 showed, there is a direct relationship between time t and the scale factor $a(t)$: as the latter monotonously increases over time in an expanding Universe, one can equivalently use time or the value of the scale factor at that time to determine a certain point in the history of the Universe. Similarly, (2.22) allows us to use redshift to that end. Within a given cosmological model, the redshift of an object is then also a measure of its distance to the observer: consider the observable Universe as a sphere, with us (the observer) at its center. For any angular direction on our sky, redshift can be used as the unit along the radial axis, or *depth*.

We have already introduced the comoving distance, χ, in Section 2.1.2. Let us now derive the relation between χ and z. From the FLRW metric (2.5) and (2.6), we get

$$\mathrm{d}\chi = \frac{c}{a}\mathrm{d}t. \tag{2.23}$$

By definition of the Hubble parameter (2.12), we have:

$$\mathrm{d}t = \frac{\mathrm{d}a}{aH(a)}. \tag{2.24}$$

Combining the two yields

$$\chi(a) = \int_a^1 \frac{c\mathrm{d}a'}{a'^2 H(a')}, \tag{2.25}$$

or, using (2.22),

$$\chi(z) = \int_0^z \frac{c\mathrm{d}z'}{H(z')}. \tag{2.26}$$

2.1.4 Structure formation

In Section 2.1.2, we described an homogeneous, isotropic Universe. While this is a good description of our own Universe at very large scales, it obviously no longer holds true at smaller scales - the very existence of our galaxy being a blatant statement to local inhomogeneity. As we mentioned at the beginning of this chapter, these present-day features of our Universe are believed to have originated with a brief period of inflation that occurred during the first few instants of its infancy, and allowed for small quantum fluctuations to be expanded into the seeds that would grow into the structure we observe today.

Beyond providing the link between the smooth and inhomogeneous appearance of our Universe at large and small scales, respectively, inflation theory is appealing in that it gives an elegant explanation to two other puzzling realizations we have come to make. First, the *flatness problem*: as already mentioned, observations indicate we live a fairly flat Universe, which would imply that it was *extremely* flat in the past. A period of exponential inflation would naturally lead to a dilution of the curvature, giving an explanation for the low values of Ω_K, both at present days and in the early, post-inflation Universe. Second, as we alluded to in Section 2.1.2, our observations at large scales reveal a surprising isotropy of the Universe. In the temperature of the CMB,

in particular, the anisotropies are incredibly small, including when looking at portions of the sky that should have never been causally connected! This is referred to as the *horizon problem*: while *we* can see photons from the CMB in two opposite directions of the sky, their respective regions of origin should be beyond the other's horizon, and there is thus no reason to expect their temperatures to be so similar. Once again, inflation provides an explanation: these regions only found themselves beyond each other's horizon because of it, but were causally connected before. Finding appropriate inflation models, and the search for signatures of this key component of our modern understanding of the Universe are some of the highest stakes in modern cosmology.

Let us define the *matter density contrast*,

$$\delta(\mathbf{r}, a) := \frac{\rho(\mathbf{r}, a) - \bar{\rho}(a)}{\bar{\rho}(a)}, \tag{2.27}$$

where ρ is the matter density and $\bar{\rho}$ its average at a given scale factor. When considering either early enough times, or large enough scales, the evolution of δ can be described by Linear Pertubation Theory. If we consider matter as an ideal, pressureless fluid, the evolution of δ on scales less than the horizon can be described by Newtonian physics (Peebles, 1980). One particular consequence worth highlighting now, as it will prove useful to our derivations in the next chapter, is that it then verifies the *Poisson equation*,

$$\nabla^2\Phi = 4\pi G a^2 \bar{\rho}\delta, \tag{2.28}$$

where Φ is the Newtonian *gravitational potential*. Under these assumptions, the linear evolution can further be shown to follow a differential equation,

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\bar{\rho}\delta = 0. \tag{2.29}$$

For a derivation of (2.29), and other details on linear perturbation theory, we refer the reader to standard cosmological textbooks such as the Dodelson (2003). For our purposes, it shall be sufficient to point out that from this simple expression, one can see that the Hubble parameter H will have a damping effect on the growth of structure, which perfectly aligns with our intuition: the expansion of the Universe suppresses the growth of structure.

When looking at perturbations, an extremely useful tool is the two-point correlation function (2PCF):

$$C_\delta(\mathbf{r}) := \langle \delta(\mathbf{r}')\delta(\mathbf{r}' + \mathbf{r})\rangle_{\mathbf{r}'} := \int_{\mathbb{R}^3} \delta(\mathbf{r}' + \mathbf{r})\delta(\mathbf{r}')\mathrm{d}\mathbf{r}', \tag{2.30}$$

where the second equality defines the angle bracket notation for averaging over $\mathbf{r}'$ (in what follows, we may omit the variable over which the averaging is done when no confusion is possible). Up to the sign, (2.30) looks exactly like the convolution product of the quantity of interest (δ in our case) with itself; it shall thus often be convenient to work in Fourier space. For any function $f : \mathbb{R}^n \mapsto \mathbb{R}$, let $\tilde{f}$ denote its Fourier transform:

$$\tilde{f}(\mathbf{k}) := \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}} f(\mathbf{r})\mathrm{d}\mathbf{r}, \tag{2.31}$$

where $\cdot$ denotes the vector inner product[4] and $\mathbf{k}$ the n-dimensional wave vector. The inverse Fourier transform is then

[4] That is, $a \cdot b := \sum_{i=1}^n a_i b_i = \langle a \odot b\rangle$, where $\odot$ is the *Hadamard*, or element-wise, product. In later chapters, we shall also use the notation $a \cdot b := \langle a, b\rangle$ for the inner product; in this chapter, however, we shall prefer the $\cdot$ notation to avoid confusion with the averaging brackets.

$$f(\mathbf{r}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \mathrm{e}^{\mathbf{i}\mathbf{k}\cdot\mathbf{r}} \tilde{f}(\mathbf{k}) \mathrm{d}\mathbf{k}. \tag{2.32}$$

With these and the Convolution Theorem (that states a convolution in direct space is a regular product in Fourier space), we can rewrite (2.30) as

$$C_\delta(\mathbf{r}) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \mathrm{e}^{\mathbf{i}\mathbf{k}\cdot\mathbf{r}} \tilde{\delta}(\mathbf{k}) \tilde{\delta}(-\mathbf{k}) \mathrm{d}\mathbf{k} = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \mathrm{e}^{\mathbf{i}\mathbf{k}\cdot\mathbf{r}} \left|\tilde{\delta}(\mathbf{k})\right|^2 \mathrm{d}\mathbf{k}. \tag{2.33}$$

Instead of the 2PCF itself, it is often convenient to work with the closely related *power spectrum*: the average, for a given scale $k := |\mathbf{k}|$, of $\left|\tilde{\delta}(\mathbf{k})\right|^2$ across all directions ω, i.e.

$$P_\delta(k) := \frac{1}{4\pi} \int \left|\tilde{\delta}(\mathbf{k})\right|^2 d\omega. \tag{2.34}$$

In the literature, another common (and equivalent) definition of the power spectrum of a statistically homogeneous and isotropic random field is

$$\langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}^*(\mathbf{k}') \rangle_\omega := (2\pi)^3 P_\delta(k) \mathbb{1}_{\{\mathbf{k}=\mathbf{k}'\}}, \tag{2.35}$$

where $\mathbb{1}_A$ is the (finite) indicator function[5], equal to 1 if A is true and 0 otherwise.

Going back to the perturbations of interest to us in this section, their scale dependence is often parametrized by assuming the Primordial Power Spectrum (PPS), P_P, follows a power law:

$$P_\mathrm{P}(k) := A_s \left(\frac{k}{k_0}\right)^{n_s - 1}, \tag{2.36}$$

where k_0 is just an (arbitrary) pivot scale, and A_s is the *amplitude* of the PPS and n_s is the *spectral index*, two of the main cosmological parameters in the standard model. Historically, n_s was often assumed to be equal to 1. In other words, the primordial perturbations were assumed to be entirely described by a a scale-independent Gaussian random field. This model, called *Harrison-Zel'dovich* (Harrison, 1970; Zel'dovich, 1972), is however now strongly excluded by observations (see Section 2.1.5), hence the newfound importance of n_s, a measure of how far the PPS deviates from scale invariance, as a cosmological parameter.

If the evolution of perturbation is assumed to be linear, and the initial density field Gaussian, each Fourier mode of the density contrast evolves independently from the others, and their values right after inflation (at $a = 0$) can be related to those today ($a = 1$) through the *transfer function*,

$$T(k) := \frac{\tilde{\delta}(k, a = 1)\tilde{\delta}(k = 0, a = 0)}{\tilde{\delta}(k, a = 0)\tilde{\delta}(k = 0, a = 1)}, \tag{2.37}$$

where $k = 0$ indicates an arbitrarily large scale. Computation of the transfer function is typically carried out using Boltzmann codes such as `CLASS` (Lesgourgues, 2011), though analytical fitting formulae (e.g. Eisenstein and Hu, 1998) can also be used if one can get away with lower accuracy. T includes effects from linear pertubation theory and encoded in (2.29), as well as those due to horizon crossing and to Baryon Acoustic Oscillations (BAO). When looking at cases where linear evolution can no longer be assumed (small scales and/or late

[5]The indicator function we have chosen to use here is more commonly found in the form of a Dirac delta function in the literature, though we will come to make use of $\mathbb{1}$ again in the remainder of this thesis (and of its infinite counterpart, ι, that we shall introduce in due time).

times), one must turn to other tools to capture the complex dynamics that arise in non-linear structure formation, for instance non-linear perturbation theory (a review of which is given by Bernardeau et al., 2002), or, ideally, N-body simulations (that can yield corrective terms to be added on top of the linear power spectrum, as is done e.g. by HALOFIT, Smith et al., 2003).

Before moving on to finally giving an overall description of the standard cosmological model, let us finish our study of the inhomogeneity of our Universe with a consideration of its metric (which will come in handy when we revisit the geodesic Equation 2.2 in the context of WL in Section 2.2). At first order, the perturbed metric of an expanding, non-smooth Universe can be described by the addition of the *Bardeen potentials* (Bardeen, 1980) to (2.5):

$$\mathrm{d}s^2 = -\left(1 + \frac{2\Psi}{c^2}\right)c^2\mathrm{d}t^2 + a^2(t)\left(1 - \frac{2\Phi}{c^2}\right)\mathrm{d}l^2. \tag{2.38}$$

Ψ and Φ are in general functions of both space and time, describing the Newtonian potential and the perturbation applied to the spatial curvature, respectively. In GR (and in the absence of anisotropic stress), $\Psi = \Phi$ (and the notations for Equation 2.38 are coherent with those of 2.28). Giblin et al. (2017) compare this linearization with fully non-linear GR simulations and offer corrections for the WL quantities we will introduce in Section 2.2.4.

2.1.5 The ΛCDM Model

The previous subsections provided us with the necessary concepts to establish the current standard model of cosmology we briefly presented at the beginning of the chapter. The different elements that make up this model are as follows.

Matter

The matter content of our Universe is itself divided in two different subtypes. *Ordinary matter* is made of baryonic elements (and electrons), the majority of which is found in the form of hydrogen. While everything material in our everyday lives is made up entirely of ordinary matter, its relative contribution (Ω_b) to the total energy density of matter (Ω_m) is dwarfed by that of *dark matter*, roughly five times higher. In the standard model, the latter is also assumed to be non-relativistic, that is, to be Cold Dark Matter (CDM). Their EOS parameter then verifies $w_m = w_b = 0$, leading to the following solution of the conservation law (2.15):

$$\rho_m(a) = a^{-3}\rho_{m,0}. \tag{2.39}$$

Radiation

Next are the relativistic components of our Universe, the majority of which are photons. Relative, massless neutrinos are also considered as radiation. They have an EOS with $w_r = 1/3$, leading to

$$\rho_r(a) = a^{-4}\rho_{r,0}. \tag{2.40}$$

Dark Energy

Recent observational evidence (Riess et al., 1998; Perlmutter et al., 1999) has shown our Universe's expansion has recently (on cosmological scales) started accelerating. *Dark Energy* is the name given to the origin of this acceleration. In its simplest form, Dark Energy is nothing but a cosmological constant[6], Λ. This amounts to a fluid with an EOS characterized by a (constant through time) $w_\Lambda = -1$, leading to

[6]The combination of a cosmological constant with non-relativistic dark matter gives the "concordance" model its name: ΛCDM.

$$\rho_\Lambda(a) = \rho_{\Lambda,0}, \tag{2.41}$$

in other words, an energy density that also remains constant with time. The second Friedmann equation (2.11) shows that, by definition of the EOS (2.16), any component with $w < -1/3$ would lead to an acceleration of the expansion. From that, and the evolution of the energy density parameters (2.39–2.41) of our three constituents alone, we can get a very clear intuition into this recent acceleration that ties into the three main eras in the cosmic history of our Universe: while radiation and matter originally dominated the energy content of the Universe[7], its scale factor was well behaved. As time passed, however, expansion naturally led to their density diminishing. Regardless of how small that of the cosmological constant was initially, the fact it remains constant by necessity meant it was to become dominant eventually - which in our case occurred about 4 billion years ago.

Through the definition in (2.19), we saw that we could treat the curvature as "just another component of the Universe." In the case of a Cosmological Constant, that we have here introduced as precisely one of the "regular" constituents, we can do the reverse: let us retrace the steps we took in Section 2.1.2 as follows. First, define the constant itself from its density:

$$\Omega_\Lambda := \frac{c^2\Lambda}{3H_0^2}. \tag{2.42}$$

Recall form (2.18) of the Friedmann equation:

$$\frac{H^2(t)}{H_0^2} = \Omega_m + \Omega_r - \frac{Kc^2}{H_0^2 a(t)^2} + \Omega_\Lambda, \tag{2.43}$$

or, going back to its original form (2.10), including a generic ρ for the homogeneous fluid content of the Universe:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{c^2\Lambda}{3}. \tag{2.44}$$

If we only add a similar $c^2\Lambda/3$ term to the second Friedmann equation (2.11), we arrive precisely at those we would have derived, using the Cosmological Principle as we did at the very beginning of Section 2.1.2, had we used the Einstein field equations in the following, infamous form:

$$G_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} - \Lambda g_{\mu\nu}. \tag{2.45}$$

If the two approaches - postulating the existence of a cosmological constant as in this version of the Einstein equations, or considering a fluid with $w_\Lambda = -1$ - are equivalent, the latter allows for more flexibility. As we have alluded to in the introductory chapter to this thesis, while ΛCDM is currently very well supported by observations, a cosmological constant remains unsatisfying for several reasons. Among them is the actual *value* such a constant would have: Planck Collaboration et al. (2018) gives $\Lambda = (4.24 \pm 0.11) \times 10^{-66}\,\mathrm{eV}^2$. Much like the extremely small value of K at early times led to the "flatness problem", some consider such a small value for the cosmological constant to constitute a fine-tuning problem.

The most natural way to move away from a cosmological constant is to parametrize w_Λ as a (non-constant) function of time. Without any physical insight into the nature of Dark Energy, it is common to turn to the linear parametrization of Linder (2003):

[7] In fact, as we mentioned at the beginning of the chapter, both did: the radiation-domination era, with scale factor $\alpha \propto t^{1/2}$, lasted until the Universe was a few tens of thousands years old, at which point it became matter-dominated, with $\alpha \propto t^{2/3}$.

$$w_\Lambda(a) := w_0 + w_a(1-a). \tag{2.46}$$

The cosmological constant case can then be retrieved by setting $w_0 = -1$, $w_a = 0$.

Massive neutrinos

While arguably already "beyond-ΛCDM" in the strictest sense, no overview of modern cosmology, however brief, would be complete without a mention of massive neutrinos. They naturally have an impact on cosmological considerations (they impact structure formation, for starters) and, conversely, cosmological observations can lead to constraints both on their mass (that are, in fact, extremely competitive with many terrestrial experiments) and on the so-called effective number of species, N_{eff}. Lesgourgues and Pastor (2006) give an in-depth review of massive neutrinos in cosmology, while the very recent work of Kreisch et al. (2019) delves into some ramifications, notably their potential role in some of the cosmological tensions that may be emerging from recent observational data. While we shall not concern ourselves with their impact in the present thesis, it should be noted that WL *is* sensible to massive neutrinos, and the considerations and issues we will address in this thesis (especially those related to the PSF) in light of the standard model remain of the same relevance if one's interest was centred on cosmology with (or aimed at the study of) massive neutrinos.

Cosmological parameters

We have now introduced all the basic components of our standard cosmological model. Let us now try and find the smallest number of *cosmological parameters* to describe all of them. We have already introduced the Hubble constant H_0 and the density parameters (at present day), Ω_Λ and Ω_m, of the cosmological constant and matter, respectively; we shall use these three as-is in our parametrization of ΛCDM. For the amount of baryons in the total matter content, we shall use the common parametrization $\Omega_b h^2$.

In Section 2.1.4, we have introduced the PPS, parametrized by its amplitude A_s and spectral index n_s. While we will keep the latter as a cosmological parameter, rather than A_s, let us introduce another parameter that is more commonly found in WL literature. When evaluated at $\mathbf{r} = 0$, the 2PCF of matter density today, as defined in (2.30), directly gives us the variance of the density contrast σ^2_∞, (recall that, by construction (2.27), its mean is 0):

$$\sigma^2_\infty := C_\delta(0) = \int_{\mathbb{R}^3} \delta(\mathbf{r})^2 \mathrm{d}\mathbf{r}. \tag{2.47}$$

It is easy to fathom, from this expression, that this quantity could be divergent. A way around this is to apply some filter W to the density contrast field, and to measure the variance of the resulting smoothed field $C_{\delta * W}(0)$, where $*$ is the convolution product. A natural choice for W is to use a 3D-spherical tophat of radius R, that is,

$$W_R(\mathbf{r}) := \frac{3\pi R^3}{4} \mathbb{1}_{\{|\mathbf{r}| \leq R\}}. \tag{2.48}$$

We can finally introduce our power spectrum normalization parameter of choice, that is, the smoothed variance of the 2PCF with radius $R = 8\mathrm{Mpc/h}$,

$$\sigma_8 := C_{\delta * W_8}(0). \tag{2.49}$$

The last cosmological parameter we need to make up a set well describing vanilla ΛCDM is one we have not encountered at all: the optical depth to reionization, τ. As we briefly mentioned in the introduction to this

Parameter	Value
Ω_Λ	0.6889 ± 0.0056
Ω_m	0.3111 ± 0.0056
$\Omega_b h^2$	0.02242 ± 0.00014
σ_8	0.8102 ± 0.0060
n_s	0.9665 ± 0.0038
τ	0.0561 ± 0.0071
H_0	67.66 ± 0.42

Table 2.1: Best-fit ΛCDM parameters and 68%-confidence regions, obtained from the combination of the full *Planck* data and BAO. H_0 is given in km/s/Mpc. Source: Planck Collaboration et al. (2018).

chapter, reionization occurred when photons emitted from some of the earliest stars caused neutral atoms to ionize. τ is directly linked to the time (or redshift) at which this occurs, and it is not by coincidence that it is the only of the 7 cosmological parameters we have chosen that we have not stumbled upon up to now: we built our exposition of the standard model throughout this chapter with WL, a probe of the late Universe, in mind. It is then quite natural that, despite its importance in the cosmic history, our WL observations contain very little information about the value of τ; and, conversely, we managed to build the basic foundations necessary for the rest of this thesis without giving reionization much attention. Fortunately, unlike lensing by the LSS, CMB measurements are an excellent probe into the value of τ, which is now extremely well constrained thanks to CMB experiments. Reionization itself, as well as its imprints on the CMB, are covered in standard cosmological textbooks (see, e.g., Dodelson, 2003, especially Section 8.7.2).

One of the main goals of modern observational cosmology is to achieve the most precise measurements of the base parameters in the ΛCDM model; the best such values, at the time of this writing and for our chosen parametrization, are shown in Table 2.1. Another central goal is to search for *deviations* from this standard model. This could potentially be achieved from within the framework of ΛCDM, for instance if two independent cosmological probes were to measure significantly different values of one of its parameters. Another approach is to try and constrain alternative approaches; the simplest of which is that of a time-varying EOS of Dark Energy, as shown in (2.46): any measured deviation from $(w_0, w_a) = (-1, 0)$ would prove the cosmological constant proposal wrong.

Whether to achieve either of these goals, or to explore more complicated Dark Energy or Modified Gravity models, it is clearly beneficial to have several independent and highly accurate cosmological probes at our disposal. We have mentioned the CMB and its use as a cosmological probe several times; we now turn to an entirely different one. Where *Planck* probed the early-time Universe with exquisite accuracy, more upcoming "Stage IV" experiments - *Euclid*, the Large Synoptic Survey Telescope (LSST), Wide Field Infrared Survey Telescope (WFIRST) and Chinese Space Station Optical Survey (CSS-OS)[8] - shall soon achieve the same with the later time Universe, in great part through the use of WL, which we shall now turn our focus to.

2.2 Gravitational lensing

As photons move along the geodesics of spacetime, massive objects can bend their path, much like optical lenses do, giving the phenomenon its name: gravitational lensing. It is often illustrated by analogy with the

[8]When referencing "Stages" of cosmological experiments, I aim to match the (actualized) definitions of the Dark Energy Task Force (Albrecht et al., 2006), that is, Stage I and II surveys as already completed and ongoing *then*, respectively, Stage III as "near-term, medium-cost" projects, and Stage IV as listed here.

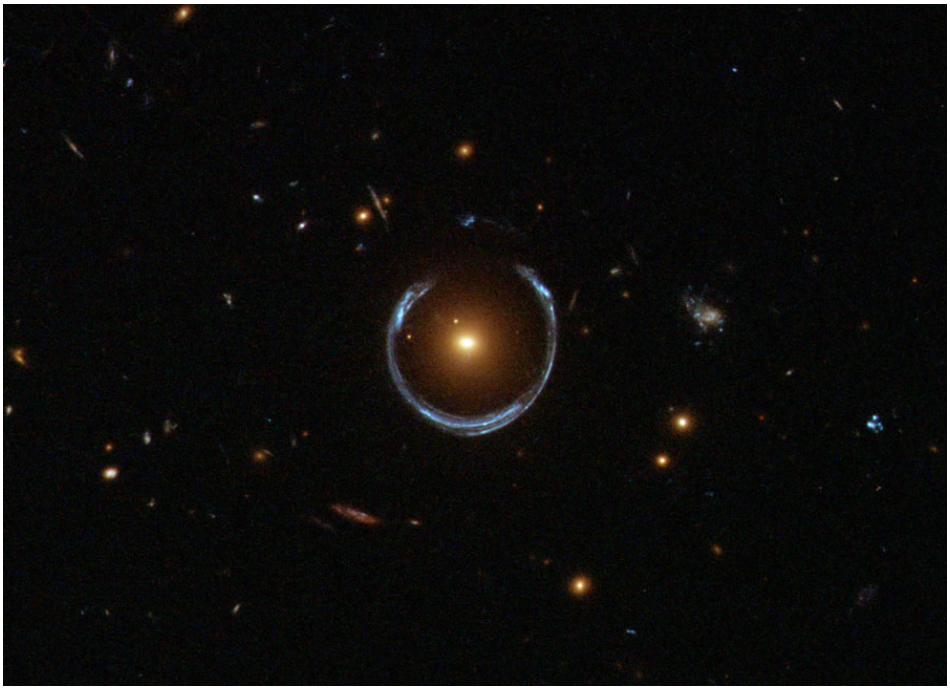


Figure 2.3: Einstein ring LRG 3-757 observed by the Hubble Space Telescope (HST). Credit: ESA/Hubble & NASA.

movement of a marble, launched in a straight direction, on a stretched sheet slightly bended by some massive object laid on it. Just over a century ago, gravitational lensing provided the earliest observational evidence of GR, when Dyson et al. (1920) measured its effect during a total eclipse.

When the right conditions are met, for instance when the observer, a massive foreground, and a luminous background galaxies are aligned, gravitational lensing can lead to spectacular displays. In an Einstein ring like the one shown in Figure 2.3, the background blue galaxy appears to wrap around the foreground red galaxy. Such effects make up what we have come to call *Strong Gravitational Lensing* (see Treu, 2010, for a review). Several hundred such objects have been found, and future wide surveys like *Euclid* and LSST are expected to bring our number of detections to the tens of thousands.

In the vast majority of cases, however, the effect of gravitational lensing is too small to be discerned by eye, and has a far smaller impact on the way galaxies appear to us than their intrinsic shape does. But while, in this regime, we cannot infer the contribution of gravitational lensing on the scale of single objects, the *statistical* study of a large number of galaxies is still a rich source of information about our Universe, its LSS, and especially on how matter is distributed within it. This constitutes the field of WL, the formalism of which we shall now introduce. We will then illustrate how to connect the main WL observable (galaxy shapes) with cosmology through the particular example of *cosmic shear* (Kilbinger, 2015), that is, the study of spatial correlations of the former.

2.2.1 Propagation of light through an in inhomogeneous Universe

To study how light propagates in an inhomogeneous Universe, following Seitz et al. (1994), let us first consider an arbitrary, fiducial light ray, $\gamma^\mu_{\rm fid}(\lambda)$. Recall from the geodesic equation (2.2) that its path is parametrized by the scalar parameter λ, such that it arrives at the observer at $\lambda = 0$. Since we are interested in how inhomogeneities impact the propagation of light, let us consider *another* light ray, $\gamma^\mu(\lambda, \boldsymbol{\theta})$, separated from $\gamma^\mu_{\rm fid}$ at the observer by the 2-dimensional (2D) angular coordinates $\boldsymbol{\theta} = (\theta_1, \theta_2)$. $\boldsymbol{\theta}$ would be the apparent position of the source on the sky if that of the fiducial ray was chosen as origin. We want to study the separation between those two rays, which depends on the contents of the Universe they went through, at any arbitrary point $\lambda > 0$. For any such, let us consider the 2D separation, $\xi(\lambda, \boldsymbol{\theta})$, projection of $\gamma^\mu_{\rm fid}(\lambda) - \gamma^\mu(\lambda, \boldsymbol{\theta})$ on the 2D screen perpendicular to the

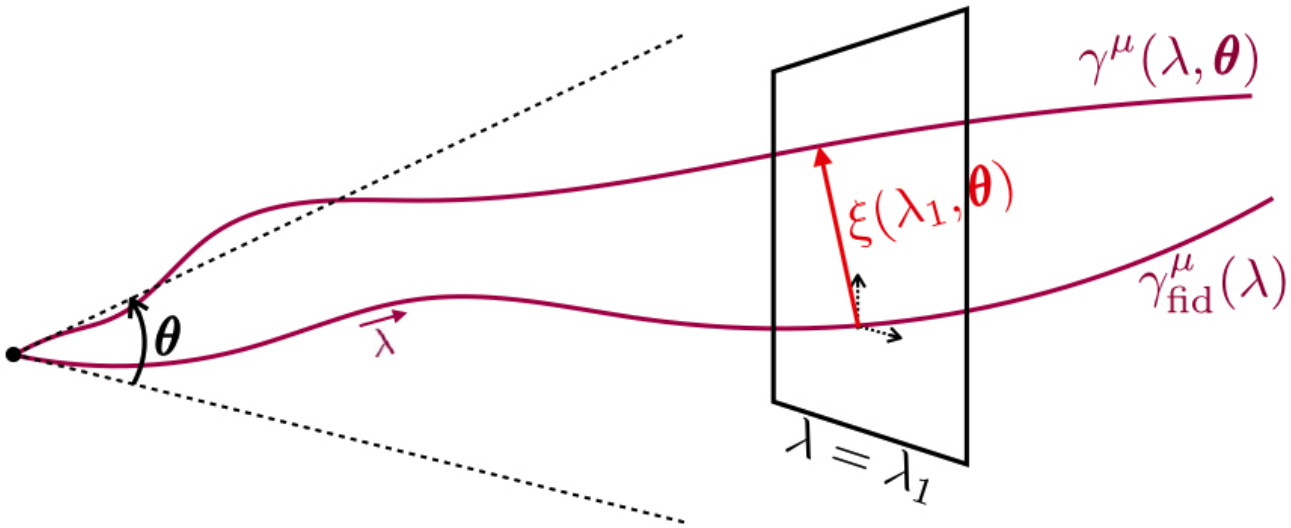


Figure 2.4: Schematic representation of the quantities of interest in the study of two light rays through an inhomogeneous Universe. The observer is at the black dot on the left. The black screen is perpendicular to the fiducial ray at an arbitrary λ_1, and the 2D separation $\xi(\lambda_1,\boldsymbol{\theta})$, in red, lies within it.

fiducial ray at λ, and tangential to the sphere of directions from the observer. A sketch of all these quantities is shown in Figure 2.4.

The evolution of ξ is described through the following *geodesic deviation equation*:

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}\lambda^2}(\lambda,\boldsymbol{\theta}) = \mathcal{T}(X_1,X_2,r)\xi(\lambda,\boldsymbol{\theta}). \tag{2.50}$$

$\mathcal{T}$ is the *optical tidal matrix*, function of physical spatial coordinates X_1 and X_2 (indicated in black dotted lines in Figure 2.4) within the tangential screen at λ, in which we omit the depedency for readability, and r, the line-of-sight component, orthogonal to said screen. In the case of our perturbed metric from equation (2.38), and if Φ can be considered a *weak field*[9] $\Phi << c^2$, Seitz et al. (1994) derive the expression of the tidal matrix,

$$\mathcal{T}_{ij}(X_1,X_2,r) = -\frac{4\pi G}{c^2}\frac{\bar{\rho}_0}{a^5}\mathbb{1}_{\{i=j\}} - \frac{1}{c^2a^2}\left(2\frac{\partial^2}{\partial X_i\partial X_j} + \mathbb{1}_{\{i=j\}}\frac{\partial^2}{\partial r^2}\right)\Phi, \tag{2.51}$$

for $i,j\in\{1,2\}$. Note that the left-hand term relates only to the (homogeneous) FLRW background and remains diagonal, while the right-hand term, called the *clump contribution*, can contain off-diagonal terms that lead to the anisotropic effects we will ultimately measure in WL studies.

Let $\boldsymbol{x} = (x_1,x_2) := \xi/a$ the *comoving separation vector*, and $\nabla_\perp := (\partial/\partial x_1, \partial/\partial x_2)$ the comoving transverse gradient in the screen perpendicular to our fiducial ray. Injecting (2.51) in (2.50) and writing it in comoving coordinates, we get

[9] An approximation of the potential of a lens with radius R and mass M is $GM/R = (c^2/2)(R_S/R)$, with R_S the Schwarzschild radius of the lens. The weak field approximation thus only breaks down for extremely compact objects, which we can afford not to concern ourselves with when looking at the gravitational lensing caused by the LSS.

$$\frac{\mathrm{d}^2\boldsymbol{x}}{\mathrm{d}\chi^2}(\chi) + K\boldsymbol{x}(\chi) = -\frac{2}{c^2}(\nabla_\perp\Phi(\boldsymbol{x},\chi) - \nabla_\perp\Phi(0,\chi)). \tag{2.52}$$

By definition of ξ (and by extension $\boldsymbol{x}$), we have the boundary conditions $\boldsymbol{x}(\chi = 0) = 0$ and $\mathrm{d}\boldsymbol{x}/\mathrm{d}\chi(\chi = 0) = \boldsymbol{\theta}$. These give us the solutions to (2.52),

$$\boldsymbol{x}(\chi) = f_K(\chi)\boldsymbol{\theta} - \frac{2}{c^2}\int_0^\chi f_K(\chi-\chi')\Big(\nabla_\perp\Phi\left(\boldsymbol{x}(\chi'),\chi'\right) - \nabla_\perp\Phi(0,\chi')\Big)\mathrm{d}\chi'. \tag{2.53}$$

In this expression, the (transverse gradient of the) gravitational potential is integrated along the perturbed light ray at $\boldsymbol{x}(\chi')$. A 0th-order approximation amounts to integrating along the unperturbed ray, $f_K(\chi')\boldsymbol{\theta}$; this is the famous *Born approximation*, shown (e.g. by Shapiro and Cooray, 2006) to be extremely accurate for WL power spectrum studies, even in the case of full-sky surveys. Under the Born approximation, (2.53) yields

$$\boldsymbol{x}(\chi) = f_K(\chi)\boldsymbol{\theta} - \frac{2}{c^2}\int_0^\chi f_K(\chi-\chi')\Big(\nabla_\perp\Phi\left(f_K(\chi')\boldsymbol{\theta},\chi'\right) - \nabla_\perp\Phi(0,\chi')\Big)\mathrm{d}\chi'. \tag{2.54}$$

Consider a source at distance χ; in the absence of gravitational lensing, it would be observed with an angular separation $\boldsymbol{\beta} := \boldsymbol{x}/f_K(\chi)$. Dividing (2.54) by the radial comoving distance f_K gives us the *lens equation*:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{2.55}$$

where we have defined the *deflection angle* as

$$\boldsymbol{\alpha}(\boldsymbol{\theta},\chi) := \frac{2}{c^2}\int_0^\chi \frac{f_K(\chi-\chi')}{f_K(\chi)}\Big(\nabla_\perp\Phi\left(f_K(\chi')\boldsymbol{\theta},\chi'\right) - \nabla_\perp\Phi(0,\chi')\Big)\mathrm{d}\chi'. \tag{2.56}$$

2.2.2 Shear and convergence

The lens equation (2.55) gives us a mapping from *lensed* (or "image") to *unlensed* (or "source") coordinates. In order to arrive at the classical lensing observables, let us now consider the linearization of this mapping. We introduce the (inverse) amplification matrix as the Jacobian

$$\boldsymbol{A} := \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}} = I_2 - \frac{\partial\boldsymbol{\alpha}}{\partial\boldsymbol{\theta}}, \tag{2.57}$$

where I_N is the $N\times N$ identity matrix. The right-hand term in the definition of $\boldsymbol{\alpha}$ (2.56) does not depend on $\boldsymbol{\theta}$ and will vanish. It is then useful to define the 2D *effective lensing potential*,

$$\psi(\boldsymbol{\theta},\chi) := \frac{2}{c^2}\int_0^\chi \frac{f_K(\chi-\chi')}{f_K(\chi)f_K(\chi')}\Phi(f_K(\chi')\boldsymbol{\theta},\chi')\mathrm{d}\chi'. \tag{2.58}$$

Up to the $\nabla_\perp\Phi(0,\chi')$ term in (2.56), $\boldsymbol{\alpha}$ is then the gradient of ψ, and we have

$$\boldsymbol{A}_{ij} := \mathbb{1}_{\{i=j\}} - \partial_i\partial_j\psi, \tag{2.59}$$

where $\partial_i := \partial/\partial\theta_i$. We can now introduce the notions of *convergence*, κ, and *shear*, γ, by parametrizing our Jacobian thus:

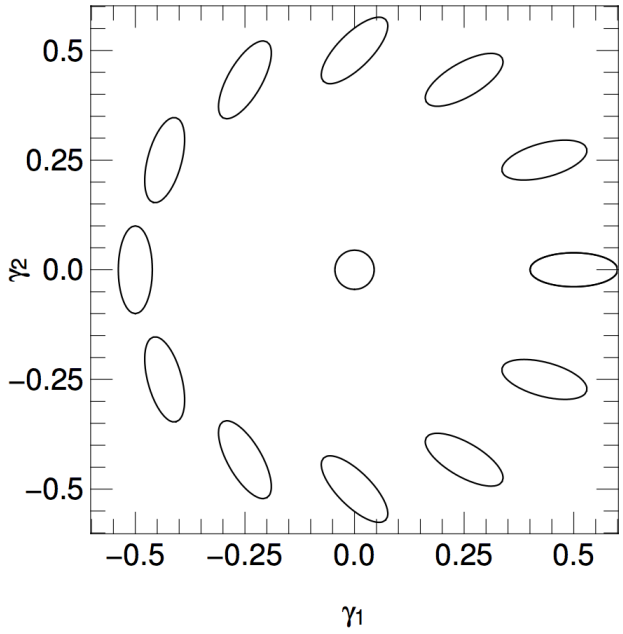


Figure 2.5: Illustration of the application of various shears. Each ellipse is a sheared version of the circular object at $(0, 0)$, shown at the position corresponding to the shear that was applied to it. Credit: Image from Kilbinger (2015), reproduced with permission.

$$\boldsymbol{A} := \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}, \tag{2.60}$$

i.e.,

$$\kappa = \frac{1}{2}(\partial_1\partial_1 + \partial_2\partial_2)\psi, \tag{2.61}$$

$$\gamma_1 = \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi, \tag{2.62}$$

$$\gamma_2 = \partial_1\partial_2\psi. \tag{2.63}$$

Convergence is the isotropic change in the observed size of an object, while shear is the anisotropic stretching due to gravitational lensing. The latter is often parametrized as a single complex number, $\gamma := \gamma_1 + \mathrm{i}\gamma_2$; its effect on a circular object is shown in Figure 2.5.

For this work, we shall focus on the alteration to the *shapes* of galaxies by gravitational lensing effects. Since κ only affects their size, let us introduce the *reduced shear*:

$$g = g_1 + \mathrm{i}g_2 := \frac{\gamma}{1-\kappa}. \tag{2.64}$$

As we shall see shortly, actual measurements of galaxy shapes in practice give us an estimate of g rather than γ itself.

2.2.3 Relation to cosmology

Now that we have introduced our main WL observables, let us establish how we can relate them to the cosmological parameters we introduced in Section 2.1.5. From the the definition of the lensing potential (2.58) and its

relationship with convergence (2.61), we can relate κ to the 2D-Laplacian of the gravitational potential:

$$\kappa(\boldsymbol{\theta},\chi) = \frac{1}{c^2}\int_0^{\chi} \frac{f_K(\chi-\chi')}{f_K(\chi)f_K(\chi')}\nabla^2\Phi(f_k(\chi')\boldsymbol{\theta},\chi')\mathrm{d}\chi'. \tag{2.65}$$

It is a reasonable assumption that $\mathrm{d}^2\Phi/\mathrm{d}\chi'^2$ would vanish when integrated along the line of sight, since negative and positive contributions would compensate each other, and we can thus replace the Laplacian in (2.65) with its 3D counterpart. This is, in turn, precisely the quantity that appears in the Poisson equation (2.28), yielding

$$\kappa(\boldsymbol{\theta},\chi) = \frac{4\pi G}{c^2}\int_0^{\chi} \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)}a^2(\chi')\bar{\rho}(\chi')\delta(f_K(\chi')\boldsymbol{\theta},\chi')\mathrm{d}\chi'. \tag{2.66}$$

If we use the matter EOS (2.39) and make the critical density (2.14) appear from the $4\pi G$ term, we get

$$\kappa(\boldsymbol{\theta},\chi) = \frac{3H_0^2\Omega_m}{2c^2}\int_0^{\chi} \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)}\frac{\delta(f_K(\chi')\boldsymbol{\theta},\chi')}{a(\chi')}\mathrm{d}\chi'. \tag{2.67}$$

We can see from this expression that the convergence is a projection of the density along comoving coordinates, with a weighting that depends on the geometry of the Universe in between source and observer.

Up to now, we have focused on a single source at an arbitrary distance χ. Let us now consider a full WL survey, probing up to limiting distance χ_{lim}. Let $n(\chi)$ denote the probabilty density function (PDF) of the galaxy sample across distances; n is more commonly referred to as the "n of z", directly related through $n(z)\mathrm{d}z = n(\chi)\mathrm{d}\chi$. The mean convergence is then

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_{\mathrm{lim}}} \kappa(\boldsymbol{\theta},\chi)n(\chi)\mathrm{d}\chi = \frac{3H_0^2\Omega_m}{2c^2}\int_0^{\chi_{\mathrm{lim}}} q(\chi)f_K(\chi)\frac{\delta(f_K(\chi)\boldsymbol{\theta},\chi)}{a(\chi)}\mathrm{d}\chi, \tag{2.68}$$

where q is the *lens efficiency*,

$$q(\chi) := \int_{\chi}^{\chi_{\mathrm{lim}}} n(\chi')\frac{f_K(\chi'-\chi)}{f_K(\chi')}\mathrm{d}\chi'. \tag{2.69}$$

Note that in this expression, our distribution of galaxies n does not depend on the direction on the sky. This no longer holds true in the vicinity of galaxy clusters, because of the proximity in between source galaxies (Schneider et al., 2002b), or between them and the lenses (Bernardeau, 1998; Hamana et al., 2002).

2.2.4 Cosmic shear

In the previous section, we established convergence as a projection of density, and thus overall mass, along a line of sight. However, since the average density contrast is zero, so is the expected value of convergence (or shear, for that matter). If we are to retrieve cosmological information, we must thus turn to higher-order statistics of our observables' distributions. The first non-trivial ones are of course the second order statistics, on which we shall focus here, in their simplest form possible. We give a brief overview of some more advanced statistics that can be used to leverage WL measurements at the end of this section.

Let us then consider the convergence power spectrum, defined similarly to (2.35) in the 2 instead of 3D case:

$$\langle\tilde{\kappa}(\boldsymbol{l})\tilde{\kappa}^*(\boldsymbol{l'})\rangle := (2\pi)^2\mathbb{1}_{\{\boldsymbol{l}=\boldsymbol{l'}\}}P_\kappa(l), \tag{2.70}$$

where the power spectrum once again depends only on the modulus of the 2D wave vector, $l := |\boldsymbol{l}|$, since like the density field δ, κ is statistically homogeneous and isotropic at large scales. Note that by working in Fourier space, we implicitly started using the important *flat-sky* approximation. Because convergence is measured in directions on the celestial *sphere*, it would be more accurate to compute the lensing power spectrum using spherical harmonics. Kilbinger et al. (2017) give both a thorough derivation in this case, as well as a study of the impact of this approximation.

By further making use of the *Limber approximation* (Limber, 1953), that is, neglecting correlations along the line of sight, Kaiser (1992) showed that the convergence and density power spectra are related through

$$P_\kappa(l) = \frac{9H_0^4\Omega_m^2}{4c^4}\int_0^{\chi_{\lim}} \frac{q^2(\chi)}{a^2(\chi)} P_\delta\left(k = \frac{l}{f_K(\chi)}, \chi\right) \mathrm{d}\chi. \tag{2.71}$$

Once again, Kilbinger et al. (2017) provide a study of the impact of the Limber approximation, and extensions thereof (LoVerde and Afshordi, 2008), on cosmological inference from cosmic shear. They conclude that working with a flat-sky is sufficient for cosmic shear in current WL surveys, though future large-scale surveys will require the consideration of a full spherical sky and a second order Limber approximation. In the present work, we shall proceed with the simple expression we obtained in (2.71).

Note that the convergence and shear power spectra are actually equal. Indeed, when taking the Fourier transform of both the convergence and complex shear, their relation to the lensing potential in equations (2.61–2.63) yield

$$\tilde{\kappa}(\boldsymbol{l}) = \frac{l_1^2 + l_2^2}{2}\tilde{\psi}, \tag{2.72}$$

$$\tilde{\gamma}_1(\boldsymbol{l}) = \frac{l_1^2 - l_2^2}{2}\tilde{\psi}, \tag{2.73}$$

$$\tilde{\gamma}_2(\boldsymbol{l}) = l_1 l_2 \tilde{\psi}, \tag{2.74}$$

which can be combined into

$$\tilde{\gamma}(\boldsymbol{l}) = \frac{(l_1 + \mathrm{i}l_2)^2}{l^2}\tilde{\kappa}(\boldsymbol{l}) := \mathrm{e}^{2\mathrm{i}\beta}\tilde{\kappa}(\boldsymbol{l}), \tag{2.75}$$

where β is the polar angle of $\boldsymbol{l}$, which in turn gives us $P_\gamma = P_\kappa$. Therefore, while it is possible to perform an "inversion" to obtain the convergence from shear measurements (for instance using Equations 2.72–2.74, as is done in the traditional method of Kaiser and Squires, 1993), we can also work directly with the 2-point statistics of shear and still retrieve the cosmological information contained in (2.71).

Our shear measurements will ultimately come from individual galaxies, which can lead to several complications when working in Fourier space, for instance the handling of masked areas where no galaxy shapes were measured. It is therefore more convenient to work in direct space, and look at the 2PCF of the shear. Instead of working directly with the complex shear, it will prove convenient to reparametrize it. For any direction $\boldsymbol{\theta}$ with polar angle φ, we define the *tangential* and *cross*-components, respectively, as

$$\gamma_t := -\mathbf{Re}\left(\gamma \mathrm{e}^{-2\mathrm{i}\varphi}\right), \tag{2.76}$$

$$\gamma_\times := -\mathbf{Im}\left(\gamma \mathrm{e}^{-2\mathrm{i}\varphi}\right), \tag{2.77}$$

where $\mathbf{Re}$ and $\mathbf{Im}$ are the real and imaginary parts of their argument, respectively. We then have three possible 2-point correlators: $\langle\gamma_t\gamma_t\rangle$, $\langle\gamma_\times\gamma_\times\rangle$, and $\langle\gamma_t\gamma_\times\rangle$. In a parity-symmetric Universe, that is, one that remains

unchanged by a mirror transformation, the latter vanishes, as γ_t is unaffected by such a transformation while $\gamma_\times$ sees its sign swapped. We can then finally introduce the two components of our shear 2PCF,

$$\xi_+(\theta) := \langle\gamma_t\gamma_t\rangle(\theta) + \langle\gamma_\times\gamma_\times\rangle(\theta), \tag{2.78}$$
$$\xi_-(\theta) := \langle\gamma_t\gamma_t\rangle(\theta) - \langle\gamma_\times\gamma_\times\rangle(\theta). \tag{2.79}$$

These are directly related to the convergence power spectrum through the *Hankel transform*,

$$\xi_+(\theta) = \frac{1}{2\pi}\int_0^{+\infty} lJ_0(l\theta)P_\kappa(l)\mathrm{d}l, \tag{2.80}$$
$$\xi_-(\theta) = \frac{1}{2\pi}\int_0^{+\infty} lJ_4(l\theta)P_\kappa(l)\mathrm{d}l, \tag{2.81}$$

where J_n is the n^{th}-order Bessel function of the first kind,

$$J_n(x) := \sum_{m=0}^{\infty}\frac{(-1)^m}{m!\Gamma(m+n+1)}\left(\frac{x}{2}\right)^{2m+n}, \tag{2.82}$$

with Γ the Gamma-function. Equations (2.80–2.81) thus give us quantities that relate the measured shear with our cosmological models.

There remains one last step to connect these to real-life WL measurements, which give us the *shapes* of galaxies rather than the shear itself. As we shall discuss further in Section 2.3.7, there are a great many definitions of what exactly constitutes the shape of a galaxy; for now, let us consider the simple case where a galaxy is assumed to have elliptical isophotes with minor axis of length a, major axis of length b, and position angle φ. Let us define its intrinsic ellipticity as

$$\epsilon^{\text{int}} := \frac{a-b}{a+b}\mathrm{e}^{2\mathrm{i}\varphi}. \tag{2.83}$$

The application of a reduced shear g (of modulus less than one) leads to its observed ellipticity being (Seitz and Schneider, 1997)

$$\epsilon = \frac{\epsilon^{\text{int}} + g}{1 + g^*\epsilon}. \tag{2.84}$$

Using these, Schneider et al. (2002a) proposed the following estimator of the 2PCF:

$$\hat{\xi_\pm}(\theta) := \frac{\sum_{(i,j)\in\Pi(\theta)} w_i w_j \left(\epsilon_{t,i}\epsilon_{t,j} \pm \epsilon_{\times,i}\epsilon_{\times,j}\right)}{\sum_{(i,j)\in\Pi(\theta)} w_i w_j}, \tag{2.85}$$

where the w_i's are weights given to each measured galaxy shape, resulting from an estimation of our uncertainty on the measurements, that arises from the steps we will detail in Section 2.3. $\Pi(\theta)$ is the set of all indices i, j such that the angular separation between galaxies i and j lies in a small bin around θ, and the tangential and cross-ellipticity components are derived the same way as their shear counterparts in (2.76, 2.77).

There are two important potential sources of bias we should mention at this point. First, while we have established the link between shear and cosmology in the last two subsections, the quantity that appears in (2.84) is not shear itself, but the *reduced* shear (2.64). Using one in lieu of the other, that is, approximating (2.84) by

$$\epsilon \approx \epsilon^{\text{int}} + \gamma, \tag{2.86}$$

can lead to significant biases in our estimation of the shear power spectrum. Corrections can reach the order of 10% of the power spectrum's amplitude at arcminute scales (Krause and Hirata, 2010; Deshpande et al., in prep).

Second, when working with galaxy ellipticities from wide surveys, it is common to assume there is no preferred intrinsic orientation of galaxies, that is, for a large number of objects,

$$\langle \epsilon^{\text{int}} \rangle = 0. \tag{2.87}$$

If this first-order assumption seems perfectly reasonable, dealing with second order quantities is trickier. From equations (2.85) and (2.86), one can see that in addition to the shear-shear correlation, terms of the form $\langle \epsilon_i^{\text{int}} (\epsilon_j^{\text{int}})^* \rangle$ and $\langle \epsilon_i^{\text{int}} \gamma_j^* \rangle$, often called "shape-shape" or II and "shape-shear" or GI, respectively, will make their appearance. These make up the problem of *intrinsic alignment*, which has recently proven to be a topic of vast interest in the WL community. Several approaches have been proposed to tackle this issue, for instance considering intrinsic alignment as an extra, "nuisance" parameter to be estimated alongside the cosmological parameters, or by using appropriate weights in the combination of shear quantities ("*nulling*"). See Troxel and Ishak (2015); Joachimi et al. (2015) for recent reviews on this topic.

In the next subsection, we will turn our attention to the practical measurement of quantities such as the galaxy ellipticities ϵ used in (2.85). With this specific and simple example, we have established how such quantities can then be used for cosmological inference. However, we should emphasize once again that WL observations can be used for several other scientific endeavors[10]. Even with the estimation of cosmological parameters in mind, WL observables can be used in many ways that go beyond the simplest of cases we have chosen to use for illustration here. We shall now give a very brief overview of some of the most significant, and highlight, once again, that the main problem we aim to address in this thesis - that of the PSF estimation - remains of similar relevance for every one of those other uses of WL data.

Using redshift information

In (2.69), all the redshift information is contained in the distribution of our galaxy sample, $n(z)$, that we integrate up to the limiting distance of our survey. This effectively makes our study of cosmic shear a 2D one. It is of course tempting to try and extend that analysis to make (more) use of the radial distribution of galaxies, which in turn would allow us to probe the LSS at different times. While 2D cosmic shear is already good at probing the overall distribution of matter (i.e. a combination of the Ω_m and σ_8 parameters), the addition of redshift information is particularly desirable if one wants to probe some other quantities. For instance, if one aims to use WL to constrain a time-varying Dark Energy EOS, as discussed at the end of Section 2.1.5, it is naturally crucial to include redshift information.

The simplest way to do so is to split the galaxies into redshift bins; this is called *tomographic* WL (Hu, 1999). This leads to one lens efficiency q_i per redshift bin, and to $n_{z\text{bins}}(n_{z\text{bins}}-1)/2$ convergence power spectra, where $n_{z\text{bins}}$ is the chosen number of redshift bins, but the overall approach remains, broadly speaking, fairly close to that we showed in the 2D case. Conversely, 3D lensing, first proposed by Heavens (2003), actually makes full use of the per-galaxy redshift estimate. While it requires more involved changes than tomography, it can also greatly improve the yield of our WL data. For instance, when using it to constrain modified gravity

[10] To name but one example, Mahdavi et al. (2007) used WL data to analyze the Abell 520 system and found what appeared to be a dark core, which could potentially be evidence against collisionless dark matter. See Okabe and Umetsu (2008); Jee et al. (2012); Clowe et al. (2012); Jee et al. (2014); Peel et al. (2017); Price et al. (2018) for further WL analysis of this object (the last two references find no significant detection of a dark core).

models, Spurio Mancini et al. (2018) recently showed that a 3D analysis of a *Euclid*-like survey will lead to a 20% decrease in errors compared to a tomographic approach.

Either of those approaches of course require an estimation of redshift for each object; in the WL context, where a very large number of galaxies are used, these are often obtained from photometric redshift (photo-z); we shall revisit this issue briefly in Section 2.3.

Derived second order functions

Beyond the standard shear 2PCF in (2.80–2.81), it can be useful to compute other, more advanced second order functions. For instance, by splitting the shear correlations into E- and B-modes (like is commonly done for CMB polarization), we would expect the latter to vanish, and it thus provides a strong tool to explore potential remaining systematics. This decomposition involves the choice of a *filter function*, several of which have been proposed in the context of WL: the top-hat (Kaiser, 1992), aperture mass (Schneider et al., 1998), optimized ring statistic (Fu and Kilbinger, 2010), Complete Orthogonal Sets of E-/B-mode Integrals (COSEBIs; Schneider et al., 2010), and spin directional wavelets (Leistedt et al., 2016).

Higher-order statistics

As a first non-trivial statistic one can derive from shear measurements, the 2PCF is both a natural and reasonable choice. Nonetheless, our WL observations contain more information than can be probed with second order statistics; in particular, higher-order statistics can probe the non-Gaussian part of the gravitational lensing signal. Other statistics have been proposed, and in some cases successfully applied to real data: the WL bispectrum Cooray and Hu (2001), peak counts (the PhD thesis of Lin, 2016, provides a great entry point and recent overview of the field), Minkowski functionals (Kratochvil et al., 2012), and recently, deep learning-based approaches (Peel et al., 2019; Merten et al., 2019). Not only can these beyond-second order approaches yield cosmological constraints that are competitive with those obtained from the 2PCF, the *combination* of both can often lead to higher constraining power than either approach would have (see, e.g., Martinet et al., 2017).

2.3 Practical weak lensing

We have established WL, in the previous section, as a rich source of cosmological information, and how the relevant quantities can be computed using measurements of a high number of galaxy shapes.

We now turn to the issue of *how* these shapes can be obtained from real-life observations. We thus give a brief overview of the critical steps involved in a *shape measurement pipeline*, that is, how one goes from raw images from a given survey to a final *shape catalog*. Each of these steps can, if not properly handled, lead to biases in the shape measurements that would then have an impact on the science performed from the output catalogs. This section is largely based on my personal experience (see Appendix B), contributing to ShapePipe (Guinot et al., in prep), a shape measurement pipeline currently being developed for the Canada France Imaging Survey (CFIS) survey.

The very first observational detections of WL by the LSS occurred in 2000 and were made independently by four different groups (Bacon et al., 2000; Kaiser et al., 2000; Van Waerbeke et al., 2000; Wittman et al., 2000). Since then, galaxy shape measurement for WL studies has been performed on many different datasets, culminating in several dedicated surveys. All these are listed in Table 2.2. At the time of this writing, the last[11] major survey to have released WL data products is the HSC survey, which provides two excellent references to get more details about the topics discussed in this section: Bosch et al. (2017) give a very detailed rundown of the main steps of their overall data processing that includes, but is not limited to, shape measurement, while Mandelbaum et al. (2017) go into more detail and provide validation for the steps central to building the final shape catalog.

[11] The fourth data release (Kuijken et al., 2019) of the Kilo Degree Survey (KiDS) came after Hyper Suprime-Cam (HSC)'s first, though

Era	Survey	Telescope/Camera	Survey ref.	Pipeline/catalog ref.	PSF
First four		WHT/PFIC		Bacon et al. (2000)	Shape interpolation
		CFHT/UH8K		Kaiser et al. (2000)	Polynomial pixel model
		CFHT/UH8K & CFH12K		Van Waerbeke et al. (2000)	Shape interpolation
		Blanco/BTC		Wittman et al. (2000)	Shape interpolation
Early applications	VIRMOS-DESCART	CFHT/CFH12k	Van Waerbeke et al. (2001)	Van Waerbeke et al. (2005)	Shape interpolation
	RCS	CFHT/CFH12k & Blanco/Mosaic	Gladders and Yee (2005)	Hoekstra et al. (2002b)	Shape interpolation
		VLT/FORS1		Maoli et al. (2001)	Shape interpolation
		Blanco/BTC & Mosaic		Jarvis et al. (2006)	PCA[a]
		WHT/PFIC & Keck II/ESI		Bacon et al. (2003)	Shape interpolation
	COMBO-17	MPG-ESO/WFI (ESO)	Wolf et al. (2001)	Brown et al. (2003)	Shape interpolation
		Subaru/SuprimeCam		Hamana et al. (2003)	Shape interpolation
	GaBoDS	MPG-ESO/WFI (ESO)	Schirmer et al. (2003)	Hetterscheidt et al. (2007)	Shape interpolation
		HST/STIS		Hämmerle et al. (2002)	Shape interpolation[b]
		HST/WFPC2		Rhodes et al. (2000)	Shape interpolation[c]
	GEMS	**HST**/ACS	Rix et al. (2004)	Schrabback et al. (2007)	Shape Interpolation[d]
Stage II	CFHTLS	CFHT/MegaCam	Hudelot et al. (2012)	Fu et al. (2008)	Shape interpolation
	COSMOS	**HST**/ACS	Scoville et al. (2007)	Schrabback et al. (2010)	PCA[e]
	CFHTLenS	CFHT/MegaCam	Heymans et al. (2012b)	Miller et al. (2013)	`LensFit`
Misc.[f]	DLS	Mayall & Blanco/Mosaic	Tyson et al. (2001)	Jee et al. (2013)	PCA
	SDSS	SDSS/Imaging Camera	York et al. (2000)	Huff et al. (2014)	PCA
	VOICE	VST/OmegaCAM	Vaccari et al. (2017)	Fu et al. (2018)	`LensFit`
	CFIS	CFHT/MegaCam	Ibata et al. (2017)	Guinot et al. (in prep)	Appendix B
Stage III	DES	Blanco/DECam	DES Collaboration et al. (2005)	Zuntz et al. (2018)	`PSFEx`
	KiDS	VST/OmegaCAM	de Jong et al. (2013)	Kuijken et al. (2015)	`LensFit`
	HSC SSP Survey	Subaru/HSC	Aihara et al. (2017)	Mandelbaum et al. (2017)	`PSFEx`
Stage IV	*Euclid*	***Euclid***/VIS	Laureijs et al. (2011)		
	LSST	LSST	LSST Science Collaboration et al. (2009)		
	WFIRST	**WFIRST**/WFI	Green et al. (2012)		
	CSS-OS	**CSS-OS**	Gong et al. (2019)		

Table 2.2: List of past, current and upcoming cosmological WL studies. When applicable, the references to both the overall survey and the actual galaxy shape measurement pipeline and/or catalog are given. For the latter, the reference given is the most up-to-date[12] at the time of this writing. The PSF model used is indicated in the last column; these are further discussed in Section 2.4.3. Space telescopes are written in bold font.

[a] An earlier processing (Jarvis et al., 2003) used an approach, close to that denoted "shape interpolation" of other early studies, based on Bernstein and Jarvis (2002).

[b] Rhodes et al. (2004) perform an independent (and better-known) analysis, complete with a tailored PSF model, wherein they use high Signal-to-Noise Ratio (SNR) stars from dense stellar fields as baseline PSF, similar to the common "shape interpolation" approach of that era, on top of which they model variations with SNR and y-axis position.

[c] Casertano et al. (2003) conducted an independent WL study of Wide Field Planetary Camera 2 (WFPC2) data; their handling of the PSF was to be described in a companion paper that remained unpublished (R.E. Griffiths, private communication).

[d] Heymans et al. (2005) provide an earlier analysis, including a detailed study of the PSF; they use a semi-time-dependent "shape interpolation" approach (that is, they split their stellar sample into 3 groups of observations taken less than 10 days apart).

[e] The original WL catalog (Leauthaud et al., 2007) was built using `TinyTim` (Rhodes et al., 2007).

[f] In this category are post-CFHT Lensing Survey (CFHTLenS) surveys where cosmological WL analysis are carried out, though they are typically not considered as "Stage III" WL surveys per se.

2.3.1 Image "preprocessing"

Since in this section, we are dealing with going from raw images to WL-relevant quantities, defining a clear line between "pre-" and "main" processing is not an easy task. Nonetheless, there are several very low-level (in the sense that they occur on the rawest, closest to telecope-outputted data) steps that are critical to successful lensing measurements, and can directly impact them (as illustrated, for instance, by Hoekstra et al., 1998, in the case of camera distortions). As an example, for Canada-France-Hawaii Telescope (CFHT) data, many of these early steps are handled by the widely used `Elixir` pipeline (Magnier and Cuillandre, 2004). See also Erben et al. (2013) for a detailed account of further steps taken (mostly within the `THELI` pipeline) on top of `Elixir`-processed images for the CFHTLenS survey.

It is naturally of crucial importance, not just for WL, but for any astronomical endeavor, to know the correspondence between an image and the part of the sky it features. Even within each image, we will need a precise mapping from pixels to real-world positions, which tends to be highly non-linear because of both the curvature of the sky and instrument-induced distortions. *Astrometry*, and especially establishing a very accurate World Coordinate System (Greisen and Calabretta, 2002; Calabretta and Greisen, 2002), are of course especially crucial for WL, where the position of galaxies will ultimately be central to the computation of shear statistics, as shown in Section 2.2.4.

Several effects related to the instrument, and especially to the Charge-Coupled Devices (CCDs) typically used in lensing surveys, are also often corrected for before moving on to the steps presented in the following subsections. These include various steps of *detrending* (flat-fielding, de-biasing, etc.) and possibly handling of other effects such as crosstalk or saturation. Estimation and subtraction of the background can also sometimes be performed at these early stages, though it is now well-established (Martinet et al., 2019) that this issue should be approached with extreme care as it can have a substantial impact on shear measurements.

Cosmic rays and satellite trails need to be addressed, as they would otherwise cause severe biases in a lensing analysis (for instance if falsely identified as extremely elongated galaxies). If single exposure images are to be used for the analysis, they should thus be identified and removed (masked) from the analysis (see, e.g., Desai et al., 2016, for an approach applied to Dark Energy Survey – DES – data).

An alternative way to deal with cosmic rays and satellite trails is the use of *stacking*, that is, the combination of several single exposures that feature the same part of the sky into a single image. Stacking comes with several other appealing consequences: as it allows to go deeper, it increases the number of detected objects; it reduces noise levels, which increases the quality of galaxy images and should thus, in principle, help with the shape measurement process; it can fill in the gaps in between the CCDs; it can help deal with undersampled images (Lauer, 1999; Fruchter and Hook, 2002; Rowe et al., 2011). Stacking however has one major drawback, especially in the case of shape measurement, linked to the core of this thesis: the handling of the PSF. We have already mentioned that it is of the utmost importance for WL - a point that we shall establish formally in Section 2.4, though it is natural to get a sense of why this would be the case, since the PSF can and does alter the very shapes we aim to measure. A single galaxy is affected by a different PSF on each single exposure it appears on, because of changing atmospheric contributions and/or because it falls at a different position within the focal plane each time. It then raises the question of how to combine the PSF models from each exposure into a single, final PSF to use in order to correct the shape measured on the stacked galaxy image. One cannot simply build a "stacked" PSF model by applying a standard PSF modelling approach directly on the stacked images, since each time a different set of pointings are combined, the resulting PSF model is different. This is illustrated in Figure 2.6 with data from CFIS: the PSF model across the central "tile", that combines stacks with different sets of available single exposures, has a great many discontinuities.

---

the associated WL products had not yet been released.

[12]Note this does not mean it is the *last* WL paper published for this particular survey/dataset. For past surveys, there typically are several more papers with cosmological analysis and re-analysis, which sometimes include larger areas than those used in the reference cited in the table (e.g. Hoekstra et al., 2002a, in the case of RCS), or the combination of several different surveys (e.g. the "100-deg^2" survey of Benjamin et al., 2007). However, when the shape measurement remains largely unchanged, I chose to cite the earlier paper that gives the most detailed account of the shape measurement pipeline.

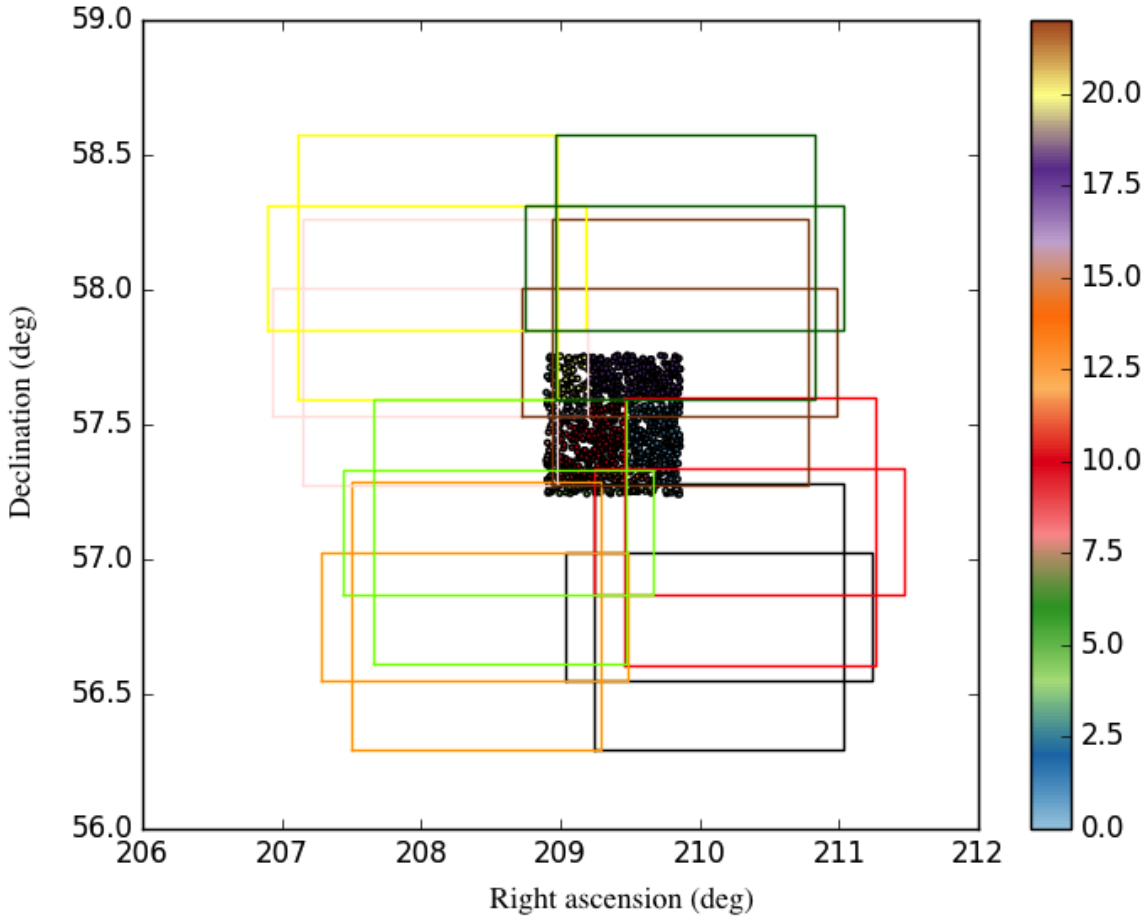


Figure 2.6: Stacking-induced discontinuities in the PSF. Each combination of two same-colored rectangles corresponds to one MegaCam pointing; the scattered objects in the middle are those detected on a single "tile", that is, a stacked image using those pointings. The color of each object corresponds to a unique combination of contributing pointings, which in turn implies the application of a different (stacked) PSF.

The PSF model therefore needs to be fitted on single exposure images. There are then three main approaches to make the most use of the multiple available exposures for shape measurement. The first one is simply not to use stacked images for shape measurement, though detection (which we shall return to in Section 2.3.3) should still be performed on these. The multiple images of the galaxy can still be leveraged to get the best possible shape measurement by simultaneously fitting each single-exposure galaxy/PSF pair of the same object. This typically implies the need for a *model-fitting* shape measurement method (see Section 2.3.7), and is for instance currently the chosen approach in the KiDS survey (Kuijken et al., 2015). If the same object is observed through several filters, shape measurement can also be improved in a similar fashion by using each of these images, as is done for instance in the Dark Energy Survey (DES) survey (Zuntz et al., 2018).

Another approach is to perform the shape measurement on the stacked image, which requires a careful combination of the single-exposure PSF models on a per-object basis. This can, in turn, be done in two ways. The first one, that seems to have fallen somewhat out of favor in the WL community in recent years, is to apply extra, well-chosen convolutions to each image so that they are all effectively affected by the same PSF, often chosen to be circular. This was, for instance, the chosen approach of Huff et al. (2014) in their WL analysis of the Sloan Digital Sky Survey (SDSS) data. Lastly, the individual PSF models can be evaluated for each object, then combined into a single PSF model, to be used with the stacked galaxy image, within any standard shape measurement approach. This requires the PSF combination to be done in the exact same fashion as that of the galaxy stacking; in particular, care should be given to the weights allocated to each single PSF. Historically, this approach has been used to combine only the PSF model quantities that were required for the correction, typically some measures of the *shape* of the PSF. This was fairly common in early studies (first two "eras" of our Table 2.2), and was also often used for space-based WL measurements with HST (e.g. by Schrabback et al., 2007). More recently, the full PSF models, sampled on small images ("postage stamps") are combined to obtain an actual stacked PSF with as many pixels, which can then be used with any shape measurement approach. This is currently done by Bosch et al. (2017) for the HSC survey.

Lastly, several *detector effects* can be taken into account at this stage. Their impact on shape measurement are somewhat similar to those of the PSF, though they are typically non-convolutional and can be flux-dependent. It is therefore impossible to simply consider them as extra effects that contribute to the final PSF and correct for them in a single step. In this work, we will always assume they have been modelled and corrected for prior to the PSF modelling stage.

Chief among these detector effects is surely the Charge Transfer Inefficiency (CTI) phenomenon that affects all CCD devices, but can become especially significant for space-based telescopes as the harsh conditions of their environment tend to lead to increasing levels of CTI throughout the lifespan of the instrument. Massey et al. (2014) proposed a modelling approach and successfully applied it to HST; the continuation of these efforts is currently ongoing specifically for application to *Euclid* (Israel et al., 2015, 2017; Nightingale et al., in prep).

Second is the Brighter-Fatter effect (BFE), wherein the size of brighter objects tends to be larger. This is, of course, particularly relevant to WL and especially PSF modelling, since the latter is typically performed using bright stars. Coulton et al. (2018) recently studied and proposed a scheme to correct for this effect, with application to HSC data.

Lastly, Boone et al. (2018) recently discovered an effect on CCD readout electronics, where a binary offset, that depends on the number of 1s in the binary representation of a "driver" pixel, occurs for the few pixels read out immediately after. It is likely one that should be considered and carefully dealt with for WL, when present.

Note that several of the steps we mentioned in this subsection can actually be improved once the others have been performed; in fact, this circularity extends to the PSF modelling step, as we shall briefly discuss in Section 2.3.5. In practice, an iterative approach can thus prove beneficial, where the corrections or calibrations discussed here are repeated on data that has already been "cleaned" from all other effects.

2.3.2 Masking

The areas covered by WL surveys are naturally becoming larger and larger, which necessitates an increasing degree of automation in the data processing pipelines. In particular, objects that appear in science exposures

and cannot be used for WL analysis need to be identified and "masked" (i.e., removed from all further analysis). This masking step itself requires automation, as it has become impossible to expect human eyes to identify and demarcate all undesirable objects.

Among everything that should be masked for a successful WL analysis, the three most significant are bad pixels, saturated stars, and nearby galaxies (see Appendix B for an anecdotal illustration of the latter's importance).

2.3.3 Object detection

Once again, the sheer size of WL surveys calls for an automated detection of both stars and galaxies. This is also typically the stage at which "postage stamps", small images centered on a particular object, are extracted and used in further analysis. This is most often performed, for WL as it is for many other types of studies on wide optical astronomical data, with Bertin and Arnouts (1996)'s celebrated `SExtractor` software. Beyond identifying individual objects and extracting postage stamps, `SExtractor` performs many other tasks such as measuring several quantities, on those objects, that will prove useful to later steps of the shape measurement pipeline (e.g. noise levels, individual SNRs, object sizes).

2.3.4 (De)blending

Another crucial information provided by `SExtractor` is the identification of *blended* objects, that is, cases where two or more distinct objects are close enough, when projected on the sky, to overlap and appear in one another's postage stamp. `SExtractor` identifies such cases, and offers tailored masks that identify which pixels are affected by the contaminating additional object(s). These can in turn either be removed from the samples entirely, or treated as stamps with several masked out pixels (which can, of course, strongly hinder the shape measurement process).

While this is an extremely powerful first order tool to deal with blends, a more advanced treatment of this issue will likely be needed for Stage IV WL surveys, especially for the ground-based LSST. As our surveys go deeper, blending naturally occurs more frequently, and experience gained on recent surveys clearly identify this issue as one of the most major for WL goals: Bosch et al. (2017) estimate that $\mathbf{58}\%$ of objects identified in the HSC wide survey were recognized blends!

Several advanced methods to deal with the identification and separation of blended objects have recently been proposed (Joseph et al., 2016; Melchior et al., 2018; Boucaud et al., 2019). These works illustrate that the use of colour information makes this complicated issue considerably easier to tackle. While the Euclid VIS observes in only one, very wide band, its spaceborne nature means it is far better at resolving objects. The (de)blending issue in WL therefore makes for great potential synergy between ground- and space-based future missions, especially between LSST and *Euclid* (Rhodes et al., 2017; Schuhmann et al., 2019).

2.3.5 Star-galaxy separation

While we have already addressed several artefacts and other potential sources of false detections that could occur during the process of Section 2.3.3, there still remains an important distinction to be made between those objects that remain: galaxies and stars. The former will of course be used to measure shapes and, eventually, cosmic shear or other cosmologically-relevant quantities. Stars tend to be both more circular than galaxies, and much *closer* (most within the Milky Way). It is therefore important not to misidentify stars as galaxies and treat their measured shapes as if they carried some sort of cosmological lensing information. Even among those objects identified as galaxies, it might be preferable to remove some from the sample that will be used for shape measurement. A good illustration of this is given by the DES Year 1 shape catalog (Zuntz et al., 2018), where two different shape measurements methods required selecting two different galaxy catalogs.

Conversely, we want clean samples of stars to fit our PSF models on; leftover galaxies within those sample would bias our PSF model. In fact, given the importance of the latter in WL studies, it is common to perform a

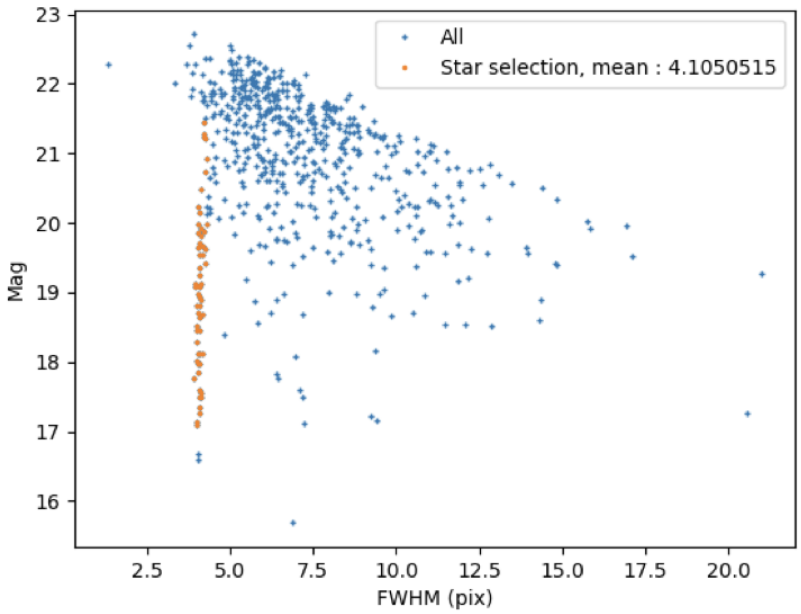


Figure 2.7: The size-magnitude diagram for a single CCD from the CFIS survey. Each scattered point is a detected object; those within the identified stellar locus are shown in orange. Both size, measured through the full width at half-maximum (FWHM), and magnitude are from `SExtractor`. Credit: A. Guinot.

further, more conservative selection of "PSF stars" for this purpose specifically, than that done for others, such as photometric measurement. For instance, the DES PSF star selection made by Zuntz et al. (2018) is different from what has been done for photometry (Drlica-Wagner et al., 2018). The most common approach is to make use of the *size-magnitude* diagram (or diagrams made from proxies of those two quantities). While galaxies come in a variety of sizes, all stars (or rather, all those unaffected by the BFE, unresolved, and unsaturated) should be of roughly the same size, namely that of the PSF. This defines the *stellar locus*, a fairly straight line within the size-magnitude diagram, the identification of which amounts to a star selection. An example is shown, from our own processing of the CFIS data, in Figure 2.7. For KiDS, Kuijken et al. (2015) refined this approach by also using the fourth-order moment measured on the object's postage stamps (see Section 2.3.7 for a definition of lower-order moments).

Another approach is the use of external star catalogs from dedicated surveys. This comes with strong confidence on the (close) lack of contamination by non-star objects, and is being made increasingly more seemly by current successive Gaia data releases (Brown et al., 2018).

This classification problem could be approached by advanced methods; a neural network is included in `SExtractor` (if ran in conjunction with the companion, PSF modelling software `PSFEx` of Bertin, 2011, which we shall return to in Section 2.4.3), though it is generally not used for PSF star selection (see the different DES star catalogs Drlica-Wagner et al., 2018; Zuntz et al., 2018, for illustration).

At the end of Section 2.3.1, we mentioned how the steps described there could benefit from an iterative, repeated processing. The same is very true for star-galaxy separation, as once a PSF model has been established, objects that appear very different from it are unlikely to be stars. Historically (e.g. in several studies of the "Early applications" era of Table 2.2), this notion was often used by performing shape "clipping", that is, by removing from the star sample all objects whose size or ellipticity differed from that of the PSF model, then refitting the model on surviving stars. Nowadays, where full-pixel PSF models are more common, bad objects can be iteratively removed from the star sample by goodness-of-fit tests, typically through a cut-off in χ^2. Such an approach is, for instance, natively handled by `PSFEx`.

2.3.6 PSF modelling

We have now established how a clean star sample has been selected from raw images. Section 2.3 gives an overview of the critical steps in a shape measurement pipeline. The PSF modelling process would typically occur at this stage, but given the nature of this work, it requires a more in-depth treatment. We thus postpone these considerations and refer the reader to Section 2.4 and subsequent chapters.

2.3.7 Galaxy shape measurement

Assuming we obtained an appropriate PSF model for each of the galaxies we selected in Section 2.3.5, we can finally turn to the practical measurement of shapes. This implies addressing two main issues. First, we need a definition of what exactly constitutes the shape of a galaxy. Second, we need to find a way to obtain this measurement while correcting for the influence of the PSF.

Defining ellipticities

In (2.83), we defined the ellipticity of a galaxy using the semi minor- and major-axes of its supposed elliptical isophotes. We then connected this quantity to the action of reduced shear through (2.84). Of course, real-life galaxies have complex brightness profiles, whether it is because of different components (e.g. a bulge and a disk) of highly different shapes, of intricate spiral arms, or simply, in the case of irregular galaxies, because they are in no way ellipsoidal. They are thus nowhere close to having elliptical isophotes, and the intuitive and geometric definition of (2.83) is of no practical use.

Let us introduce a definition of "ellipticity" that is computable on any brightness profile, $I(x_1, x_2)$, where x_1 and x_2 are arbitrary orthogonal coordinates (in practice, they will typically be chosen to move along the $x-$ and $y-$axes corresponding to the pixel grid on our discrete images). Let Q_{ij} denote the (centered) *quadrupole moments* of our image I. Q is defined as

$$Q_{ij} := \frac{\int_{\mathbb{R}^2} I(x_1, x_2)(x_i - x_i^c)(x_j - x_j^c)\mathrm{d}(x_1, x_2)}{Q_0}, \tag{2.88}$$

for $i, j \in \{1, 2\}$, where

$$Q_0 := \int_{\mathbb{R}^2} I(x_1, x_2)\mathrm{d}(x_1, x_2) \tag{2.89}$$

and the two

$$x_i^c := \frac{\int_{\mathbb{R}^2} I(x_1, x_2)x_i\mathrm{d}(x_1, x_2)}{Q_0} \tag{2.90}$$

are called the image *centroid*. Note a strictly equivalent definition sometimes found in the litterature uses

$$\mu_{ij} := \int_{\mathbb{R}^2} I(x_1, x_2)(x_1 - x_1^c)^i(x_2 - x_2^c)^j\mathrm{d}(x_1, x_2), \tag{2.91}$$

which directly leads to

$$Q_0 = \mu_{00}, \tag{2.92}$$
$$\mu_{10} = \mu_{01} = 0, \tag{2.93}$$
$$Q_{11} = \mu_{20}/\mu_{00},\ Q_{12} = Q_{21} = \mu_{11}/\mu_{00},\ Q_{22} = \mu_{02}/\mu_{00}. \tag{2.94}$$

Where, by definition of the centroid, the first order centered moments (μ_{01}, μ_{10}) vanish. This notation has the advantage of illustrating how characterizing the ellipticity amounts to the knowledge of 6 individual quantities: either the 6 moments of order up to 2 or, equivalently, (a specific measure of) its flux, its two centroid coordinates, and three second order moments.

We can now define the ellipticity of our image I as

$$\epsilon := \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}}. \tag{2.95}$$

This definition of ellipticity is a generalization of (2.83), since they coincide in the particular case of objects with elliptical isophotes. The action of a reduced shear (2.84) also holds true regardless of the initial image's profile I (up to its quadrupole moments being defined).

Another definition commonly found (perhaps even more so than that of Equation 2.95) in the lensing literature is

$$e := \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22}}. \tag{2.96}$$

The two are directly related through (Bartelmann and Schneider, 2001)

$$\epsilon = \frac{e}{1 + \sqrt{1 - |e|^2}}, \; e = \frac{2\epsilon}{1 + |\epsilon|^2}, \tag{2.97}$$

and analogous to (2.84), the action of a reduced shear g on an object with intrinsic ellipticity e^{int} is

$$e = \frac{e^{\mathrm{int}} - 2g + g^2 e^*}{1 + |g|^2 - 2\mathbf{Re}(ge^*)}. \tag{2.98}$$

The two notations can therefore be used fairly interchangeably, as long as one remains self-coherent (i.e. avoids mixing-and-matching of quantities computed with different definitions).

Another useful quantity is a measure of the *size* of an object, defined from its quadrupole moments Q_{ij}, as

$$R^2 := Q_{11} + Q_{22}. \tag{2.99}$$

In the case of an object that follows a parametric profile, this quantity is directly related to the size parameters. For instance, for a 2D Gaussian profile

$$I(x_1, x_2) := A \exp\left(-\frac{(x_1 + x_2)^2}{2\sigma^2}\right), \tag{2.100}$$

we have $R^2 = 2\sigma^2$. Much like equations (2.95–2.96) gave us ellipticity definitions applicable on any arbitrary profile, R^2 is a measure of size that can (in principle) be computed on any object. Throughout this thesis, when referring to stars or PSFs, the term "shape" will mean the set of both ellipticity e and size R^2 of the object.

Before we can use these definitions to measure ellipticities, two outstanding issues remain to be addressed. First, images will be discretely sampled on a finite set of pixels, P. Second, noise will be present in our images. Dealing with the discrete sampling may seem straightforward enough at first: we only have to replace the integrals in the definition(s) of our image moments with a discrete sum over the set of available pixels P. For instance, our definition of μ_{ij} (2.91) would then become

$$\mu_{ij} := \sum_{(x_1,x_2)\in P} I(x_1,x_2)(x_1 - x_1^c)^i (x_2 - x_2^c)^j. \tag{2.101}$$

We shall henceforth use the same notations Q_{ij}, μ_{ij}, and x_i^c to refer to either the continuous or discrete versions of these quantities.

In practice, our measurements will be made from postage stamps centered on the objects of interest (see Section 2.3.3). This means the set P is finite, which is not undesirable, as the pixel intensity actually emanating from objects of interest in a WL analysis (distant galaxies and unresolved stars) is likely to go below noise level after a small number of pixels, and it limits the number of other objects whose own profile impacts the pixel values. However, the quadrupole moments (2.88) are quantities of circular symmetry, while our postage stamps are square. This leads to truncation effects that can be quite significant when measuring ellipticities from (discrete) quadrupole moments on a postage stamp, especially in light of the fact that by construction, the first ellipticity component will likely correspond to elongation along the x and y axes of our stamps, while the second to those in the diagonal directions, i.e. where the truncation occurs the closest and farthest, respectively.

Second, quadrupole moments are notoriously sensitive to noise. It is easy see why by looking at their definition (2.88): the furthest away from the center, the stronger the influence a single pixel will carry (because of the term in x_1^2, x_1x_2, or x_2^2), yet the more its value will be driven by noise. Direct measurements of ellipticities from quadrupole moments are thus strongly noise-dominated in practical applications.

Both of these considerations are alleviated by the addition of a *window function*, similar to how we solved the problem of the power spectrum normalization in Equation 2.1.5. Let W denote such a function; we use the same notations to define *weighted* quadrupole moments and related quantities, for instance

$$\mu_{ij} := \int_{\mathbb{R}^2} I(x_1,x_2)W(x_1,x_2)(x_1 - x_1^c)^i (x_2 - x_2^c)^j \mathrm{d}(x_1,x_2). \tag{2.102}$$

The addition of this window function naturally biases our estimation of the ellipticity - we shall return to this issue in Section 2.3.8 and Section 2.4.1, as it arises in the case of galaxies and PSFs, respectively. Moreover, it raises the question of the choice of W. It is typically chosen to be a 2D Gaussian centered on the object's centroid, which still leaves the question of its size to be addressed. Historically, this size would often be fixed a priori, and chosen to be the same for measurements on both galaxies and the stars used to model the PSF. This eventually proved to cause issues, especially as WL studies started being carried out with space instruments, where the moments measured from the (mostly optical) PSF are highly dependent on the chosen size for the window function (Hoekstra et al., 1998). More recently, Hirata and Seljak (2003); Mandelbaum et al. (2005) suggested *matching* the size of the window function to that of the object being measured - we shall refer to this approach as Hirata-Seljak-Mandelbaum (HSM), and use it for all direct (that is, with no PSF correction) shape measurements derived from quadrupole moments for the rest of this thesis. Let us note that, as proposed by Melchior et al. (2011); Viola et al. (2011), a natural extension is to make the window function match the object not only in size, but also in ellipticity, by iteratively shearing it.

"Shape measurement," or, how to correct for the PSF

The effect of the PSF is that of a convolution, i.e. if an object had an initial (in image space, that is, possibly already gravitationally sheared) brightness profile I, and was affected by a PSF H, the observed image would be

$$Y = I * H, \tag{2.103}$$

where $*$ is the convolution product. Removing the effect of the PSF therefore amounts to the task of *deconvolution*; a very well-studied topic in image processing (see, e.g., Starck and Murtagh, 2007). However, the

needs that come with a WL-centric analysis make it impossible to directly apply this plethora of tried-and-tested deconvolution tools, as they come with no guarantee that the *shape* of objects, as we just defined them, remains unchanged through the process. This led to new approaches being developed, specifically with lensing in mind, as early as the mid-1990s. Broadly speaking, these *shape measurement* methods (which should be understood as including the removal of PSF effects, rather than just measuring the shape of an object taken on its own, as can be done using the weighted moments approach) fall into two categories.

Moments-based methods make use of the quadrupole moments, as defined in (2.88), of both the observed galaxy and the corresponding PSF, and correct for the latter directly in moments space. The classical example of such an approach is the famous Kaiser-Squires-Broadhurst (KSB), first proposed by Kaiser et al. (1995). Throughout the years, several variants and improvements of KSB were proposed, for instance by Luppino and Kaiser (1997); Hoekstra et al. (1998); Kaiser (2000); Viola et al. (2011), and used on most of the early cosmic shear studies; see Table 2.2, where all entries where the PSF is modelled through "shape interpolation" used moments-based shape measurement approaches.

Model-fitting methods, first proposed as early as the end of last century by Kuijken (1999), makes use of the full PSF image instead of just its measured moments. An analytical profile for the galaxy is chosen, for instance a Sérsic profile (Sérsic, 1963). The parameters of this profile (plus those used in an extra shearing process if the profile is circular) are then fitted by convolving it with the PSF, then sampling it at the same pixel scale as the observations and comparing them. Because of this process, model-fitting methods are also often referred to as *forward-fitting*. This class of approaches had come to become the preferred one up to very recently, and has been used on several recent surveys in the following incarnations: `LensFit` (Miller et al., 2013) in CFHTLenS and KiDS (Kuijken et al., 2015), `im3shape` (Zuntz et al., 2013) and `ngmix` (Sheldon, 2015) in DES (Zuntz et al., 2018).

Both approaches come with their own advantages and pitfalls. Moments-based methods are typically less computationally demanding, as model-fitting often requires iteratively solving an inverse problem (see Section 3.1.1 for an introduction). As we already discussed above, the former come with the issue of having to select a window function, therefore introducing bias; the latter, with the need to select a model for the galaxies, which in turn introduces *model bias* (as real-world galaxies do not quite follow simple analytical profiles). Those two sources of bias have long been known to be very similar in the lensing community; with the correct formalism, one can in fact show they are mathematically identical, as done by Simon and Schneider (2017).

A lot of effort has been made by the lensing community to compare different shape measurement methods, originally in the hopes of determining the best-suited, or at least understanding their differences. To this end, and to investigate other aspects of practical WL measurements, a series of challenges were organized, starting with the Shear TEsting Program (Heymans et al., 2006; Massey et al., 2007), followed up by the GRavitational lEnsing Accuracy Testing (GREAT) series, each of increasing complexity: GREAT08 (Bridle et al., 2009, 2010), GREAT10 (Kitching et al., 2011, 2012), which also contained a "Star Challenge" to compare PSF interpolation methods (Kitching et al., 2013), and the latest to date, GREAT3 (Mandelbaum et al., 2014, 2015). This quest for the best per-object estimate of shape has, very recently (and in part thanks to the experience gained from these challenges), been replaced by a new paradigm (Mandelbaum, 2018): the outputs of the shape measurement method itself is of little import, as long as the actual *shear* estimates we will derive from it (typically through some sort of averaging, e.g. in a 2PCF) are viable. In other words, rather than having an accurate shape measurement, we would rather have one that we can properly *calibrate*.

Before turning to how this calibration is carried out in practice in the next subsection, it is worth mentioning two more shape measurement approaches. Neither really fits into either the moments-based or the model-fitting categories. Moreover, rather than attempting to measure the shapes of individual objects, then estimate the contribution of gravitational lensing-induced shear from these, they instead directly attempt to measure the latter, in turn bypassing many of the issues we have presented in this subsection. The first is the Bayesian Fourier Domain (BFD) approach of Bernstein and Armstrong (2014) (see also Bernstein et al., 2016, for extensions). The second relies on machine learning algorithms trained on simulations; two different approaches have been proposed by Springer et al. (2018) and Tewes et al. (2019). The latter and BFD are both planned to be used,

along with more classical methods of the two classes described above, within *Euclid*.

2.3.8 Calibration

For an inferred shear $\hat{g} = (\hat{g}_1, \hat{g}_2)$, let us extend the bias resulting from all the steps we have outlined in the present section to first order,

$$\hat{g}_i \approx (1 + m_i) g_i + c_i, \tag{2.104}$$

where $i \in \{1, 2\}$, g_i is the true shear component, and m_i and c_i are the *multiplicative* and *additive* biases, respectively. The aim of the *calibration* step is to both evaluate and correct for these terms. The bias incurred on a specific galaxy depends on many of its properties, so it is important to estimate m and c as functions of these properties, rather than just single constants across the whole catalog. The dependency on the *size* of galaxies is the most obvious of these, so it has been common in calibration schemes to estimate different m and c values per object size bin. The requirements on future surveys will likely require a calibration scheme that varies with a great many parameters. Pujol et al. (2017) carried out an in-depth study of several of these. See also Hoekstra et al. (2015, 2017); Martinet et al. (2019), where the authors explore many sources of bias (especially, in the latest of the series, that of undetected galaxies).

In practice, until the very recent proposal for an alternative approach, the estimation of m, c was carried out using image simulations. These are made to be as realistic, and include as many real-world effects, as possible. The entire shape measurement pipeline is then ran on those simulations, and since we have access to the inputted shear values, we can actually compute the bias parameters from (2.104). When building those simulations, an invaluable resource is the freely available GalSim[13] software (Rowe et al., 2015), the original development of which started as a byproduct of the creation of the GREAT3 dataset (though it has since then far outgrown this initial purpose). The latest practical calibration of WL measurements through simulations was carried out by Mandelbaum et al. (2018) for the HSC survey.

Calibration through simulations comes with two major caveats. The first one is the computational cost associated with generating enough simulations to achieve sufficiently accurate calibrations; Pujol et al. (2019) recently proposed an approach to drastically reduce that number. The second is that the ultimate calibration scheme we get from simulations will only ever be as good as the simulations themselves, and it is of course hard to gauge exactly *how* realistic a particular simulation is. There are many specific points that could cause concern in this area; a good example is the profiles used to simulate the galaxies. It is well-known that analytical profiles are far too simplistic to accurately describe the intensity profile of real-life galaxies. Worse, if we use model-fitting methods, we would end up using the same (or very similar) profiles both to generate and measure the shapes of galaxies, therefore entirely missing the model bias that would definitely occur in applications to real data. A way around this issue is to use images of real galaxies, taken with an instrument with higher resolution, and reprocess them (including the application of the shear to be measured) so they match what would be observed by the survey at hand. An example is shown in Figure 2.8, where I have taken a galaxy observed with HST's Advanced Camera for Surveys (ACS), and processed it so it appears as it would if observed with *Euclid*'s VIS instrument. While we have no information on the intrinsic shape of the galaxy, we can still use many of them to infer shear and perform calibration. While this approach frees us from the lack of realism introduced by using analytical galaxy profiles, it comes with its own limitation, namely the number of real galaxies observed, much lower than that we aim to observe with future surveys. This can (and should) be somewhat alleviated by using several different "realizations" of the same real object, including by rotating the original profile (which should be done regardless, to remove the impact of intrinsic galaxy shapes). Another promising and far more flexible option has recently been proposed by Ravanbakhsh et al. (2017), who use generative networks trained on real data to create any arbitrary number of galaxy profiles that are indistinguishable from those we observed.

[13] https://github.com/GalSim-developers/GalSim

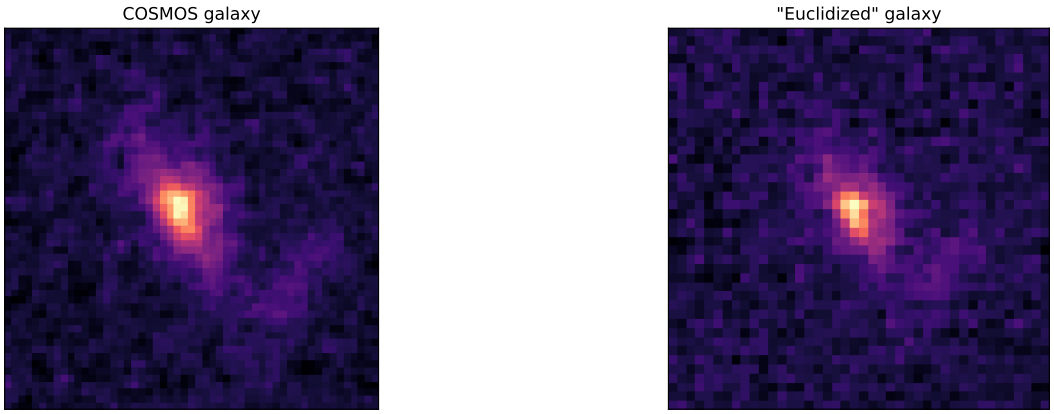


Figure 2.8: "Euclidization" of a COSmic evOlution Survey (COSMOS) galaxy. The galaxy, as observed by HST, is shown on the left; its reprocessed version, as it would appear if observed by VIS, on the right.

As we hinted at at the beginning of this subsection, another recent but far-reaching development is the advent of *meta-calibration* (hereafer MetaCal). First introduced in Huff and Mandelbaum (2017) and Sheldon and Huff (2017), MetaCal foregoes the need for simulations entirely by instead measuring the *shear response*[14] on observed objects. The general procedure is as follows: the PSF is removed from each object through an analytical deconvolution *via* the convolution theorem, that is, an element-wise division in Fourier space. This leads to a non-physical image (as the observation was noisy), that can, however, still be sheared. Several different shear values are applied, then the sheared (and original) images are reconvolved with a PSF that is larger than the one used at the deconvolution step. This removes the spurious modes created by the deconvolution. By measuring the shape of each of these reconvolved galaxies, one can compute the numerical derivative of this shape as a function of applied shear, and thus build a *shear responsivity matrix* that can then be used for calibration. MetaCal has already been successfully applied to real data within the DES survey (Zuntz et al., 2018). Several aspects of MetaCal we have not mentioned in the brief overview of the method given here require careful handling, especially the issue of noise (as the deconvolution/reconvolution process introduces spurious correlations). We refer the reader to the three papers cited in this subsection (as some new developments are proposed in the real-data application of Zuntz et al., 2018).

2.3.9 Obtaining redshift information

One last item is worth mentioning, despite not being part of the shape measurement pipeline per se: redshift estimation for the galaxies in our catalog. As mentioned at the end of Section 2.2.4, whether they are used through tomography or a full 3D analysis, redshifts are needed in order to maximize the amount of cosmological information yielded by a WL survey. In fact, even for a 2D cosmic shear analysis, as we might recall from (2.69), we would still need the $n(z)$ for our galaxy sample.

While the best way to measure redshift is through spectroscopic measurements, the sheer size and number of objects involved in WL surveys requires some other, faster (and less costly) way of obtaining per-object redshift estimates. This is typically done through photo-z, that is, the estimation of an object's redshift from the measurement of its photometry in several different bands - which, in practice, amounts to a very low resolution sampling of its spectral energy distribution (SED). To this end, ground-based surveys often observe in several different bands. In the case of *Euclid*, however, WL measurements will be made from the VIS instrument, which contains only a single, very wide (550-900nm) band. The success of *Euclid*'s WL science objectives is thus

[14] Note the aforementioned approach of Pujol et al. (2019) also makes use of the idea of shear responsivity, though within a simulation-driven calibration framework.

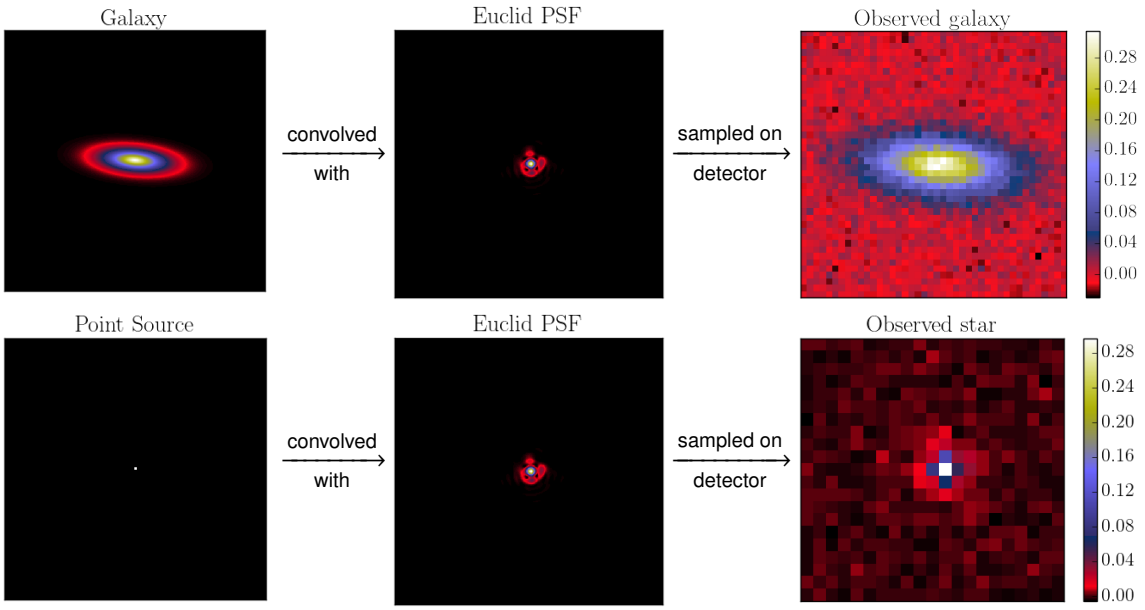


Figure 2.9: Illustration of the imaging process on a galaxy (top left) and a point source (bottom left). Both undergo convolution by the PSF (center), and are then sampled on a finite number of pixels (right).

dependent on photometric measurements obtained from the ground. This is, for instance, one of the main goals of the CFIS survey, further discussed and analyzed in Appendix B.

Once the photometric data has been acquired, there are two main ways to convert them into redshift estimates: template-fitting (see, e.g., Ilbert et al., 2006) and machine learning (e.g. the well-known artificial neural network (ANN) approach of Collister and Lahav, 2004). Salvato et al. (2019) give a very up-to-date (at the time of this writing) review of the field. For an example of a recent application to a WL survey, we once again refer the reader to the first HSC data release, for which the photo-z measurements are described in Tanaka et al. (2017).

Lastly, let us briefly mention *clustering redshifts*, an alternative approach where spatial cross-correlations with a sample of spectroscopically-identified redshifts are used. See Morrison et al. (2017) (and references therein) for an overview of the method, an open-source code[15], and a comparison to KiDS photo-z.

2.4 Point Spread Function

Let us now turn to the main topic of the rest of this thesis: PSF estimation. The images we use for WL purposes are obtained from telescopes, which means they - like any other image from an optical imaging device - are slightly distorted. The origins of this distortion are various: diffraction effects, imperfect optics, for instance due to the polishing of mirrors, and possibly stray light propagating within the instrument. All these make up the *optical PSF*, that we shall focus on in most of the following chapters. For ground-based surveys, the effect of the atmosphere further contributes to the final PSF, and is the main driver for the development of space-based telescopes. In space, there are other contributions to the PSF, such as the jitter of the spacecraft once in orbit (Ma et al., 2008), and the effect of the Attitude and Orbit Control System (AOCS). When these are included, we call the resulting PSF the *system PSF*.

The effect of the final PSF on an image is a convolution; by definition, the PSF is the kernel of that convolution. The *imaging process*, that is, how the actual, continuous intensity profile of an object I on the sky (which we consider as already including the effects of gravitational lensing) is transformed by the imaging process is

[15] https://github.com/morriscb/The-wiZZ/

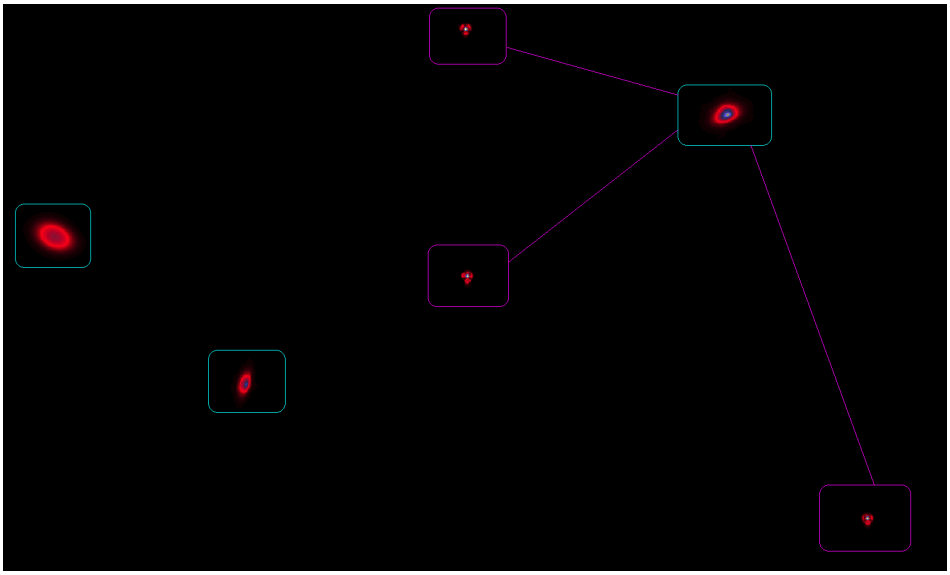


Figure 2.10: Cartoon representation of the PSF field estimation problem. Stars, indicated in magenta, provide a measurement of the PSF, that must then be used to determine the PSF at the position of the galaxies, in cyan.

illustrated in Figure 2.9. Galaxies undergo a convolution by the PSF that alters their shape, and are further sampled on a finite number of pixels, which also contain additive noise. In our case of optical images taken with CCDs, it will always be a safe assumption that the noise is drawn from a white Gaussian distribution.

Applying a convolution to a point source simply results in the convolution kernel. Unresolved stars in the field should be point sources, and therefore give us measurements of the PSF; however, they also undergo the imaging process, and we thus only have access to noisy realizations of our PSF. The *PSF modelling* task is, starting from these, at the bottom right of Figure 2.9, to build an estimator of the *true* PSF, in the central column, that we can then combine with the observed galaxy in the top right, through some shape measurement scheme (see Section 2.3.7), to obtain an estimate of the true galaxy in the top left, or at least of its shape.

The PSF varies as a function of position within the focal plane; this is especially true for ground based surveys, as light arriving with different incidence angles has gone through different atmospheric turbulences, but also holds true for the optical PSF. Stars therefore do not directly give us (degraded) samples of the PSF H, but rather samples, at a fixed set of locations, of a spatially varying function $H(x, y)$. PSF modelling therefore also entails an interpolation, as illustrated in Figure 2.10. In the case of ground-based surveys, where the atmospheric component varies quickly, this estimation needs to be carried out on a per-image basis. Space-based telescopes are a lot more stable, and it is in principle possible to combine several exposures to build the PSF model. However, even a space telescope undergoes changes as time passes, which in turn leads to a varying PSF, as was realized early on in WL studies using the HST (see, e.g., Heymans et al., 2005; Schrabback et al., 2007). In this case, the leading factor was found to be a slight defocus that altered the PSF of ACS-acquired images. The temperature of the spacecraft, which varies with time, and particularly with the relative position of the Sun, also impacts its individual components, and thus the PSF.

Beyond the very intuitive nature of the importance of the PSF modelling step to WL studies that we have already mentioned, its practical implications were made evident very early on. One of the “first four” detections of cosmic shear was that of Van Waerbeke et al. (2000), using observations made at the CFHT, and was quickly followed-up by Visible and InfraRed Multi-Object Spectrographs (VIRMOS)-Dark matter from Ellipticity Sources CARTography (DESCART; Van Waerbeke et al., 2001), one of the very first cosmological WL-oriented surveys, that yielded significant constraints on cosmological models. Hoekstra (2004) later showed that the PSF model used was insufficient to capture the full extent of spatial variations. This prompted a reanalysis of *the same data*, including a more accurate PSF model, in Van Waerbeke et al. (2005), which drastically changed the

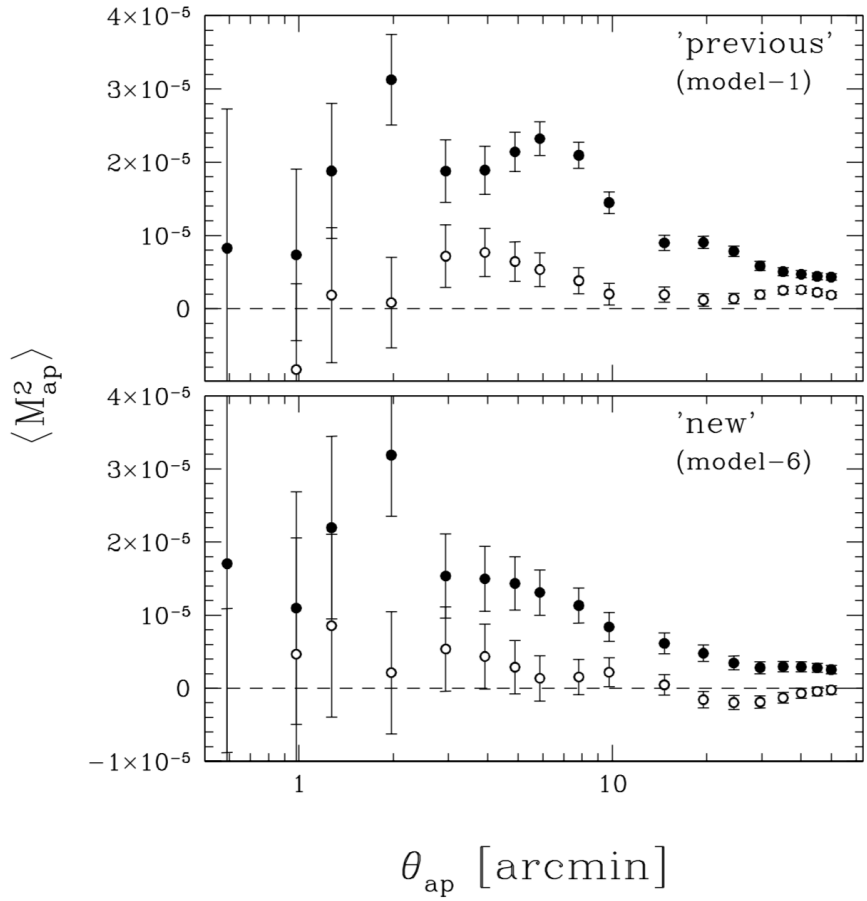


Figure 2.11: Aperture mass measurements from the VIRMOS-DESCART survey. Top: original processing. Bottom: reanalysis, including a corrected PSF model. Credit: Image from Van Waerbeke et al. (2005), reproduced with permission.

values of the mass aperture statistics (one of the derived second order statistics we discussed at the end of Equation 2.2.4), as shown in Figure 2.11. The apparent bump at scales of $\theta_{ap} \approx 5$arcminutes in the original analysis, for instance, completely disappears with the new model, and was thus likely due solely to the PSF model imperfections. To complement this striking (and early) example, in Section 2.4.1, we will present a more general framework to measure the impact of PSF modelling error on WL.

There are two different approaches to the PSF modelling problem. Our knowledge of the instrument can be used to build a model of the optical PSF, which gives us a simulator of the PSF, the parameters of which can then be fitted to the stars. For this reason, these are often called *parametric*, or physical, models. Since it is, in principle, impossible to know the atmospheric component of the PSF in ground-based surveys a priori, this approach has, so far, only been used with space-based telescopes (though see our discussion of `Piff` below). For WL, the only example of such, up to now, has been the HST. Parametric models of its ACS instrument are obtained using the `TinyTim` software (Krist, 1995).

In the *non-parametric*[16], or empirical, approach, knowledge of the instrument (if available) is not used, and the PSF model is built instead using *only* the available data, that is, the unresolved stars in the field.

[16]When I began my work on the Euclid PSF, the name we used for these methods was "data-driven". Since physical models also use data to fit the parameters, this naming scheme was faded out in favour of the parametric/non-parametric one. Of course, one could argue that non-parametric models still do contain *some* parameters, as we shall see in Chapters 3 and 5.

As mentioned, this has been the preferred approach to deal with ground-based surveys, since the atmospheric contribution can then be estimated on a per-image basis. Even for space-based surveys, however, it is desirable to pursue a non-parametric approach: in the aforementioned case of HST, both Jee et al. (2007) and Hoffmann and Anderson (2017) found that they outperformed the `TinyTim` physical model, for WL and for photometric purposes. Moreover, a physical model is only as good as our assumptions about the physics of the instrument. Any effect left unaccounted for will be entirely missed, while in principle non-parametric approaches could recover them. They also should be less susceptible to unexpected surprises, for instance a non-nominal launch of the instrument that has a higher-than-expected impact on its components. Lastly, and most importantly, developing both approaches allows for their eventual combination, likely leading to a better PSF model than either method taken separately.

After introducing the formalism for propagation of PSF errors in the next subsection, in Section 2.4.3, I attempt to give an exhaustive list of all PSF modelling approaches (both parametric and non-parametric) used in, or developed for, cosmological WL applications.

2.4.1 Impact of PSF modelling on shear measurements

Beyond the intuitive notion that an incorrect PSF model would interfere with galaxy shape measurement, and thus directly impact the cosmological results derived from cosmic shear, and the very early practical realization that this was indeed very true (see Figure 2.11), a need for a way to *quantify* the impact of PSF modelling errors on shear measurements naturally arises. It could allow us to control the ultimate impact of our PSF models on science results, or even to set *requirements*. There is no such thing as a perfect model, but if we know how the PSF errors propagate, we can start from a requirement, say, on the tightness of our constraint on the Dark Energy EOS parameters in (2.46), allocate a budget to different sources of systematic errors, including those related to the PSF, and use that propagation scheme to convert this budget into a maximum allowed PSF error. We can then derive objectives for the quality of the PSF model.

Such a propagation framework was first proposed by Paulin-Henriksson et al. (2008), working with the quadrupole moments we introduced in Section 2.3.7. This work was further expanded by Massey et al. (2012). While the latter also included both detector effects and those due to the shape measurement approach in their work, in this section, I will consider only the PSF-related terms of their proposed formalism.

Let us then consider a galaxy with complex ellipticity, defined from its quadrupole moments as in (2.96), $e^{\rm gal}$ (which already includes the contribution of gravitational effects) and size, defined in (2.99), $R^2_{\rm gal}$. After undergoing the first step of the imaging process of Figure 2.9, that is, convolution by a PSF of ellipticity and sizes $e^{\rm PSF}, R^2_{\rm PSF}$, the resulting profile has shape

$$e^{\rm obs} = e^{\rm gal} + \frac{R^2_{\rm PSF}}{R^2_{\rm gal} + R^2_{\rm PSF}}(e^{\rm PSF} - e^{\rm gal}) \tag{2.105}$$

and

$$R^2_{\rm obs} = R^2_{\rm gal} + R^2_{\rm PSF}. \tag{2.106}$$

The quantity of interest in (2.105) is $e^{\rm gal}$, while what we have access to in practice are measurements of all observed and PSF quantities. Using (2.106), we can rewrite

$$e^{\rm gal} = \frac{e^{\rm obs}R^2_{\rm obs} - e^{\rm PSF}R^2_{\rm PSF}}{R^2_{\rm obs} - R^2_{\rm PSF}}. \tag{2.107}$$

From this equation, it would appear the only quantities of interest in a PSF model are its size and ellipticity; however, this is due to our assumption of being in an ideal case of quantities derived from unweighted quadrupole moments, as we shall see shortly.

As a first step, however, let us keep considering these quantities. To study the impact of errors in the reconstructed PSF, let us consider an imperfect PSF model. Since its ellipticity and size are the only quantities related to the PSF appearing in (2.107), let us simply consider $\delta(R^2_{\rm PSF})$ and $\delta e^{\rm PSF}$, the errors in those quantities present in our model. Note that at this stage, whether these errors actually occur from the model itself (that is, the PSF model has a different shape than the true PSF), or from our incapacity to correctly measure the shape of our PSF (e.g. because we cannot measure unweighted quadrupole moments in practice, even on a noiseless PSF model) makes no difference.

The key idea to propagate these errors into our galaxy shape measurement is to perform a Taylor expansion of (2.107) with regards to the shape of the PSF. At first order, this gives

$$\hat{e}^{\rm gal} \approx e^{\rm gal} + \frac{\partial e_{\rm gal}}{\partial\left(R^2_{\rm PSF}\right)}\delta(R^2_{\rm PSF}) + \frac{\partial e_{\rm gal}}{\partial e^{\rm PSF}}\delta e^{\rm PSF}, \tag{2.108}$$

where $\hat{e}^{\rm gal}$ is the object ellipticity computed using the imperfect model rather than the true PSF. The two partial derivatives are readily computable from (2.107), and yield (dropping the gal notation for quantities relative to the galaxy)

$$\hat{e}_i = e_i\left(1 + \frac{\delta(R^2_{\rm PSF})}{R^2_{\rm gal}}\right) - \left(\frac{R^2_{\rm PSF}}{R^2_{\rm gal}}\delta e_i^{\rm PSF} + \frac{\delta(R^2_{\rm PSF})}{R^2_{\rm gal}}e_i^{\rm PSF}\right), \tag{2.109}$$

where we have separated the ellipticity into its individual components $i \in \{1, 2\}$.

This is the famous expression derived by Paulin-Henriksson et al. (2008) (their equation 8, or the 26$^{\rm th}$ in Massey et al., 2012). This expression has been widely used by the WL community, for instance to study the number of stars required to achieve a given PSF model accuracy (Paulin-Henriksson et al., 2009), to quantify the quality of a given PSF model on real data (Rowe, 2010; Jarvis et al., 2016, see Section 2.4.2), and, significantly, to set requirements on WL surveys (Cropper et al., 2013).

There are a few more interesting observations worth making from (2.109). First, remember the usual linear parametrization of shear bias of (2.104). Taking the average of (2.109) across a large number of objects of size $R^2_{\rm gal}$, and under the common assumptions of (2.86–2.87), we get the predicted contributions (of the PSF modelling) to multiplicative and additive biases,

$$m^{\rm PH} := \frac{\delta(R^2_{\rm PSF})}{R^2_{\rm gal}}, \tag{2.110}$$

and

$$c_i^{\rm PH} := -\left(\frac{R^2_{\rm PSF}}{R^2_{\rm gal}}\delta e_i^{\rm PSF} + \frac{\delta(R^2_{\rm PSF})}{R^2_{\rm gal}}e_i^{\rm PSF}\right), \tag{2.111}$$

respectively. Since we split (2.107) into both components instead of using the complex notation, we can point out that the expected multiplicative bias should be the same across both, while the additive bias depends on the PSF's own ellipticity, leading to a different c_i for $i \in \{1, 2\}$. We can also already see, from (2.109), how the way the PSF impacts shape measurement is twofold: both the PSF's intrinsic shape, $e^{\rm PSF}, R^2_{\rm PSF}$, and our *knowledge* of those quantities will impact the measured shapes. From a mission design perspective, this can be directly

linked to the questions of hardware and software: a WL survey should, ideally, aim to have both a very small and isotropic PSF (hardware/instrument design), and an excellent PSF model to recover these values with high accuracy (software/data processing).

Cropper et al. (2013) use (2.109) (among many other factors) to develop a top-down approach: starting from science requirements (i.e. constraints on cosmological parameters), they obtain requirements on both the instrument and our ability to process its data. This is, in a way, the opposite of the bottom-up approach of the Massey et al. (2012) companion paper we have been focusing on. This can be achieved by injecting (2.109) into the definition (2.80–2.81) of the shear 2PCF, then propagating the terms arising because of the errors to our constraints on cosmological parameters, for instance by using the Fisher matrix formalism (though this choice should be made with care, see Wolz et al., 2012; Euclid Collaboration et al., 2019).

As we have already mentioned earlier, unweighted quadrupole moments can never be computed in practice because of noise, and thus require apodisation, that is, the use of a window function. This raises the question of how much trust can be put into the predictions we get from (2.109) if the left hand side is, in practice, measured using a window function (or, equivalently, a specific analytical model). In fact, even the terms on the right-hand side likely cannot be estimated with unweighted moments. This is especially true in the case of a diffraction-limited PSF, which has diverging second order moments, which means e^{PSF} is not even defined!

Melchior et al. (2011) studied the impact of the addition of a window function in the measurement of quadrupole moments and (especially) ellipticity. They show, analytically, that the addition of a window function causes a mixing with higher-order moments of the same object. Following Massey et al. (2012), let us consider these effects as prefactors, P_e and P_R, corresponding to the changes in e^{PSF} and $R^2_{\mathrm{PSF}}/R^2_{\mathrm{gal}}$, respectively. Note that since the only size terms in (2.109) are ratios of PSF and galaxy sizes, the second prefactor is related to that ratio, while the first is only related to PSF quantities. This means that P_R is not just a function of the PSF, but also of the object being measured. Both are, of course, dependent on the shape measurement method chosen (and on its hyperparameters, like the window size). We can then add these terms in (2.109), yielding

$$\hat{e}_i = e_i\left(1 + \frac{\delta(R^2_{\mathrm{PSF}})}{P_R R^2_{\mathrm{gal}}}\right) - \frac{1}{P_R P_e}\left(\frac{R^2_{\mathrm{PSF}}}{R^2_{\mathrm{gal}}}\delta e_i^{\mathrm{PSF}} + \frac{\delta(R^2_{\mathrm{PSF}})}{R^2_{\mathrm{gal}}}e_i^{\mathrm{PSF}}\right). \tag{2.112}$$

For a Gaussian PSF, P_e and P_R are exactly equal to 1 and (2.112) reverts to (2.109). Neglecting those terms is thus likely acceptable for ground-based surveys, where the atmospheric contribution smooths out the PSF. It is far less evident in the case of space-based surveys, however. Beyond the aforementioned undefined ellipticities, the PSF of spaceborne instruments often suffers from far stronger *ellipticity gradients*, that is, the ellipticity of their profile varies with the radius around center used to compute it (see Figure 2.12 for an illustration). In other words, since the weighted measurements of e^{PSF} are strongly dependent on the chosen window function, so is P_e. In Chapter 3, we will address this issue, and indeed show that the simpler formulation of (2.109) no longer holds in the case of *Euclid*.

While the addition of P_R and P_e will prove sufficient to make this point, it should be noted that Massey et al. (2012) in fact go further in that they also consider the error made in determining P_R itself. This is achieved by considering a weighted version of (2.107), then adding the first-order term in $\delta(P_R)$ in (2.108). They note that the same treatment could be applied to errors in P_e, though these would act in the exact same way as δe^{PSF}. In light of the results of next chapter, however, I think even this additional term might be worth explicitly taking into account, as even for one given PSF model, P_e would still vary from a shape measurement method to another.

A couple of alternatives to this formalism have already been suggested to quantify the impact of PSF errors on shear measurement. Lu et al. (2017) propose to use real data, in a somewhat similar heuristic to that used in calibration schemes (see Section 2.3.8). As is commonly done to assess the quality of a PSF model a posteriori, they separate their star sample into two, and only use one of the sets to build their PSF model.

[17] As specified in the internal "EUCL-EST-TN-1-002" document.

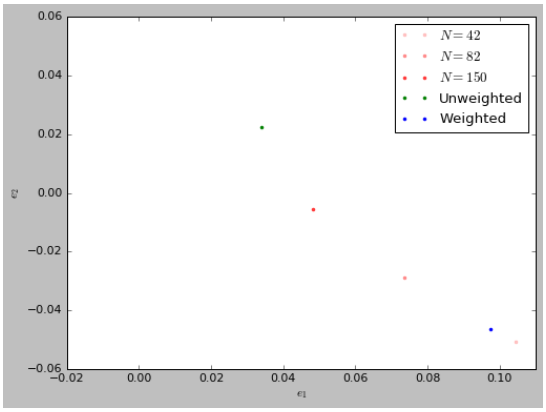


Figure 2.12: Illustration of the "ellipticity gradient" in a space telescope's PSF. Each point shows the two ellipticity components measured on the same simulated Euclid PSF, from a noiseless, very large image with extremely small pixel sampling rate (much more so than that of the VIS detector).
The green point shows the measured ellipticity with no weight function (save for the truncation at the edges of the large stamp), the blue one with a Gaussian weight with the size recommended for internal use within the Euclid Consortium (EC)[17], and the red ones with different square truncations with N pixels around the center.

They then perform galaxy simulations, complete with an applied shear, as would be done in a simulation-driven calibration approach, but they convolve the profile with the stars they left aside. They then perform their shape measurement using both the "true PSF" (i.e. the observed star) and their reconstruction of it. The difference in shear biases found between the two sets of measurement gives their estimation of the contribution of the PSF[18]. While this approach is definitely interesting, especially in its use of real data to try and measure the bias brought by our PSF modelling, the use of noisy star images as ground truth PSFs would be impossible if these were undersampled (as will be the case for *Euclid*). Even in the absence of undersampling, the use of noisy convolution kernels can be problematic. A PSF model that somehow perfectly reproduced all stars in the field, noise patterns and all, would lead to a measured bias that is strictly zero in this case. Except we know stars are imperfect samples of the true, underlying PSF, so we do not really want our models to perfectly reproduce their properties. For instance, the measurement of ellipticities on all stars in any given exposure is extremely noisy.

The other approach has very recently been proposed by Paykari et al. (2019). They construct an "end-to-end" framework where all identified sources of errors and/or uncertainty on our measurement of shear are applied on an object-by-object basis. Their approach can thus be used to account for more effects than just those related to the PSF (though it is one of the two they chose as example for this first application), and comes with the added benefits of treating all considered effects simultaneously, thus also capturing potential cross-effects, and taking into account effects that arise as a function of sky position (e.g. because of where the CCD mosaic falls with each pointing, which in turn creates patterns in where PSF and detector effects occur across the whole survey).

2.4.2 Real data diagnostics

If we assume (2.109) to hold true, we can use it to build PSF model diagnostics applicable to real data. We can see that the quantities impacting galaxy shape measurement are the residuals in ellipticity and size, $\delta(e_i^{\rm PSF}), \delta(R^2_{\rm PSF})/R^2_{\rm PSF}$. All the experiments we will carry out in the main chapters of this thesis use simulations where we had access to the ground truth PSF. We can thus compute the true values for these residuals. On real data, these quantities must be estimated using stars.

As discussed in Section 2.3.7, measurements of shape parameters from quadrupole moments are ex-

[18]Note we will use a similar difference in measured biases in our own simulations (to remove all effects due to the shape measurement itself and not the PSF model) in Chapter 3.

tremely sensitive to noise. Building PSF diagnostics based on (2.109) will thus require averaging over large numbers of objects. Rowe (2010) proposed a set of such statistics, expanded by Jarvis et al. (2016). These are defined as follows:

$$\rho_1(\theta) = \left\langle \delta e^*_{\rm PSF}(x)\delta e_{\rm PSF}(x+\theta)\right\rangle, \tag{2.113}$$

$$\rho_2(\theta) = \left\langle e^*_{\rm PSF}(x)\delta e_{\rm PSF}(x+\theta)\right\rangle, \tag{2.114}$$

$$\rho_3(\theta) = \left\langle \left(e^*_{\rm PSF}\frac{\delta R^2_{\rm PSF}}{R^2_{\rm PSF}}\right)(x)\left(e_{\rm PSF}\frac{\delta R^2_{\rm PSF}}{R^2_{\rm PSF}}\right)(x+\theta)\right\rangle, \tag{2.115}$$

$$\rho_4(\theta) = \left\langle \delta e^*_{\rm PSF}(x)\left(e_{\rm PSF}\frac{\delta R^2_{\rm PSF}}{R^2_{\rm PSF}}\right)(x+\theta)\right\rangle, \tag{2.116}$$

$$\rho_5(\theta) = \left\langle e^*_{\rm PSF}(x)\left(e_{\rm PSF}\frac{\delta R^2_{\rm PSF}}{R^2_{\rm PSF}}\right)(x+\theta)\right\rangle, \tag{2.117}$$

where * denotes complex conjugation. The ρ statistics are therefore made up of all the combination of 2-point auto- and cross-correlations between the shape residual terms (in ellipticity and size) appearing in (2.109), as well as those with the actual PSF shape. These are the terms that appear when injecting equation (2.109) into the computation of the shear 2PCF, ξ_+, defined in (2.80). Their impact on the latter can be expressed as

$$\begin{aligned}\delta\xi_+(\theta) =& 2\left\langle\frac{R^2_{\rm PSF}\delta R^2_{\rm PSF}}{R^2_{\rm gal}R^2_{\rm PSF}}\right\rangle\xi_+(\theta) + \left\langle\frac{R^2_{\rm PSF}}{R^2_{\rm gal}}\right\rangle^2\rho_1(\theta) - \alpha\left\langle\frac{R^2_{\rm PSF}}{R^2_{\rm gal}}\right\rangle\rho_2(\theta)\\ &+\left\langle\frac{R^2_{\rm PSF}}{R^2_{\rm gal}}\right\rangle^2\rho_3(\theta) + \left\langle\frac{R^2_{\rm PSF}}{R^2_{\rm gal}}\right\rangle^2\rho_4(\theta) - \alpha\left\langle\frac{R^2_{\rm PSF}}{R^2_{\rm gal}}\right\rangle\rho_5(\theta),\end{aligned} \tag{2.118}$$

where α is the *PSF leakage* term defined in (3.46) (see Jarvis et al., 2016, for a recipe to estimate it from real data).

Equation (2.118) can then be used to transfer requirements on the accuracy of the 2PCF into requirements on the ρ statistics. Another appeal of the ρ statistics is that they can be used, as originally suggested by Rowe (2010), as a mean to compare different PSF models. They can, in fact, even be used to compare the quality of PSF models across different surveys. See Appendix B for an application to the CFIS survey.

2.4.3 Past and current PSF models

Let us now review all PSF models for WL, starting with all those already applied on real data, as shown in Table 2.2 (and footnotes therein).

"Shape interpolation"

As can be seen in Table 2.2, the vast majority of early WL studies used what I have named "shape interpolation" for PSF modelling. These early works used the KSB method (or some variant thereof) to measure the shapes of galaxies, and KSB only requires specific quantities related to the PSF to carry out the correction, rather than a full pixel model. These are the anistropic *smearing*, and the isotropic damping (as even a round PSF causes a galaxy to become more circular). "Shape interpolation" as a PSF model thus amounted to measuring the shape of stars, then interpolating those quantities across the field. This interpolation was, in most cases, carried out polynomially. A significant exception to that is the early CFHT/CFH12K-based surveys we mentioned at the beginning of this section, for which Hoekstra (2004) built a tailored interpolation function. Chang et al. (2012) also proposed `PSFENT`, an interpolation scheme for star shapes based on maximum entropy. This approach

was meant to handle the high (spatial) frequency variations induced by atmospheric distortions in ground-based surveys with short exposure times (Heymans et al., 2012a) better than polynomial models could.

On ground-based surveys, this shape interpolation was carried out exposure-by-exposure. HST-based studies made use of the telescope's stability, and used exposures with dense stellar fields (e.g. observations of globular clusters) to build the PSF model; see for instance Rhodes et al. (2004)'s model for the Space Telescope Imaging Spectrograph (STIS) instrument (where corrective terms are added to a polynomial model of shapes), and Schrabback et al. (2007), where different PSF models are built on each available dense star field, then the few stars from each WL-useable exposure are used to select the closest model among those, in order to recover the PSF model corresponding to the closest defocus value.

Another interesting exception to this early trend of interpolating only the shape of the PSF is that of Kaiser et al. (2000) (another of the "First four"), where a spatially-varying model of the PSF in full pixel domain was used (with further treatment, closer to the standard shape interpolation approaches of the time, for some identified failure modes). The PSF at field position (x, y) is then constructed as

$$\hat{H}(x, y) := \sum_j A_j(x, y) S_j, \tag{2.119}$$

where each S_j is a full image (rather than just a set of measurements of the shape of the PSF). The A_j coefficients were chosen to be linearly varying. This is extremely similar to the non-parametric approaches that would appear several years later, and which we shall now describe.

Principal Components Analysis (PCA)

In most cases, building a non-parametric PSF model boils down to selecting a way to construct both the A and S terms in (2.119). The most natural way to determine the latter from data is the use of PCA (Dunteman, 1989), which can be either carried out on some quantities related to the PSF (e.g. some measurements of shape, like in the above methods), or on actual star images themselves. In this case, the elements of S are once again full images, sometimes referred to as *eigenPSF*s.

The use of PCA was first proposed, specifically for use in WL, by Jarvis and Jain (2004), though it was already the method of choice for PSF modelling in the general SDSS pipeline (Lupton et al., 2001, where they refer to PCA as the Karhunen-Loève transform), and was used for the subsequent WL measurements by Lin et al. (2012) and Huff et al. (2014) (the former further perform a correction of shape residuals, akin to the shape interpolation method described above). PCA also led to the final model used for WL analysis with the HST (Schrabback et al., 2010). Their latest use on real data was in Jee et al. (2013)'s processing of the Deep Lens Survey (DLS) data.

The choice of a set of A functions for each principal component, or *eigenPSF*, S_j, is then a matter of traditional spatial interpolation. Gentile et al. (2013) give a review of several such standard approaches, with application to GREAT3 data.

`PSFEx`

Bertin (2011)'s `PSFEx` can be considered the gold standard in (freely available) non-parametric PSF estimation. Designed as a companion software to `SExtractor`, the two were initially intended to be used in conjunction to perform both the PSF modelling and the shape measurement (by model-fitting) steps. Though `PSFEx` has never been used in this manner for WL on real data, it is for instance one of the methods applied to simulations by Martinet et al. (2019) in their study of undetected galaxies and their impact on shear measurements. It is applied, in a slightly modified form (to only use the PSF modelling part), in several current WL surveys, namely DES (Zuntz et al., 2018) and HSC (Mandelbaum et al., 2017). See also Appendix B for our own application to CFIS.

The PSFEx model is polynomial, up to a user-set degree d, in any chosen set of SExtractor parameters. By default, these are, naturally, field positions, leading to

$$\hat{H}^{\mathrm{PSFEx}}(x,y) = \sum_{\substack{p,q\geq 0 \\ p+q\leq d}} x^p y^q S_{pq}. \tag{2.120}$$

The set of basis functions S can be chosen either by fitting the parameters of a chosen model or, by default, by learning a full pixel model. A first guess S_0 is then estimated from a median image of all available observations (reshifted to the same grid), and the final $S := S_0 + \Delta S$ is then learned by minimizing

$$\min_{\Delta S} \chi^2(\Delta S) + \|T\Delta S\|_{\mathrm{F}}^2. \tag{2.121}$$

We will return to the χ^2 term and define it explicitely in Equation 3.3.1 of next chapter. For now, let us simply state that it is, as indicated by its name, a data fidelity term that ensures the model recovers the profile of the observed stars. In Section 3.1.2, we shall introduce and discuss the need for *regularization* in inverse problems; the right-hand term in (2.121) is one such, namely a Tikhonov regularization. In the case of PSFEx, T is chosen to be scalar, and this term favors solutions S with smaller l_2 norms.

This default configuration (polynomial dependence on positions and data-derived pixel basis functions) is the one in which PSFEx was always ran in in WL surveys. Note that in this case, in order to rewrite (2.120) in the general form of (2.119), we only have to define[19]

$$\forall i, A_i := (x_i^p y_i^q)_{p+q\leq d}. \tag{2.122}$$

By doing so, we recast the problem solved when building the PSFEx model as one of *matrix factorization*, as is our own proposed approach we will present in Chapter 3.

LensFit

LensFit (Miller et al., 2007; Kitching et al., 2008) is a Bayesian shape measurement approach of the model-fitting class. It was the method used to process the CFHTLenS data, for which a non-parametric PSF modelling approach was also developed (Miller et al., 2013). The name LensFit is now often used to identify the whole suite of methods, which also includes others steps described in Section 2.3 beyond PSF modelling and shape measurement. It was one of the entries in the GREAT3 challenge, under the name "MaltaOx" (Mandelbaum et al., 2015). It is also currently being used in KiDS, as described in Kuijken et al. (2015), which incidentally provides greater detail of the PSF part than the initial Miller et al. (2013) paper. Here, we shall focus on the PSF modelling part of LensFit, which I will simply refer to as LensFit.

Unlike PSFEx, which builds a set of basis functions that contain as many pixels as the desired final model, the LensFit model is fitted pixel-by-pixel. The value of each pixel is a polynomial combination of positions, up to a user-set order d (called n in Kuijken et al., 2015). A significant improvement over PSFEx, which needs to be fitted on each CCD separately, is that LensFit can be fitted jointly on the whole exposure. The discontinuities in the PSF caused by the jump from one CCD to the next are handled by fitting the polynomial coefficients up to a certain (also user-set) degree, $n_{\mathrm{c}} \leq d$, separately on each detector.

One way to recast the LensFit model into our notation from (2.119) is thus as follows. Consider each S_j to be an image with a single non-zero pixel. There would then be as many as the desired number of pixels in the final model (in KiDS, this number was chosen to be 35×35, at the same pixel scale as that of the instrument). Let c denote the CCD number, and N_{D} the number of CCDs in the camera mosaic. We then have

[19] Where, here and throughout, I shall try to use i to index different objects (stars or galaxies), and j to index elements of S, which I shall carry on referring to as *eigenPSF*, even when they are not actually chosen through PCA.

$$\hat{H}^{\mathrm{LensFit}}(x,y) := \sum_j \left[\sum_{\substack{p,q\geq 0 \\ p+q\leq n_\mathrm{c}}} \left(x^p y^q \sum_{k=1}^{N_\mathrm{D}} \mathbb{1}_{\{c(x,y)=k\}} a_{kpq}^{(j)} \right) + \sum_{\substack{p+q>n_\mathrm{c} \\ p+q\leq d}} x^p y^q a_{0pq}^{(j)} \right] S_j. \tag{2.123}$$

Note that, with these notations, `LensFit` uses predetermined S_j's and the free parameters to be fitted to star observations are, for each pixel j,

$$A_j := \left(a_{kpq}^{(j)} \right)_{k,p,q}. \tag{2.124}$$

This is the opposite of `PSFEx`, where the A were predetermined by (2.122), and S learnt from the data. When fitting the A_j's, each star is given a weight

$$w_i := \frac{\nu_i^2}{\nu_i^2 + 50^2}, \tag{2.125}$$

where ν_i is the integrated star SNR on its exposure.

The number of free parameters per pixel is then

$$N_\mathrm{coeff} = \frac{1}{2}[(d+1)(d+2) + (N_\mathrm{D}-1)(n_\mathrm{c}+1)(n_\mathrm{c}+2)]. \tag{2.126}$$

In both CFHTLenS and KiDS, the parameters were chosen to be $d = 3, n_\mathrm{c} = 1$, after comparison of the quality of the models resulting from several different values on a subset of the data. This led, for instance, to a number of free parameters $N_\mathrm{coeff} = 103$ for KiDS (carried out with the OmegaCAM instrument that has $N_\mathrm{D} = 32$ CCDs). While this may seem unreasonably high, it should be kept in mind that the fitting will be carried out using stars from the whole exposure at once. Assuming a uniform distribution of stars on science exposures, this means a factor of about N_D more datapoints than would be available on each CCD.

Let us mention one last important specificity of `LensFit` that arises because it treats each pixel separately. So far in this section, we have used (x, y) to denote the position, within an image, of an object. The true centroid of this object is highly unlikely to fall precisely at the center of a pixel. We must thus consider each real object to be sampled with an intra-pixel *shift*, corresponding to the difference between an arbitrary position within the pixel grid and that of its true centroid. This is illustrated in Figure 2.13. When comparing an observed star with a PSF model, one of the two needs to be shifted so that both are sampled on the same grid. In most cases (e.g. `PSFEx`), the *model* is shifted to match each star. The opposite is true of `LensFit`, however, where the observations are all shifted to a common grid. While this allows for each pixel to be treated independently, this shifting, carried out using a sinc function, implies performing an intra-pixel interpolation on noisy observations.

`TinyTim`

We have already cited `TinyTim` (Krist, 1995) as the only example of a physical PSF model applied in a WL context. The software can be used to build a model of the ACS on board the HST, the only spaceborne telescope used for cosmological WL to date. Its specific use (and modifications) in this case are detailed by Rhodes et al. (2007). Like later, PCA-based models used in the COSMOS surveys, `TinyTim` is ran with various defocus levels, which were found to be the main driver in the temporal variations of the PSF. All the resulting models are compared with the few available stars in the extragalactic exposures to select the best match, which corresponds to finding the closest defocus level. This PSF model was the one used to build the first COSMOS shape catalog (Leauthaud et al., 2007), though as we already mentioned, later analysis found a non-parametric approach to yield more accurate models.

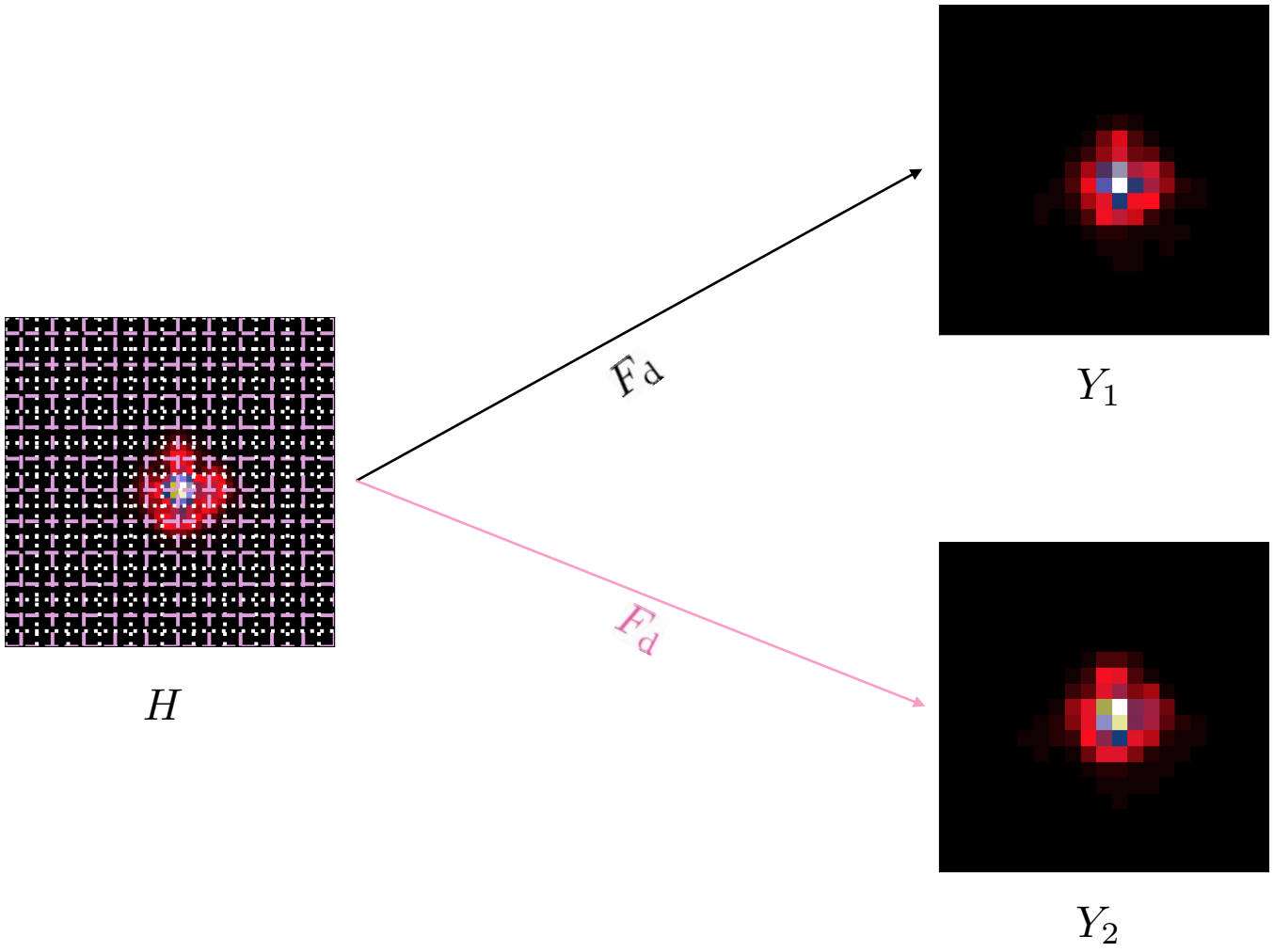


Figure 2.13: Illustration of intra-pixel shifts. The same PSF, H, is sampled on two different pixel grids, indicated by the dotted white and dashed pink lines. Application of the degradation operator, F_{d}, yields two different observations at detector pixel scale, Y_1 and Y_2. Despite their coming from the same object H, and being shown in postage stamps of the same size, they appear very different because their respective centroids fall at different positions with regard to, say, the central pixel of the stamp. Any PSF modelling approach based on star observations must account for this effect through the use of intra-pixel *shifts*.

Shapelets

Shapelets (Refregier, 2003) are a particular set of basis functions, that were originally considered for use in galaxy shape measurement. In this case, the PSF itself can be decomposed in the shapelet basis (Refregier and Bacon, 2003), and the correction can then be applied in shapelet space.

Every PSF modelling method we have presented up to now has been used in cosmic shear surveys. To the best of my knowledge, this is not the case of a shapelet-based PSF model. It has, however, been used by Romano et al. (2010) to perform a WL study of the Abell 611 object (which they also compared to a KSB measurement, with standard "shape interpolation" as PSF model). They build a shapelet decomposition of their stars, then interpolate the shapelet coefficients with a fourth-order polynomial, which amounts to using the shapelet basis functions as S in (2.119). A shapelet-based entry was also submitted to the GREAT10 Star Challenge under the name "Gaussianlets" (Kitching et al., 2013).

Fourier

Zhang et al. (2015) is the latest in a series of papers proposing a shape measurement approach, `Fourier_quad`, that was one of the main entries in the GREAT3 challenge (Mandelbaum et al., 2015). Like other moments-based methods introduced in Equation 2.3.7, it relies on computing the quadrupole moments of the galaxy images, though this computation is done in Fourier rather than direct space. To carry out the PSF correction, the Fourier transform of the PSF model at the position of the galaxy is used, which in turn naturally brings forth the idea of performing the PSF modelling step in Fourier space. Similar to `LensFit` in the direct domain, a pixel-by-pixel interpolation is carried out from the Fast Fourier Transform (FFT; Cooley and Tukey, 1965) of the star postage stamps, potentially after applying some noise mitigation scheme (Zhang et al., 2019).

One particular consequence of carrying out the PSF modelling in Fourier space is that application of the FFT leads to centered postage stamps, and thus circumvents the issue of intra-pixel shifts and interpolation of noisy objects that, as mentioned, `LensFit` suffers from. Moreover, Lu et al. (2018) recently proposed an approach to remove correlations occurring because of remaining errors in the PSF model. While this is currently the only approach that can deal with PSF modelling errors a posteriori, it is method-dependent, i.e. it can only be applied in the specific case of a Fourier PSF model used with `Fourier_quad` as the shape measurement.

Machine Learning

Let us once again turn to our equation (2.119), and equate the non-parametric PSF modelling to that of finding an appropriate representation space. As mentioned, PCA is one of the most classical ways to obtain such a representation. More recently, *unsupervised* machine learning methods have often been used to carry out this task. Their application to the PSF modelling problem was already considered by the `MegaLUT` team of the GREAT3 challenge and briefly mentioned in Mandelbaum et al. (2015).

An approach based on ANNs is presented in the PhD thesis of Kuntzer (2018). This approach requires a very large training set of "true PSFs", which typically would have to be obtained through simulations. This then raises the concern of whether these PSF simulations can be made to be sufficiently realistic. Another approach, proposed by Herbel et al. (2018), uses a Convolutional Neural Network, also trained on simulations, to produce PSF models from a set number of parameters that can be measured from stars on actual data. This in turn allows them to produce PSF models that match the properties seen on data, with their use in simulations (rather than shape measurement) in mind.

VIS physical model

We now turn to two PSF models that have, at the time of this writing, not yet been described in published work, but are under active development and planned to be used in some of the main ongoing and upcoming WL surveys. The first is a physical (or parametric) model of the VIS PSF, which will be described in detail in

Duncan et al. (in prep)[20]. It contains both a physically-motivated "simulator" of the PSF, the VIS equivalent to ACS's `TinyTim`, and a fitting approach to select the parameters of this model from observed, undersampled stars in *Euclid* images (including some planned calibration fields). An early version of this model is used and briefly described in Paykari et al. (2019).

The bulk of the original work we will present in later chapters of this thesis also aims at building a VIS PSF model, albeit within a non-parametric framework. Our approach and this one are currently being developed entirely independently, which is desirable both for redundancy and cross-validation. As we already mentioned, however, ultimately *combining* the two approaches is likely to lead to a better PSF model than either could achieve separately.

`Piff`

The idea of such an hybrid approach, combining a physical and an empirical model of the PSF so that the latter can capture and correct for effects missed by the former, was already suggested over a decade ago by Jarvis et al. (2008). It is one of the planned features of `Piff` (PSFs In the Full FOV), a new PSF modelling Python package. Its development, led by Mike Jarvis, is already at quite an advanced stage[21].

Miscellaneous

Let us now briefly mention a few other noteworthy works in the field of PSF modelling. In the list we have just built, several models create a full image of the PSF, often by working directly from the pixels of the star images to construct their own basis S. This is, for instance, the case of `PSFEx`, when ran in its default configuration. While this setting is the most widely used, `PSFEx` also allows for the use of predetermined basis functions (namely, either the polar shapelets introduced by Massey and Refregier, 2005, or a user-provided set). In a similar vein, and analogously to what is done when performing model-fitting for galaxy shape measurement, one can in principle select an analytical profile, fit its free parameters to the observed stars, then interpolate those over the field positions to obtain a PSF model. This was attempted in the variable PSF branch of the GREAT3 challenge by the CEA-EPFL team, who used an elliptical Moffat profile (Moffat, 1969).

When using such an approach, like with many others (e.g. PCA or shape interpolation), the choice of the interpolation function to use, that is, of the A coefficients in (2.119) remains free, and thus falls into the classical problem of *spatial interpolation*, for which numerous methods exist. Reviews of several of these are given by Bergé et al. (2012); Gentile et al. (2013) and, in the specific case of the Fourier approach described above, by Lu et al. (2017). A particularly elegant alternative to these classical methods was recently proposed by Ngolè and Starck (2017). It relies on notions of manifold learning and on numerical OT, specifically Wasserstein barycenters, which we shall return to and use profusely in Chapters 4 and 5.

The only examples we have given of physical PSF models so far are those for the HST (`TinyTim`) and *Euclid* (Duncan et al., in prep). Davis et al. (2016) proposed such an approach for Dark Energy Camera (DECam), though as DES is a ground-based survey, a full PSF modelling would also require handling atmospheric effects, once again highlighting the importance of hybrid approaches. In this area, Xin et al. (2018) propose a two-part model for the SDSS PSF, which is, interestingly, somewhat opposite to the hybrid case considered by Jarvis et al. (2008). The latter suggested using the knowledge of the instrument, that is, of the *optical* PSF, as a base model, to be complemented by an empirical approach to account for the remaining atmospheric (or otherwise) effects. In contrast, this approach instead uses a physically-motivated model of atmospheric turbulences (namely the von Kàrmàn turbulence model, De Karman and Howarth, 1938), further convolved with a model aiming at capturing the (unknown) instrumental PSF.

"Pi of the Sky" is a *very* wide field of view (FOV) - 1.5 steradians! - experiment with the early detection of (optical counterparts to) Gamma Ray Bursts as one of its main science drivers (Burd et al., 2005). This design leads to a very highly distorted PSF toward the edges of the camera. The task of modelling such a PSF is of

[20] In the meantime, *Euclid* members may refer to the "`EUCL-OXF-TN-draft`" document.
[21] `https://github.com/rmjarvis/Piff`

course quite different from that of WL-oriented surveys, and entails dealing with dramatically stronger variations across the field. Piotrowski et al. (2013) tackle this challenge using a physical model based on laboratory measurements of the instrument, improved by a fitting procedure using observed stars.

Lastly, let us mention the work of Suksmono (2013), who uses a Compressed Sensing approach to interpolate PSF ellipticity field at all positions. This approach should therefore fall under the "shape interpolation" category, yet is remarkable for two reasons. First, it makes full use of the definition of ellipticity as a complex number, by considering the PSF ellipticity field as complex-valued random field. Second, as a compressed sensing approach, it makes use of the same notion of *sparsity* that we shall introduce and use in our own approach in the next chapter.

2.5 *Euclid*

We now turn our attention to the *Euclid* mission, main focus of the work presented in this thesis. As with the beginning of this chapter, I shall start from the very general, with a broad overview of the mission, then quickly narrow our scope down to those parts that my PhD work has related to. This will enable us, in particular, to reach an understanding of some of *Euclid*'s specificities, and how they make the modelling of its PSF quite the unique task, from which originates the need for new methods beyond those I presented in Section 2.4.3. For a far more detailed (albeit somewhat aged) overview, we refer the reader to the Euclid Redbook (Laureijs et al., 2011).

Euclid is a medium-class ESA mission, part of its "Cosmic Vision" program for the $2015-2025$ decade. The main observations will be carried from an eponymous space telescope, the launch of which is planned on a Soyuz rocket in June of 2022. As already mentioned, complementary observations from several ground-based surveys are critical to the overall mission, especially to obtain photo-z measurements. In this section, however, I will focus on the *Euclid* spacecraft.

The spaceborne observatory will carry out its work, for an expected nominal mission span of 6 years, from the L2 Sun-Earth Lagrange point. It will carry a 1.2m primary mirror, along with two instruments: the visible imager VIS, and a Near Infrared Spectrometer and Photometer (NISP). The scientific exploitation of *Euclid* data, and the design and construction of these two instruments is the responsibility of the EC, made up of about 1500 members from 14 participating European countries, as well as Canada and the United States of America.

The main science objectives of *Euclid* are in the field of cosmology, and especially aimed at a better understanding of the "Dark Universe", that is, the distribution of dark matter and the nature of Dark Energy (see Section 2.1.5). This will be achieved through the use of two main cosmological probe: galaxy clustering (including the study of BAO), and WL. Several other scientific results will also be derived, both in cosmology (e.g. *via* clusters of galaxies) and other areas of astronomy and astrophysics (e.g. exoplanets and solar system objects).

In addition to the three Euclid Deep Fields (EDFs, dubbed EDF North, EDF South and EDF Fornax) that will be observed repeatedly in order to go deeper and probe the high-redshift Universe (which will *also* be useful for WL, especially calibration; see e.g. Martinet et al., 2019), the *Euclid Wide Survey* aims to maximize the yield of its two main cosmological probes. As both WL and galaxy clustering rely on the measurement of a large number of objects, the aim is to observe as much of the sky as possible. Figure 2.14 shows the footprint of the Euclid Wide Survey. Since the main objects of interest are distant galaxies, several zones are excluded both to minimize the amount of contaminants and to maximize the quality of such observations. The former criterion naturally leads to the exclusion of the plane of our own galaxy (where most observed objects would be intra-galactic) and the Large Magellanic Cloud (LMC), the latter to that of the ecliptic plane (to avoid contamination by zodiacal light cause by our own Sun) and regions of strong reddening (caused by intra-galactic dust and measured to high accuracy by *Planck*). Note that *Euclid will* carry out some observations in those excluded areas, including some calibration fields which might prove particularly useful for PSF modelling. After exclusion of these zones, the Wide Survey amounts to a total of $15,000\mathrm{deg}^2$, which will allow for the detection of billions of galaxies, and a highly precise redshift measurement for tens of millions of them from the spectroscopic part of NISP. The expected number density of galaxies suitable for shape measurement is of $30\mathrm{arcmin}^{-2}$, a vast

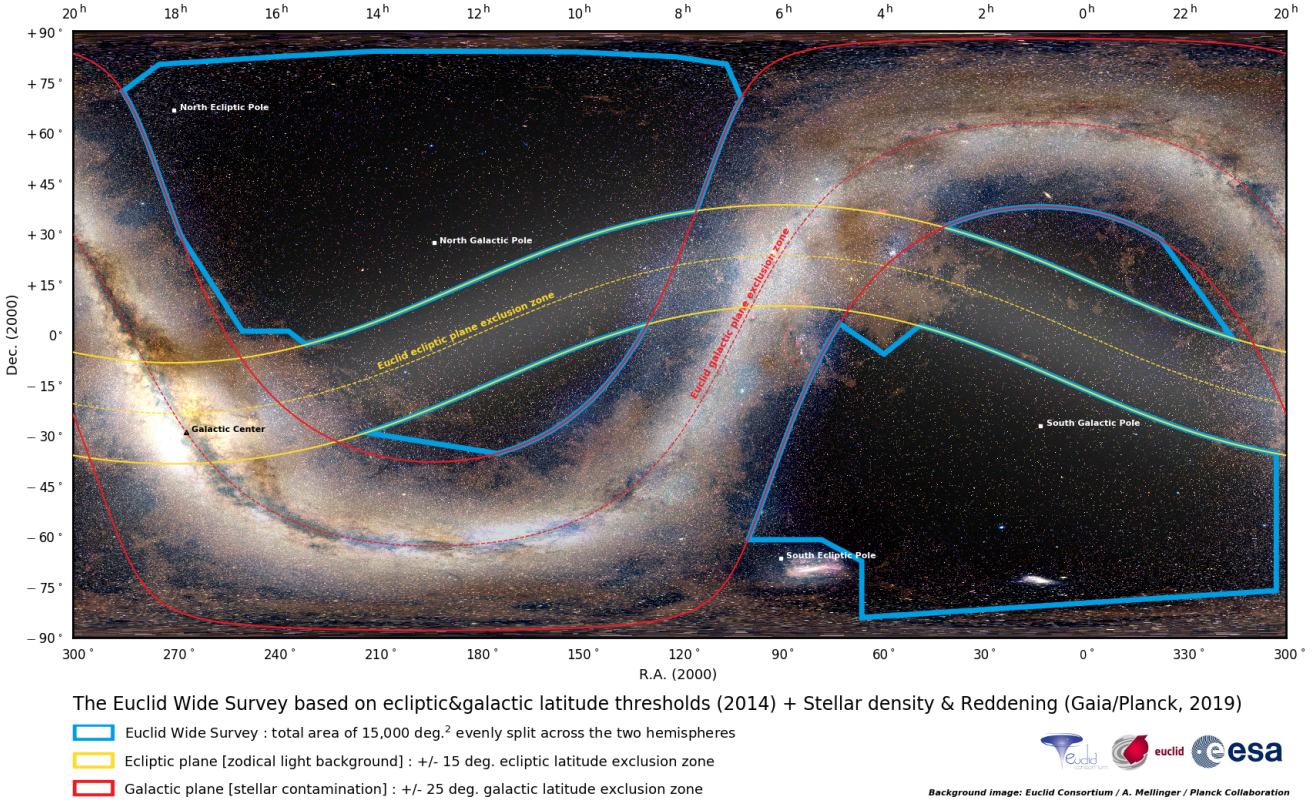


Figure 2.14: Full-sky map showing the Euclid Wide Survey (contoured in blue), and the main excluded zones: galactic and ecliptic planes. The LMC, at the bottom-left of the southern extragalactic region, is also excluded. Credit: Jean-Charles Cuillandre.

improvement even over the already impressive, current (effective) number density of $21.8\mathrm{arcmin}^{-2}$ for the HSC survey (Mandelbaum et al., 2017).

With WL measurements as one of its core objectives, *Euclid* has been designed with accurate shape measurement as one of the main drivers. While the photometric part of NISP will contribute, along with complementary observations from the ground, to the determination of photo-z, the shape measurement itself will be made using images from the VIS instrument. The VIS focal plane, represented in Figure 2.15, is made up of a mosaic of 6×6 CCDs, each of them comprising 4096×4132 pixels. The pixel scale is of about $0.1\mathrm{arcsec}$, leading to a very wide FOV of about $0.57\mathrm{deg}^2$. For comparison, the FOV of the HST's ACS instrument, from which the best space-based WL catalogs to date were produced, is about 180 times smaller.

By design, the VIS PSF will allow for an exquisitely small (with an FWHM of less than $0.2\mathrm{arcsec}$), and highly stable PSF. Nonetheless, because of the properties of the instrument we just detailed, its modelling will involve both the handling of spatial variations that occur both across a very large FOV *and* that include high spatial frequency variations caused by polishing errors in the mirrors. Star images will further be undersampled, which means a *super-resolution* step will be required to achieve a properly sampled PSF model.

Moreover, all shape catalogs up to now (see Table 2.2) have been established either from the ground, or with the HST. In either case, observations were carried out using fairly narrow filters. Once again, VIS was designed to optimize shape measurement over a very large number of objects, which, by necessity, means over very *faint* galaxies. A single, very wide band spanning most of the optical domain (from 550 to $900\mathrm{nm}$) was thus chosen. This, however, comes at the cost of needing to handle the *chromatic* variations of the PSF.

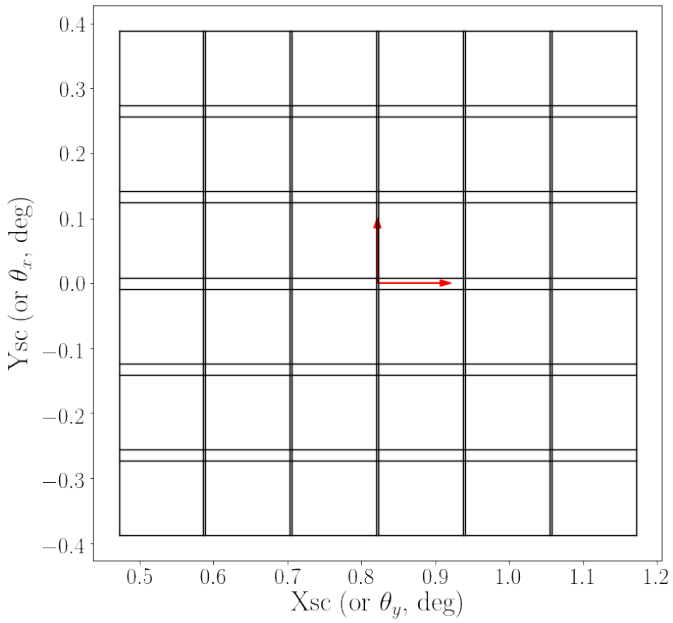


Figure 2.15: Schematic view of the VIS focal plane. Each almost-square region corresponds to one of the 36 CCDs. The gaps in between them are larger in one of the two directions (corresponding to either "Ysc" or "θ_x", two different conventions used internally within the EC).

None of the methods I presented in Section 2.4.3 can address all of these issues, save for the parametric approach of Duncan et al. (in prep). As we have already argued and illustrated throughout this section, it is highly desirable to *also* have a non-parametric approach to model the VIS (or any) PSF. Most of the work presented in this thesis, starting with the next chapter, is precisely aimed at working towards a non-parametric PSF model that can handle all of *Euclid*'s specificities.

Chapter 3

Monochromatic PSF field estimation

Contents

In this chapter, we start our PSF modelling efforts in a simplified setting. In particular, while the PSFs we will use to test our model are simulations of the VIS PSF, we do not yet take into account the chromatic variations, and instead assume the PSF to always be sampled at the same wavelength.

We begin, in the first half of this chapter, with a brief overview of the mathematical tools our proposed approach relies on. In Section 3.1, I introduce the fields of Inverse Problems, the use of sparsity as a regularization, and how such problems can be solved in practice. Section 3.2 gives a very brief introduction to the few concepts of Graph Theory we will use. Lastly, in Section 3.3, we finally apply all of these tools to the modelling and the interpolation of a (monochromatic) Euclid PSF. These last two subsections largely correspond to (the arXiv v1 of) Schmitz et al. (2019) (Paper II hereafter). They are the result of work done in collaboration with Jean-Luc Starck and Fred Ngolè, and that benefited from numerous comments from EC members, especially Lance Miller, Christopher Duncan, Henk Hoekstra, Peter Schneider and several of the other authors of Paper II; a complete list is given in Appendix A.

3.1 Sparse Signal Processing

The use of sparse representations has proven beneficial to a great many applications, and has deeply impacted the vast field of signal processing. It has also affected our everyday lives, for instance through considerable breakthroughs in data compression. A related area where sparsity has had a profound impact is that of *inverse problems*, which are precisely what our PSF modelling task will amount to. As with the previous sections, I will start with a general overview of the field, then quickly move on to the specific concepts we will require in order to build our first Euclid PSF model. For a more extensive view of Sparse Signal Processing, see the classical textbook of Mallat (1999), and that of Starck et al. (2015).

3.1.1 Inverse problems

We have already come to mention inverse problems a few times throughout this thesis; for instance, when performing galaxy shape measurement with model-fitting methods (Equation 2.3.7). Some instances of PSF modelling also rely on solving an inverse problem. In both cases, starting from observations (galaxies and stars, respectively) that are assumed to have resulted from some process F (called the *forward operator* or, in some cases, the "observation model"), the aim is to *invert* that process in order to recover some knowledge, X, that could not be measured directly. As an illustration, recall our Figure 2.13. There, an initial, finely sampled PSF H went through the process, F_d, of sampling on a pixel grid, leading to observations Y_i with coarser pixels. Suppose we started from these observations Y_i, considered the forward operator $F := F_\mathrm{d}$, and tried to recover the undegraded PSF, $X := H$, we would effectively be solving an inverse problem. It is, in fact, close to the one we will actually solve in Section 3.3, although it will then take a much more complex (and realistic) form. In the case of our other example, suppose we were fitting a Sérsic profile with intensity I, size R_e and (Sérsic) index n as free parameters, given an ellipticity (e_1, e_2) by application of a shear transformation. Starting with noisy galaxy observations and knowledge of the PSF H, the model-fitting step consists in solving the inverse problem with $X := (I, R_e, n, e_1, e_2)$, and

$$F : X \mapsto F_\mathrm{d}\left(\mathcal{S}(I, R_e, n) \circ \mathcal{G}_{(e_1,e_2)}\right), \tag{3.1}$$

where F_d is still the operator that samples on finite pixels,

$$\mathcal{S}(I, R_e, n) : (x, y) \mapsto I \exp\left(-b_n\left[\left(\frac{\sqrt{x^2+y^2}}{R_e}\right)^{1/n} - 1\right]\right) \tag{3.2}$$

is the Sérsic profile (with b_n a normalization factor), and the shearing operator is defined as

$$\mathcal{G}_{(e_1,e_2)} : (x, y) \mapsto \begin{pmatrix} 1-e_1 & -e_2 \\ -e_2 & 1+e_1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{3.3}$$

In practice, the forward operator is often not invertible. In those cases, inverse problems are solved by choosing some *cost function*, $\mathcal{L}$, that measures how well the chosen X reproduces the observations. The aim is then to find the value(s) of X that *minimizes* this cost, that is, to solve

$$\min_X \mathcal{E}(X) := \mathcal{L}(F(X), Y). \tag{3.4}$$

We already gave an example of such in (2.121), the cost function solved by `PSFEx` to build its PSF model. Solving (3.4) can be achieved by using *optimization* methods; we will present those used in the present case in Section 3.1.4.

3.1.2 Regularization

Before turning to the practical solving of inverse problems, let us first consider another example. Equation (2.119) allowed us to cast the PSF modelling problem as the choice of two matrices, S and A. We then saw, through `PSFEx` and `LensFit`, examples where A was fixed and S learned from the data, or S fixed and A learned, respectively. When performing PCA, neither are set in advance, and both S and A are computed from the data, with S being the eigenvectors and A the corresponding codes (that is, the coefficients given to each entry of S to reproduce the original datapoints).

PCA thus offers one specific way of solving the general inverse problem that amounts to finding two matrices, S and A, such that observations X are reconstructed through their combination:

$$X \approx SA. \tag{3.5}$$

This class of inverse problems, called *matrix factorization*, sets the framework we will use in the second half of this chapter (and generalize in the next).

Hadamard (1902) defined "well-posed" problems as those for which a solution exists, is unique, and well-behaved (in that it varies continuously with the initial conditions, or observations). Problems that do not meet these conditions are *ill-posed*. Such is the case of most inverse problems.

Consider the case of matrix factorization. For a given set of observations Y, let us choose

$$\mathcal{L}(Y, S, A) := \|Y - SA\|_{\mathrm{F}}^2, \tag{3.6}$$

with $\|.\|_{\mathrm{F}}$ the matrix Fröbenius norm, as our loss function. Suppose we found one solution, S^*, A^*, to our inverse problem. For any scalar α, define

$$\tilde{S^*} := \alpha S^*, \tag{3.7}$$

$$\tilde{A^*} := \frac{1}{\alpha} A^*. \tag{3.8}$$

$$\tag{3.9}$$

We have $\mathcal{L}(Y, S^*, A^*) = \mathcal{L}(Y, \tilde{S^*}, \tilde{A^*})$, and these new quantities are therefore also a solution. Matrix factorization is then indeed ill-posed, as there exists infinitely many solutions.

Recall how, after introducing the Einstein equations (2.4) in Chapter 2, we found ourselves in the need for additional assumptions (which we found in the Cosmological Principle) in order to be able to find solutions. Similarly, in order to solve ill-posed inverse problems, it is common to use regularization, that is, the addition of *constraints* on our final solution. A classical example is the Tikhonov regularization, i.e. the addition of an l_2 term to (3.4). The problem solved by `PSFEx` uses such a regularization, as shown in (2.121).

3.1.3 Sparse representation

Another extremely useful constraint is the use of *sparsity*. Consider the discrete, 1-dimensional (1D) signal $X(t) := \sin(\pi/2 + 2\pi t)$, shown in purple Figure 3.1a. We show its FFT, $\tilde{X}(k)$, in Figure 3.1b. Most of $\tilde{X}$'s coefficients are 0: we then say X is *sparse* in Fourier space.

Suppose we only have access to noisy observations Y, shown in black, in both spaces, in Figure 3.1. In order to recover our initial signal, we can make use of its sparse representation by adding a constraint. The most natural way is to use the l_0 pseudo-norm,

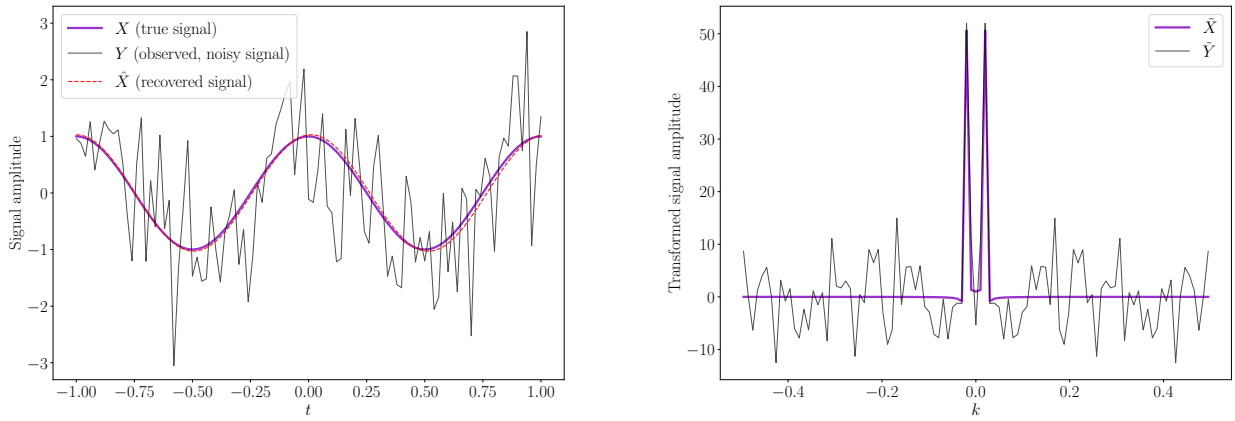
$$\|X\|_0 := \sum_i \mathbb{1}_{\{x_i \neq 0\}}, \tag{3.10}$$

that returns the count of non-zero entries. Let us then build our cost function as:

- an l_2 term to ensure our reconstruction is close to the observations, the *data-fidelity*;
- an l_0 term to favour solutions with few non-zero coefficients in Fourier space, the *sparsity constraint*.

The problem to solve would then be

$$\min_X \mathcal{E}(X) := \|Y - X\|_2^2 + \lambda \|\hat{X}\|_0. \tag{3.11}$$



(a) Signals in the direct domain.

(b) (Real part of) signals in the Fourier domain.

Figure 3.1: Toy example of signals with a sparse representation.

λ is a new hyperparameter that regulates the trade-off between how close we want to match the observed data, and the importance of our additional constraint. It is typically chosen by using some knowledge about (or estimation of) noise properties. We show the solution (see Section 3.1.5) to (3.11) in red in Figure 3.1a, for $\lambda = 20$. The signal is remarkably well recovered despite the poor apparent quality of the observations.

This successful recovery is because the assumption of sparsity in Fourier domain was exactly correct. That is not necessarily the case for all signals, of course. However, the same idea can be used by finding other, appropriate representation spaces for the signal of interest. *Wavelets* are often used to achieve this. They come in several different forms, but can be considered, roughly speaking, as a generalization of the Fourier transform. They have proven invaluable, for instance, in treating natural images. We once again point the reader to Mallat (1999)'s textbook for more details on wavelets and their applications. These have spun many fields, including astrophysics, where the *Starlet*, or Undecimated Isotropic Wavelet transform (Starck et al., 2011) has been used to tackle many inverse problems. The Starlet decomposition of a Euclid PSF is shown in Figure 3.2 as example.

3.1.4 Convex optimization

We now turn to the issue of solving (3.4). In some cases, the solution can be found analytically, e.g. when finding the minimum amounts to solving a linear system. In other cases, we must often turn to iterative schemes. If $\mathcal{E}$ is differentiable, a necessary condition for X^* to be a minimum is

$$\nabla\mathcal{E}(X^*) = 0. \tag{3.12}$$

If $\mathcal{E}$ is also *convex*, this condition is sufficient; if it is *strongly* convex, any such point will be the *unique* global minimum. The field of *convex* optimization aims at solving such problems. Nesterov (2018) provides a recently updated textbook covering the field. Let us note that even when the function we aim to minimize is *not* convex, it can still be possible (and sensible) to apply convex optimization tools. For instance, in some cases, we might not be interested in finding the global minimum of (3.4), and a local minimum might be sufficient for our purposes. This is often the case when the problem we are solving is aimed at learning a new representation space for our data, e.g. in Chapter 4. We will address the non-differentiable case later in Section 3.1.5.

In convex optimization, finding the point that verifies (3.12) is often achieved through *gradient descent*: starting at some initialization $X^{(0)}$, we will iteratively follow the direction of "steepest descent", that is, the one opposite the gradient:

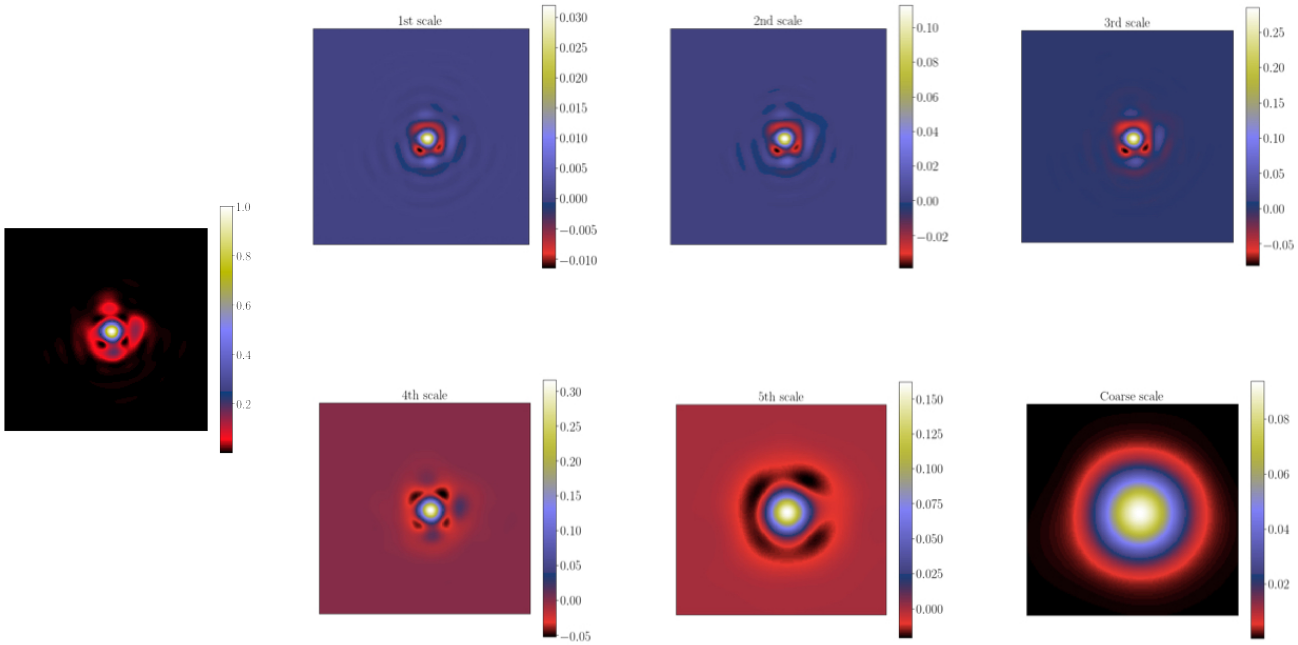


Figure 3.2: Example image (left) and its Starlet decomposition. The wavelet coefficient at each successive scale $j \leq 5$ is an image, of the same size as the original, that captures decreasingly finer variations. The "coarse scale", at the bottom right, contains the remaining information that is, by construction, a smooth approximation of the input image.

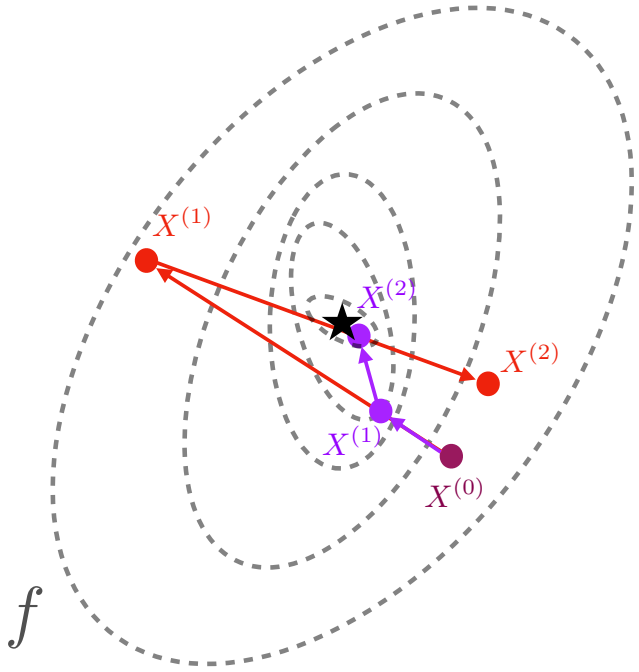


Figure 3.3: Illustration of gradient descent methods. The grey dotted lines denote isolines of the objective function, $\mathcal{E} = f$, and the black star its global minimum. The first two iterations (3.13) are shown, for an appropriate step size α (in purple) and one that is too large and leads to oscillations around the objective (in red).

$$X^{(t+1)} := X^{(t)} - \alpha \nabla \mathcal{E}(X^{(t)}), \tag{3.13}$$

where α is the *step size* and determines how far along this path we will move. As α becomes infinitely small, we are sure that $\mathcal{E}(X^{(t+1)}) \leq \mathcal{E}(X^{(t)})$. However, the larger α is, the faster we will move toward the minimum. This is illustrated in Figure 3.3. If the gradient of $\mathcal{E}$ is L-Lipschitzian, that is,

$$\forall (X, Y), \|\nabla \mathcal{E}(X) - \nabla \mathcal{E}(Y)\| \leq L\|X - Y\|, \tag{3.14}$$

then choosing $\alpha \leq 2/L$ ensures convergence. If that is not the case, or there is no way to compute L, α can instead be chosen empirically at each iteration using a *line search* method, for instance Armijo's rule (Armijo, 1966).

Finding X^* that verifies (3.12) amounts to finding the zeros of $\nabla \mathcal{E}$. A classical method to find the zeros of a differentiable, smooth function g is the *Newton* method,

$$X^{(t+1)} := X^{(t)} - \frac{g(X^{(t)})}{g'(X^{(t)})}. \tag{3.15}$$

When applicable, the Newton method comes with theoretical guarantees of quadratic convergence. It is therefore tempting to try and apply it in the case of our optimization problem by taking $g := \nabla \mathcal{E}$. This is thus only possible if $\mathcal{E}$ is *twice*-differentiable, and in this case its inverse Hessian matrix, $\left(\nabla^2 \mathcal{E}(X^{(t)})\right)^{-1}$ appears as the g' term in (3.15). Note that in this case, we can see the application of the Newton method as a particular case of the generic gradient descent in (3.13), with a particularly efficient choice of α, at each iteration, as a function of the Hessian.

It is often intractable to compute or invert the Hessian of $\mathcal{E}$ in practice. In the same way as α can be empirically chosen at each iteration when L cannot be directly computed, we can instead use an empirical approximation of the Hessian that is updated at each iteration. This scheme is called the *Quasi-Newton* method. It is a very efficient way to solve optimization problems when the cost function is smooth and there are no added constraints, i.e. when the need for regularization that we saw in Section 3.1.2 does not arise. The most famous Quasi-Newton method is the Limited memory BFGS (L-BFGS) (Morales and Nocedal, 2011), which we will use in Chapter 4 to solve a non-constrained optimization problem.

3.1.5 Proximal methods

Standard gradient descent and Newton methods both require the computation of the gradient of the cost function. This is impossible if it includes a non-differentiable term, as is often the case when we use regularization. For instance, recall our example of denoising with a sparsity constraint in Fourier domain of (3.11); the l_0 pseudo-norm is non-differentiable, and neither of the above approaches can be directly applied. Proximal methods (Parikh and Boyd, 2014) allow for such problems to be solved, and are an extremely powerful tool to solve constrained optimization problems. They rely on the the *proximal operator* (Moreau, 1962, "prox" hereafter), defined for a function g as

$$\text{prox}_g(Y) := \underset{X}{\text{argmin}} \left(\frac{1}{2} \|X - Y\|_2^2 + g(X) \right). \tag{3.16}$$

For a list of the properties of the prox as an operator, we refer the reader to Combettes and Pesquet (2011). Here, we focus on their use in solving minimization problem.

Several functions g that are commonly used for regularization of inverse problems have a closed-form prox. We shall call these functions *proximable*. Let us mention a few significant examples of proximable functions that we will use in the present work. For a set C, we define its indicator function as

$$\iota_C(X) = \begin{cases} 0 & \text{if } X \in C, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.17}$$

Note it is different to the (finite) indicator function we have introduced in (2.35). The addition of an $\iota_C(X)$ term to our cost function therefore enforces the constraint that the final solution should belong to C. This allows for a wide range of regularization. For example, if we want our solutions in $\mathbb{R}^p$ to contain only positive values, we can set $C := \mathbb{R}^p_+$. For a convex set C, the prox of ι_C is simply the orthogonal projection onto C.

Other important examples are sparsity constraints. In the example of Section 3.1.3, we introduced the l_0 pseudo-norm as a way to enforce sparsity. If $g(X) := \lambda \|X\|_0$, with λ a scalar parameter as in (3.11) then its prox is the *hard-thresholding* operator, defined as

$$\text{HT}_\lambda(X) = \left(\mathbb{1}_{\{|x_i| \geq \lambda\}} x_i \right)_i . \tag{3.18}$$

The l_0 norm is the natural way to enforce sparsity, but has the major caveat that it is not convex. It is therefore common to use the l_1 norm, as a convex relaxation, in its stead. Donoho and Huo (2001) show that the l_1 norm (or any l_p for $0 \leq p \leq 1$) does also promote sparsity and can, in some cases, recover the exact solution that would have been obtained by using the l_0. With $g(X) := \lambda \|X\|_1$, the prox is once again analytically known. It is the *Soft-Thresholding* operator,

$$\text{ST}_\lambda(X) = \left(\mathbb{1}_{\{|x_i| \geq \lambda\}} \left(x_i - \frac{|x_i|}{x_i} \lambda \right) \right)_i , \tag{3.19}$$

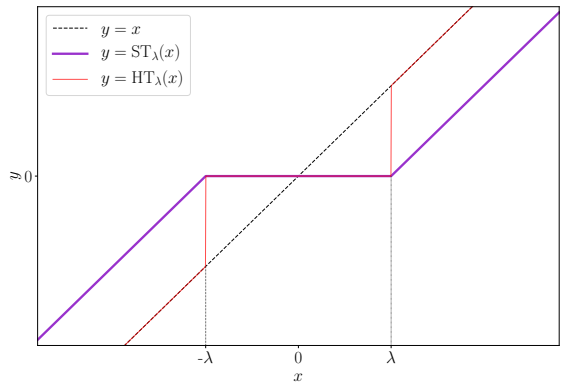


Figure 3.4: Hard- and soft-thresholding operators, for a given λ parameter.

shown along with HT, in the 1D case, in Figure 3.4.

Proximal methods aim at solving minimization problems that include both a differentiable and a proximable terms. For their theoretical foundation, proofs of properties and convergence, we refer the reader to the aforementioned general references (Combettes and Pesquet, 2011; Parikh and Boyd, 2014), or to those related to each specific proximal algorithm. Intuitively, one can think of them as a generalization of the gradient descent approach to the case where the loss function is not differentiable. As an illustration, consider problems of the form

$$\min_X \mathcal{E}(X) := f(X) + g(X), \tag{3.20}$$

with f smooth, differentiable and convex, and g proximable. The Forward-Backward splitting algorithm offers an iterative scheme to solve these problems, namely,

$$X^{(t+1)} := \mathrm{prox}_{\alpha g}\left(X^{(t)} - \alpha \nabla f(X^{(t)})\right). \tag{3.21}$$

Note the term inside the prox is the usual gradient descent iteration of (3.13) (the "forward step"). The application of the prox ("backward step") then ensures the constraint expressed by g is enforced. This is illustrated in Figure 3.5 in the case where $g := \iota_C$. Note that in this case, the forward-backward splitting is equivalent to the long-used "projected gradient descent".

In the case where there are more than one proximable function,

$$\min_X \mathcal{E}(X) := f(X) + \sum_k g_k(X), \tag{3.22}$$

e.g. because we are enforcing several constraints simultaneously, the problem can be solved by the Generalized Forward-Backward proposed by Raguet et al. (2013). Lastly, Condat (2013) propose a method to solve problems of the form

$$\min_X \mathcal{E}(X) := f(X) + \sum_k g_k(L_k X), \tag{3.23}$$

where the $(L_k)_k$ are bounded linear functions. This can, for instance, be used when we want to apply a sparsity constraint to our solution, but within a given representation space. For instance, in our previous example, the l_0

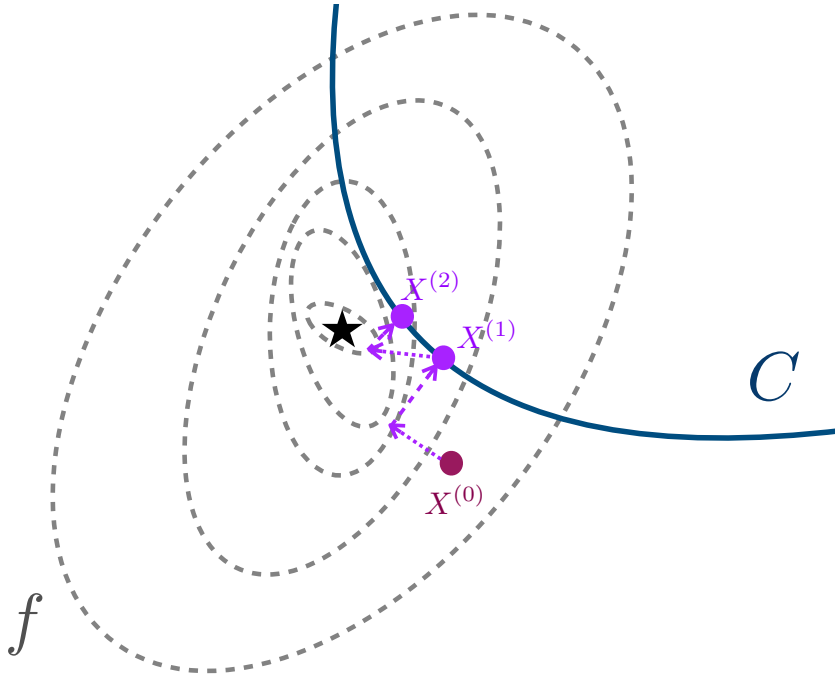


Figure 3.5: Illustration of the first two forward-backward splitting iterations, for $\mathcal{E} = f + \iota_C$, where f is the same as in Figure 3.3. The dotted arrows denote the forward step, corresponding to a gradient descent iteration (3.13), and the dashed arrows to the application of prox_{ι_C}.

norm was applied to $\tilde{X}$, the FFT of our solution X. Similarly, when using a non-orthogonal wavelet transform, Φ (e.g. the Starlet transform), the prox of $g : X \mapsto \lambda\|\Phi X\|_1$ does not have a closed form, and is often estimated iteratively. Alternatively, one can take $L_1 := \Phi$ in (3.23), as was for instance done by Farrens et al. (2017) to deconvolve galaxy images from a spatially-varying PSF. All three of these proximal methods are implemented in the freely available `ModOpt`[1] python package, that we used for the applications presented in the rest of this thesis.

3.1.6 Alternated minimization

In some cases, such as the matrix factorization example of Section 3.1.2, we want to minimize some cost $\mathcal{E}$ that is function of more than one variable, and not multiconvex. That is, while the problem might be convex with regards to each variable taken separately, it is not *jointly* convex in all of them. A standard example is that of *dictionary learning*, a particular case of matrix factorization, where we want to learn both a dictionary D and a set of codes Λ[2]. This problem typically takes the form

$$\min_{D,\Lambda} \|Y - D\Lambda\|_F^2 + g_D(D) + g_\Lambda(\Lambda), \tag{3.24}$$

where the g terms enforce some constraint on either or both of D and Λ. Several well-known variants of dictionary learning can be recovered with different choices of g_D, g_Λ. For instance, the problem of Non-negative Matrix Factorization (NMF) is that where both are chosen to be ι_+, by which we denote the indicator function of $\mathbb{R}_+^p$ for any p, the dimensionality of its inputs (in this case, either the dictionary or the codes). See Section 4.1.1 for a quick overview of dictionary learning methods.

[1] `https://github.com/CEA-COSMIC/ModOpt`

[2] In this context, Λ denotes a matrix containing a set of weights, and has nothing to do with the hypothetical cosmological constant discussed in Equation 2.1.5. This will also be the case throughout Chapter 4.

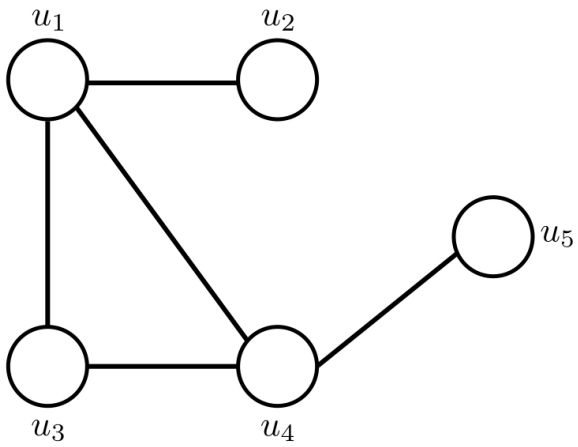


Figure 3.6: An example unweighted graph.

Such problems are often solved by *alternated minimization*, that is, by iteratively solving for one parameter (e.g. the dictionary D) while considering the others (say, the codes Λ) fixed. A great body of literature exists on alternated schemes, including proofs of convergence under certain assumptions. See, for instance, the Proximal Alternating Linearized Minimization of Bolte et al. (2010).

3.2 A primer on graph theory

In this section, we give a very brief introduction to some graph theory concepts that we will use in our building of a PSF model in next section. Let us first define the Laplacian matrix. Consider an unweighted graph such as that shown in Figure 3.6 (where each vertex i is identified through u_i). We define the *degree*, $d(i)$, of vertex i as the number of edges connected to it. In our example, we have $d(1) = d(4) = 3, d(2) = d(5) = 1$ and $d(3) = 2$. The degree matrix D is simply the diagonal matrix with $D_{ii} = d(i)$. It contains information about the graph's connectivity, but none about its actual structure (as we cannot know which vertices contribute to another's degree). This is contained in the *adjacency* matrix A, which in the unweighted case is simply defined as

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between vertices } i \text{ and } j, \\ 0 & \text{otherwise.} \end{cases} \tag{3.25}$$

We can now define the *Laplacian* matrix of a graph as

$$L = D - A. \tag{3.26}$$

In the example of Figure 3.6, the Laplacian would be

$$L = \begin{pmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 2 & -1 & 0 \\ -1 & 0 & -1 & 3 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}. \tag{3.27}$$

The Laplacian matrix is a central tool to graph theory. In the next section, we will introduce the *PSF Graph*, wherein each vertex will correspond to a star, and the edges will be weighted by a function of the distance in-between them. Generalizing the definition of L to the weighted case is an intuitive procedure. While the adjacency matrix entries up to now only contained a binary information (either two vertices are connected, or they are not), in the weighted case we replace that with the weight on the corresponding edge: the entries of A now tell us quantitatively *how* connected two vertices are. Similarly, for the degree matrix to quantify the amount of connectivity of a given vertex rather than just count the number of edges, we define the degree of node i as $d(i) = \sum_j A_{ij}$, that is, the sum of the weights carried by all edges connected to vertex i.

As a pathway toward defining wavelets on graphs, Hammond et al. (2011) introduced, by analogy with the usual transform, the Fourier transform on graphs. For a graph G with Laplacian L, let $(V_l)_l$ denote its eigenvectors. For any function f defined on the vertices of G, we define its Fourier transform as

$$\hat{f}(l) := \langle V_l, f \rangle. \tag{3.28}$$

Using this definition, one can then set constraints (e.g. sparsity) on the Fourier transform of a graph-valued function. We will use this tool to capture the spatial variations of the Euclid PSF in the next section.

3.3 Non-parametric Euclid PSF field recovery

In this section, we start building our non-parametric PSF modelling approach. We first consider a simplified setting that does not yet include all of *Euclid*'s specificities. It will already require us to use the tools we have presented in the first half of this chapter, and will also provide an opportunity to revisit the issue of PSF error propagation of Section 2.4.1. Indeed, we will compare our proposed approach to `PSFEx` in a *Euclid*-like setting, leading to two different imperfect PSF models we can use to explore the impact of PSF errors on galaxy shape measurement. In particular, by propagating the errors of both PSF models through different shape measurement methods, we examine whether the usual assumption that these two issues can be treated separately still holds for *Euclid*.

The aspects of the PSF, and other factors that could impact its modelling, that we leave aside in this section are the following. Binary stars can impact the PSF model if they are not removed from those objects used to fit the PSF model. Since previous work (Kuntzer et al., 2016; Kuntzer and Courbin, 2017) deals with the identification of such objects, we will assume throughout this thesis that they have already been removed. More generally, in this thesis, we will always work under the assumption that our star catalogs are empty from contamination (except for our work on CFIS data in Appendix B). Detector effects such as CTI and the BFE are also assumed to have been treated separately (see Section 2.3.1). Effects due to the AOCS and guiding errors, on the other hand, would contribute to the PSF as an extra convolutive term. Here, we focus on the *optical PSF*. In practice, the kernel of the extra convolution should be obtainable from precise time-series measurements of the instrument's AOCS. The *system PSF* could then be recovered by simply convolving together this kernel with our own optical PSF model.

In Section 2.5, we discussed the large band of the VIS instrument and its implications on the chromatic variations of the PSF (Cypriano et al., 2010; Eriksen and Hoekstra, 2018). No existing non-parametric approach can model these variations. We will propose the first one in Chapter 5. It relies on yet another class of mathematical tools, namely numerical OT, that we will introduce and apply in Chapter 4. In this section, we will

start our efforts by using the simulated VIS PSF at a single wavelength, which will allow us to use PSFEx for comparison.

The simulations used in this section (further described in Section 3.3.4 and Ngolè et al., 2015) did not yet include manufacturing and polishing errors. These can induce variations of the PSF that occur on very small spatial scales. As discussed below, the proposed method can by construction handle these high-spatial frequency variations (with the strong caveat that observations need to fall within the area of variation for our model to account for it).

Lastly, the modelling we carry out in the present work is done at a single point in time, with a low number of observed stars. As we already saw was the case for the HST, the telescope will in truth vary with time, which means the PSF modelling should be performed on each exposure separately (or on a set of exposures taken closely enough in time that the temporal variation can be neglected). However, another approach is to include the temporal variation within the non-parametric model itself, and fit it either to several exposures simultaneously, or in an online way, updating the model with each new available exposure. Not only could this improve the quality of the PSF model, it might also help mitigate two serious limitations of the non-parametric method: its quality depending on the number of stars available, and the aforementioned need to observe one precisely at the position of high-spatial frequency variations (which should be constant with time).

In our simplified setting, the two main issues to tackle are the spatial variations of the PSF and the undersampling of star images. Of the methods we presented in Section 2.4.3, PSFEx remains the most widely used method and is, to the best of our knowledge, the only one to deal with both the super-resolution and the spatial variation steps at the same time. Mandelbaum et al. (2017) found the PSFEx-based model of the HSC PSF to perform poorly when seeing becomes better than a certain value, close to the threshold at which PSFEx automatically switches to the super-resolution mode (Bosch et al., 2017), and could indicate issues with PSFEx's handling of super-resolution (and the need for other non-parametric methods to deal with this problem). Super-resolution is a well-studied problem in image processing, where sparsity-based methods (Starck et al., 2015) have been shown to perform extremely well. Ngolè et al. (2015) showed this to hold true in the particular case of PSFs. However, contrary to the typical setting of the super-resolution problem where the object of interest is observed several times with slight shifts (Rowe et al., 2011), in the case of *Euclid*, we instead have several undersampled observations of the PSF *at different positions* in the FOV. Ngolè et al. (2016) recently introduced Resolved Components Analysis (RCA), a method specifically designed to handle such a problem, but estimating the PSFs only at star positions.

In this section, we expand the RCA method by capturing spatial variations of the PSF through a set of *PSF graphs*. We can thus estimate the PSF at any arbitrary position in the field, while preserving all the properties of the RCA-recovered PSF field. This leads to a new approach that can deal, similarly to PSFEx, with both super-resolution and spatial variations. The python library is freely available[3].

The official *Euclid* requirements for WL are that the overall multiplicative shear bias should be lower than 2×10^{-3}. As discussed in Section 2.4.1, using the approach of Cropper et al. (2013), this in turn leads to stringent requirements on the PSF model accuracy: the root mean square (RMS) error on each ellipticity component $\left(e_i^{\mathrm{PSF}}\right)_{i\in\{1,2\}}$ should be lower than 5×10^{-5}, and that on the relative size $\delta R^2_{\mathrm{PSF}}/R^2_{\mathrm{PSF}}$ (as defined in Section 2.3.7) lower than 5×10^{-4}. As we will see in the experiments of this section, the proposed approach cannot yet, on its own and using only a small number of stars, achieve these requirements. It does, however, already provide an improvement over PSFEx.

The rest of the section is organized as follows: Section 3.3.1 describes the formalism of the PSF recovery field problem we adopted, Section 3.3.2 gives a quick overview of the RCA method and Section 3.3.3 presents the new PSF field recovery method. In Section 3.3.4, we apply both PSFEx and our approach to recover simplified *Euclid*-like PSFs and compare the resulting models. We then use them for galaxy shape measurement and study the impact of PSF modelling errors in Sections 3.3.5 and 3.3.6. We give some partial conclusions to this chapter in Section 3.3.7.

[3] https://github.com/CosmoStat/rca

3.3.1 Modelling the PSF field from stars

Notations

Let us first describe the problem at hand. Let $\mathcal{H}(u)$ denote the (unknown) PSF that we wish to estimate; $\mathcal{H}$ is a function of $u = (x, y)$, a 2D vector containing the position within the image. It is the full, untruncated PSF intensity profile, and thus outputs a continuous image at any position u. Here and throughout this paper, all such positions will be given in "image" coordinates (i.e. within the instrument's CCDs), since the position of objects on the sky has no influence on the PSF they are affected by. Similarly, here we consider $\mathcal{H}$ to be a single PSF that aggregates all effects (e.g. diffraction, imperfect optical elements). In particular, we do not consider the intermediary, relative position of incoming light rays from a given object on each individual optical component. We also consider the spatial variations of the PSF to be slow enough that the entirety of an object whose center lies at position u is affected by the same $\mathcal{H}(u)$.

Assume we observe a set of n_{stars} stars across the FOV, at positions $\mathcal{U}_{\text{stars}} := (u_1, \dots, u_{n_{\text{stars}}})$. Each star i gives us a noisy, undersampled observation of $\mathcal{H}$:

$$Y_i = \mathcal{F}\left(\mathcal{H}(u_i)\right) + N_i, \tag{3.29}$$

where N_i is a noise vector and $\mathcal{F}$ is the degradation operator, i.e. the effect of the realization on the instrument's CCDs. In our case, we separate its effects in two distinct operators,

$$\mathcal{F} = F_{\text{d}} \circ \mathcal{F}_{\text{s}}. \tag{3.30}$$

$\mathcal{F}_{\text{s}}$ is a discrete sampling into a finite number of pixels, which turns each continuous image $\mathcal{H}(u)$ into a truncated $p \times p$ image sampled at our target pixel size. F_{d} is composed of a sub-pixel shift (as illustrated in Figure 2.13), and a further downsampling matrix M (i.e. the pixel response of our instrument) that can lead to undersampling. Denoting by D the downsampling factor caused by M, the available observations Y_i are thus $Dp \times Dp$ images. In the following, we will treat them as flattened vectors of size D^2p^2.

The problem at hand is composed of the two following parts:

- from observations $Y := (Y_1, \dots, Y_{n_{\text{stars}}})$, build an estimator $\hat{H}$ of the true PSF $\mathcal{H}$ at corresponding positions $\mathcal{U}_{\text{stars}}$;
- recover the PSF at the galaxy positions, $\mathcal{U}_{\text{gal}} \neq \mathcal{U}_{\text{stars}}$.

In our present case of undersampled observations, while still discretized, we want our PSF model $\hat{H}$ to be sampled on a finer grid than observations $(Y_i)_i$, that is, to counter the effect of F_{d}.

`PSFEx`

We gave a brief introduction to `PSFEx` in Equation 2.4.3. Recall from (2.121) that it built its model by solving an inverse problem with a χ^2 data fidelity term and a Tikhonov regularization. Using the notations we just introduced, the former can be defined as

$$\chi^2(\Delta S) = \sum_{i=1}^{n_{\text{stars}}} \left\| \frac{Y_i - F_{\text{d}}\left([S_0 + \Delta S]A_i\right)}{\hat{\sigma}_i} \right\|_2^2, \tag{3.31}$$

where $\hat{\sigma}_i^2$ contains the estimated per-pixel variances.

Here, we include the flux normalization, sub-pixel shifting and potential downsampling (if super-resolution is required) operators in F_{d}. Shifting of the PSF models to the same grid as those of the observed stars is performed, both within `PSFEx` and for our proposed approach in the upcoming section, through the use of a Lanczos interpolant (as opposed to `LensFit` that uses, as described in Equation 2.4.3, a full sinc interpolation).

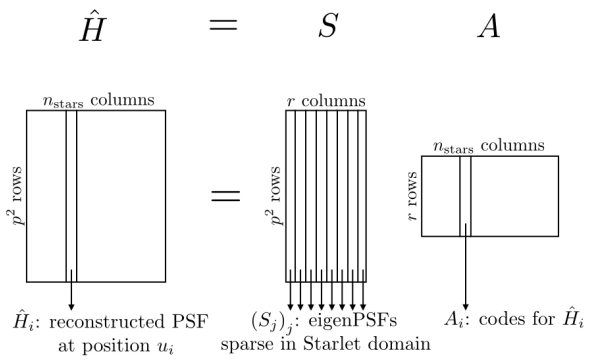


Figure 3.7: RCA's matrix factorization.

3.3.2 Resolved Components Analysis

Overview

RCA (Ngolè et al., 2016), like many other methods (including `PSFEx`, as shown in Equation 2.4.3), achieves super-resolution through matrix factorization. The PSF at the position u_i of each star is reconstructed through a linear combination of a set of eigenPSFs, S_j:

$$\hat{H}_i^{\text{RCA}} := \hat{H}^{\text{RCA}}(u_i) = \sum_{j=1}^{r} S_j A_{ij} = S A_i, \tag{3.32}$$

where each eigenPSF S_j is an image of the same size as the PSF images. Introducing the set of all reconstructed PSFs at star positions, $\hat{H} = \left(\hat{H}_1, \ldots, \hat{H}_{n_{\text{stars}}}\right)$, we thus have the matrix formulation illustrated in Figure 3.7.

Because our data is undersampled, a strong degeneracy needs to be broken: infinitely many finely sampled PSFs would manage to reproduce the observed undersampled stars. RCA manages to break this degeneracy by enforcing several constraints on both S and A that reflect known properties of the PSF field.

1. *Low rank*: the PSF variations across the field should be capturable through a small number of eigenPSFs r. This can be enforced by choosing S to be of dimension $p^2 \times r$, with $r \ll p^2$.

2. *Sparsity*: the PSF should have a sparse representation in an appropriate basis, which can be enforced through a sparsity constraint on the eigenPSFs.

3. *Positivity*: the final PSF model should contain no negative pixel values.

4. *Spatial constraints*: variations of the PSF across the field are highly structured, and the smaller the difference between two PSFs' positions u_i, u_j, the smaller the difference between their representations $\hat{H}_i, \hat{H}_j$ should be.

The last of these constraints is achieved through a further factorization of A itself, described in the following subsection.

Spatial regularization on graph

The spatial variations of the PSF across the FOV is highly structured, with both smooth variations that take place across the whole field, and some much more localized changes that only affect PSFs in a small part of it. If we had access to evenly spaced samples, this would amount to variations occurring at different (spatial) frequencies. We could then capture these variations by making each of our eigenPSFs contain information related to a given spatial frequency. Our sampling of the PSF is, however, dependent on the position of stars in the field over which we have no control.

RCA overcomes this hurdle through the introduction of *graph harmonics*: each row A_k of A, which contains the weights given to all observed star positions for eigenPSF k, is associated with a graph. For $k \in \{1, \dots, r\}$, let P_{e_k,a_k} denote the Laplacian (up to a constant multiplication on the diagonal) of the graph associated with A_k (and thus to the kth eigenPSF). We define it as

$$\begin{aligned}(P_{e_k,a_k})_{ij} &:= \frac{-1}{\|u_i - u_j\|_2^{e_k}} \qquad \text{if } i \neq j,\\ (P_{e_k,a_k})_{ii} &:= \sum_{\substack{j=1\\ j\neq i}}^{n_{\text{stars}}} \frac{a_k}{\|u_i - u_j\|_2^{e_k}}.\end{aligned} \tag{3.33}$$

From the definition of the Laplacian for a weighted graph given in Section 3.2, we can see that these P_{e_k,a_k} are precisely the Laplacian matrices of a fully connected graph with the edge between i and j weighted by $1/\|u_i - u_j\|_2^{e_k}$, multiplied (entry-wise) by a matrix $\tilde{L_{\text{D}}}$, defined as

$$\tilde{L_{\text{D}}} := a_k \text{Id} - \begin{pmatrix} 0 & -1 & -1 & \dots & -1 \\ -1 & 0 & -1 & \dots & -1 \\ \vdots & & \ddots & & \vdots \\ -1 & \dots & -1 & 0 & -1 \\ -1 & \dots & -1 & -1 & 0 \end{pmatrix}. \tag{3.34}$$

In other words, each of our r PSF graphs are fully connected graphs with the edge between positions u_i and u_j given a weight of $1/\|u_i - u_j\|_2^{e_k}$. The role of e_k in associating each of our graphs to a certain spatial frequency is straightforward: the higher its value, the stronger the decay in edge weight as the distance between two vertices increases, leading to the graph capturing lower spatial frequencies. Comparing (3.34) to (3.26) gives an intuitive (though unrigorous) interpretation as to the role of a_k: it amounts to multiplying the degree matrix of our graph by a_k, in turn affecting its overall connectivity.

By carefully choosing the parameters of our set of graphs, $(e_k, a_k)_{k\in\{1,\dots,n_{\text{stars}}\}}$, we can thus make each of them sensitive to different ranges of distances, which leads to the harmonic interpretation. See Ngolè et al. (2016, particularly Sections 5.2, 5.5.3 and Appendix A) for more details, as well as a scheme to select appropriate $(e_k, a_k)_k$ from the data.

We enforce the link between A's rows and their corresponding graph through the addition of a constraint on the former. Namely, we want to preserve the graph's geometry through A so that the amplitude of coefficients associated with a certain eigenPSF varies with the right spatial harmonics. We achieve this as follows: since P_{e_k,a_k} is real and symmetric, we decompose it as

$$P_{e_k,a_k} := V_{e_k,a_k} \Sigma_{e_k,a_k} V_{e_k,a_k}^\top, \tag{3.35}$$

where Σ_{e_k,a_k} is diagonal. Let $V := (V_{e_1,a_1}, \dots, V_{e_r,a_r})$ the matrix made up of the eigenvectors associated with each of our r PSF graphs. Our spatial constraint can now be expressed by further factorizing A by $V^\top$, and forcing the

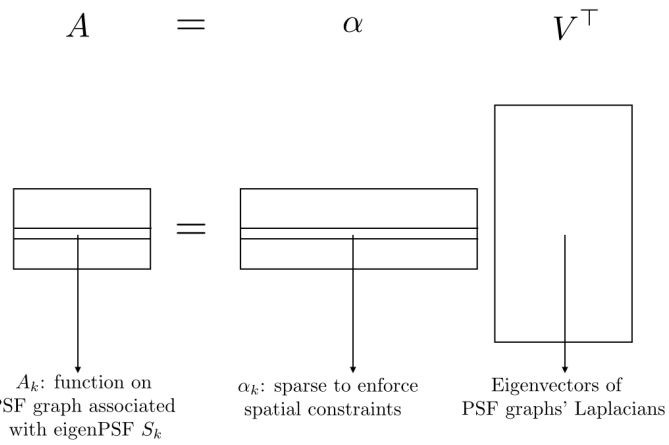


Figure 3.8: Matrices involved in the spatial constraint.

resulting coefficients α to be sparse. This is illustrated in Figure 3.8. At the end of Section 3.2, we introduced the Graph Fourier transform of Hammond et al. (2011). Here, each line A_k containing the coefficients of each star for a particular eigenPSF is a function on the associated PSF Graph. From (3.28), we can see that the matrix V introduced in this subsection is nothing but the concatenation of the eigenvectors associated to each eigenPSF's graph. Factorizing A by $V^\top$ and imposing the rows of the resulting matrix α to be sparse thus simply amounts to imposing the coefficients associated to our eigenPSFs to be sparse in the Fourier domain of each associated graph (themselves capturing, by construction, a particular spatial frequency).

As mentioned at the beginning of this section, manufacturing and polishing defects in the VIS instrument will inevitably lead to very localized, but strong variations of the PSF at some (fixed) positions. While these are not yet included in the simulated PSFs we use in Section 3.3.4, they should be naturally handled by our proposed approach with the addition of extra eigenPSFs. Each of these additions would diminish the role of constraint 1 (low rankness), but the added graph (and corresponding eigenPSF) would capture only those very localized changes in the PSF. However, all this can only be accomplished if some of the observed stars do fall within the area where these variations occur. As we already discussed, this caveat could be alleviated by adding a temporal component to our model, and fitting it on stars extracted from several different exposures.

Optimization problem

Combining the factorizations illustrated in Figures 3.7 and 3.8, reconstruction of the PSF field at the star positions through RCA amounts to solving the following problem:

$$\min_{S,\alpha}\left(\frac{1}{2}\|Y - F_{\mathrm{d}}(S\alpha V^\top)\|_{\mathrm{F}}^2 + \sum_{i=1}^{r}\|w_i \odot \Phi s_i\|_1 + \iota_+(S\alpha V^\top) + \iota_\Omega(\alpha)\right), \tag{3.36}$$

where $(w_i)_i$ are weights, $\odot$ denotes the Hadamard (or entry-wise) product, Φ is a transform through which our eigenPSFs should have a sparse representation (in our case, Φ will always be the Starlet transform, Starck et al., 2011), ι_+ is the previously introduced positivity indicator function, that is,

$$\iota_+ : X \mapsto \begin{cases} 0 & \text{if no entry of } X \text{ is strictly negative,} \\ +\infty & \text{otherwise.} \end{cases} \tag{3.37}$$

Ω is a sparsity enforcing set,

$$\Omega := \left\{\alpha, \forall i \in \{1, \dots, r\}, \|(\alpha^\top)_i\|_0 \leq \eta_i\right\}, \tag{3.38}$$

and ι_Ω its indicator function. Here, α belongs to Ω if each of its row i has at most η_i non-zero entries.

Breaking down (3.36) into its four summands, we can get a sense of how solving it would yield a PSF model that fits the observations while satisfying the list of constraints we introduced at the beginning of Section 3.3.2. Indeed, the first term is the data fidelity term, which ensures we recover the observed star images after applying the correct undersampling operator. The second term promotes the sparsity of our eigenPSFs, thus satisfying constraint 2. The third term ensures our PSF model only contains positive pixel values, enforcing constraint 3. The fourth term forces the learned α to be sparse, in turn satisfying constraint 4 as detailed in item 3.3.2. Lastly, as mentioned above, constraint 1 is achieved by setting a low enough number of eigenPSFs r.

Finding the eigenPSFs and their associated coefficients for each star amounts to solving (3.36). This can be done through alternated minimization, that is, by solving in turn for S then for α iteratively. Each minimization is performed through the use of the proximal methods we presented in Section 3.1.5. For more details on how parameters $(e_k, a_k)_k$, r, $(w_i)_i$ and $(\eta_i)_i$ are set, we refer the reader to Ngolè et al. (2016).

3.3.3 PSF Field Recovery from Graph Harmonics

Spatial interpolation of the PSF

We now turn to the problem of interpolating our PSF model from the positions of stars, $\mathcal{U}_{\text{stars}}$, to that of galaxies, $\mathcal{U}_{\text{gal}}$.

As discussed in Figure 2.4.3, several standard methods exist to perform spatial interpolation, that is, to estimate the (unknown) value of some function f at a new position $u = (x, y)$ given its measurements at several other positions: $(f(u_k))_k$. The most natural (and the one used by `PSFEx`) is probably the use of a polynomial function of positions, i.e.

$$\hat{f}(u) = \sum_{\substack{i,j \geq 0 \\ i+j \leq d}} x^i y^j Q_{ij}, \tag{3.39}$$

where the maximum polynomial degree d is a user selected parameter and the $(Q_{ij})_{i,j}$ are chosen such that $\hat{f}(u_k) \approx f(u_k)$ at every position where f was observed (in our case, $u_k \in \mathcal{U}_{\text{stars}}$). Note that the particular set-up of `PSFEx` shown in (2.120) can be recovered when choosing $f := \hat{H}^{\text{PSFEx}}$ and Q_{ij} the `PSFEx`-learned, image-sized components.

An alternative to the polynomial approach is the use of Radial Basis Function (RBF) (Buhmann, 2003). An RBF is a kernel φ that only depends on the distance between two points. The polynomial formulation of (3.39) can then be replaced by

$$\hat{f}(u) = \sum_{i=1}^{n_{\text{neighbors}}} Q_i \varphi(\|u - u_i\|), \tag{3.40}$$

where we sum over observed positions of the closest neighbors of u, and the $(Q_i)_i$ are, once again, chosen so that $\hat{f}$ coincides with the observed f at all sampled positions. Broadly speaking, the idea behind these schemes is that the closer a position u_i is, the more its observed value $f(u_i)$ should contribute to the estimated $\hat{f}(u)$, and RBF interpolation can be thought of as a generalization of inverse distance weighting schemes. Note that an assumption underlying the use of RBFs is that the PSF's amount of similarity to its neighbors in f is isotropic, i.e., the same in every direction.

Because of its simplicity and good performance exhibited in Gentile et al. (2013), we chose to use RBF interpolation in the present work, and selected the commonly used thin plate RBF kernel (see Ngolè and

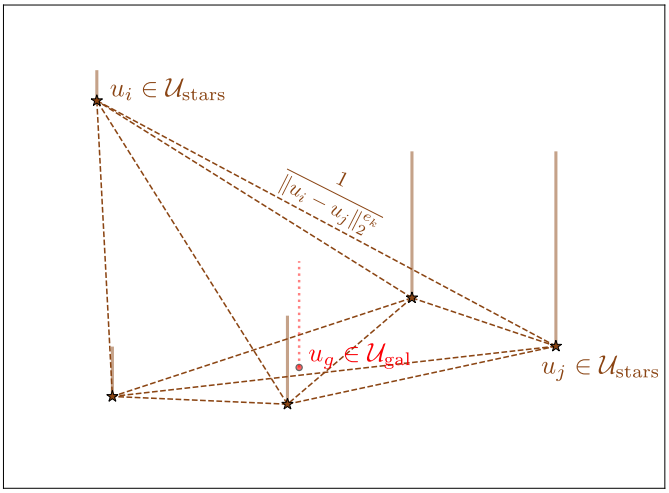


Figure 3.9: Graphical representation of the PSF graph associated with eigenPSF S_k. The height of the vertical bar at each position u_i corresponds to the amplitude of coefficient A_{ki}.

Starck, 2017, Section 3.2, for a quick discussion of its physical interpretation). In what follows, we always set $n_{\text{neighbors}}$ to 15.

Spatial Regularity using RCA Graph Harmonics

Aside from the choice of the spatial interpolator discussed above, one must also decide *which* function f to interpolate. In our case, where the PSFs are images of p^2 pixels, the simplest approach would be to consider each of these pixels as a scalar function and interpolate it independently from the others. This is what is done by the `LensFit` model (see Equation 2.4.3). While simple, this approach is extremely sensitive to single-pixel fluctuations, which are not unexpected in our data-driven estimations of the PSF, for instance if some noise-related artefacts remain.

As mentioned, instead of using p^2 different f scalar functions, `PSFEx` instead considers f to be R^{p^2}-valued. By construction, it performs a polynomial interpolation of its learned components. Spatial interpolation can also be carried within any chosen basis of representation – a typical example being the use of PCA we discussed in Equation 2.4.3, wherein spatial interpolation is carried over the coefficients associated with the first few principal components.

Our proposed approach is to perform this spatial interpolation step within the Graph Harmonics framework of RCA. We showed in item 3.3.2 that the rows of matrix A are functions on a set of graphs, each containing the spatial information related to one particular eigenPSF. This is illustrated in Figure 3.9: the coefficients related to eigenPSF S_k encode the spatial variations for a given range of distances. By performing the spatial interpolation in each of the r rows of the RCA-learned A matrix, we are moving along each of the corresponding PSF graphs. For any new position u, we can then reconstruct a new set of coefficients $A_u \in \mathbb{R}^p$ through r RBF applications as in (3.40), and reconstruct the PSF as

$$\hat{H}(u) := S A_u. \tag{3.41}$$

This amounts to adding a new point on the PSF graphs, as shown in red in Figure 3.9. Since S was learned from the observed stars and A_u preserves the graph harmonics, this step ensures the constraints we highlighted at the end of Section 3.3.2 are still applied to the new PSF at the galaxy positions. In particular, the spatial constraints are preserved thanks to the PSF graphs.

Note that an additional advantage to this approach lies in the fact that the most computation-intensive steps are performed during the reconstruction of the PSF field through RCA (Section 3.3.2). In a *Euclid*-like

framework where star images are undersampled, if we were to use RCA to perform the necessary super-resolution step, the dictionary S and the graph harmonics encoded in A would already be computed. The proposed method can thus perform the spatial interpolation step in a particularly appropriate representation at no additional computational cost save for that of fitting the RBF weights. Conversely, if one wanted to use any other representation, even one as simple as PCA would require the extra computation a spectral value decomposition (SVD).

3.3.4 Comparison of PSF models

Data set

The PSFs we use are simulations of *Euclid*'s VIS PSFs (as described in Ngolè et al., 2015, Section 4.1), located in the central part of the FOV, sampled at a single wavelength of $600\,\mathrm{nm}$. As mentioned at the beginning of Section 3.3, this is a simplification of the true Euclid PSF, since we neglect its chromatic variations, and the detector and guiding effects are absent from the simulations. This data set contains 597 such PSFs, each consisting of a 512×512 stamp with a pixel size of approximately 0.0083 arcsecond, i.e. sampled on a much finer grid than the *Euclid* pixel size. See Figure 3.10 for two examples chosen at some of the top-right- and bottom-left-most positions.

As previously discussed, in a real-life observing situation, the only information (in the non-parametric framework of this thesis) from which we would derive our PSF models would be obtained from stars within the field, which lie at positions different from those where we wish to estimate it. We thus randomly split our sample of PSFs into two parts:

- a training set of 300 PSFs, the position of which we will refer to as "star positions";
- a test set with the remaining 297 PSFs, the position of which we will refer to as "galaxy positions".

The number of stars in our training sample is of order 10 times smaller than the expected average number of usable stars present in a VIS science exposure, though using all the available stars simultaneously would require taking into account the variations of the PSF across different CCDs. In Figure 3.3.4, we will only use the PSFs at star positions to try and produce estimations of the PSF at the galaxy positions. Conversely, from Section 3.3.5 on, we will solely focus on and use objects at the galaxy positions.

Euclid's sampling frequency is at 0.688 of the Nyquist rate, which sets our goal in terms of super-resolution at achieving an upsampling factor of $1/D = 2$ (Cropper et al., 2013). To simulate observed stars, we sample all 300 PSFs in the training set at the nominal *Euclid* pixel scale of 0.1 arcsecond. This is achieved by first applying a mean filter (which amounts to the approximation that the VIS pixel response is a perfect top hat), then sampling pixels at the correct rate. We apply a random sub-pixel shift to each resulting image, then truncate the stamps to be of size 21×21 around the pixel closest to the object's centroid. Indeed, in observing situations, our PSF models would likely (definitely, in the cases of both `PSFEx` and RCA) be fitted on image stamps containing a suitable star extracted from the full image as described in Section 2.3.3. Our models would thus necessarily need to deal with the resulting truncation effects. Lastly, we add various levels of white Gaussian noise with standard deviation σ, yielding five different sets of observed stars at average SNR of $10, 20, 35$ and 50, where the SNR is defined as

$$\mathrm{SNR} = \frac{\|x\|_2^2}{\sigma^2 p^2}, \tag{3.42}$$

for image x of size $p \times p$. An example of the resulting star images is shown in Figure 3.11.

From these, we will estimate the PSF at twice the *Euclid* pixel sampling and at the galaxy positions in Figure 3.3.4. Because of the upsampling, the resulting PSFs will be stored in stamps of 42×42 pixels.

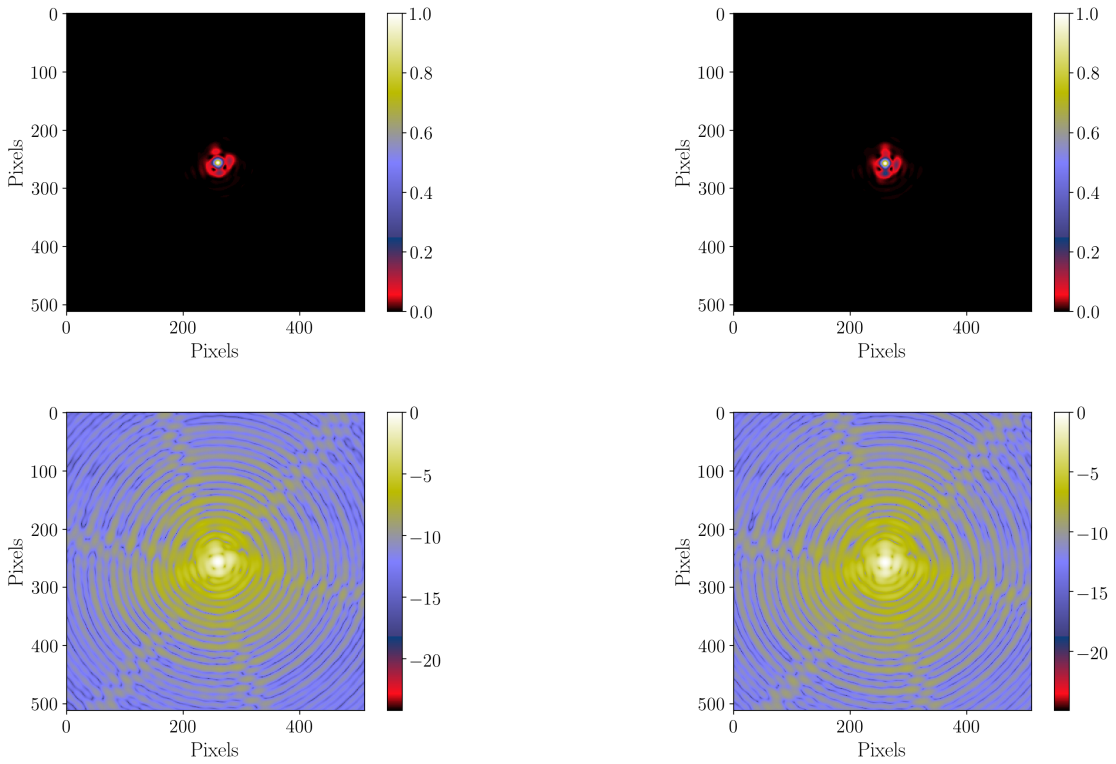
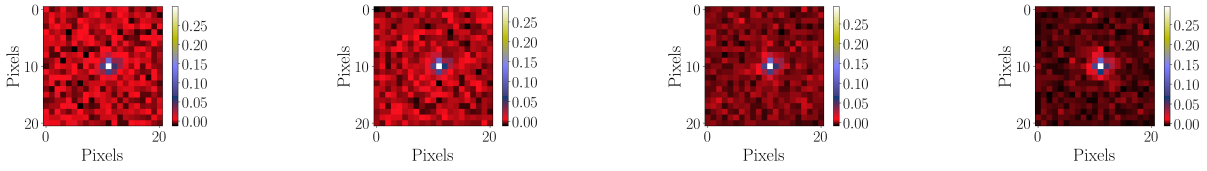


Figure 3.10: Visual examples of the simulated Euclid PSF in the natural (top row) and logarithmic (bottom row) domains, at the original pixel sampling of the simulation (about 12 times finer than *Euclid*). Each stamp is approximately 4.25 arcsecond across.

For comparison purposes, we also prepare a set of "known" PSFs $\hat{H}^{\mathrm{kn}}$ at those positions, by sampling the 297 test PSFs at half of *Euclid*'s pixel size and truncating the resulting images to 42×42 pixels. While not the ideal case (where the continuous PSF image would be perfectly known), this fiducial, unattainable case amounts to the best possible PSF our approaches to super-resolution, denoising and spatial interpolation could possibly achieve. Note that this would require some extra conditions to be met, e.g. by the population of random shifts undergone by the undersampled images (which is, under the safe assumption that shifts are randomly distributed, also directly related to the number of observed stars). Using the notations of Section 3.3.1, $\hat{H}^{\mathrm{kn}}$ would be the the PSF obtained if only $\mathcal{F}_{\mathrm{s}}$ remained while F_{d} had been perfectly corrected for. In other words, the only effects degrading these PSFs are those of sampling (at our target of half *Euclid*'s pixel size), and truncation at the best possible stamp size given that of our star images.



(a) SNR 10 (b) SNR 20 (c) SNR 35 (d) SNR 50

Figure 3.11: Example observed (undersampled) star stamps, at the various noise levels considered, from which the PSF models will be estimated. Each stamp is approximately 2.1 arcsecond across.

PSF modelling

We first assume the star images described in Section 3.3.4 were already extracted, and we perform both super-resolution and spatial interpolation using RCA, as described in subsections 3.3.2 and 3.3.3. This yields a set of 297 RCA-estimated 42×42 PSFs, $\hat{H}^{\text{RCA}}$, at galaxy positions per SNR level.

PSFEx was designed to run on catalogs extracted using companion software SExtractor, and thus requires a little more setting up. For each SNR level, we first create a fake full image of $12\,000 \times 12\,000$ pixels, into which the 300 stars are placed at their respective positions[4]. We then run SExtractor on the resulting images, with parameters selected so that all stars are detected and extracted correctly, and no spurious detections occur. When ran, PSFEx performs a further selection across all objects extracted by SExtractor, which is usually desirable to have the PSF model fitted to appropriate stars. In our case, however, since we already know our SExtractor catalog to be perfect, we tune PSFEx's selection parameters so that as many stars as possible are used. One is nonetheless rejected at SNR 50. The parameters related to the model are the following:

```
PSF_SAMPLING     .5
PSF_SIZE         42,42
PSFVAR_KEYS      X_IMAGE,Y_IMAGE
PSFVAR_GROUPS    1,1
PSFVAR_DEGREES   2
```

Namely, PSFEx learns a set of PSF basis elements $\left(S_{ij}\right)_{i,j}$ such that the PSF at position x, y is estimated as in (2.120), with $d = 2$ (repeating the experiments with $d = 3$ led to very poor PSF models). All other PSFEx parameters are left to their default value. Again, this gives us one set of estimated PSFs per SNR level, $\hat{H}^{\text{PSFEx}}$, with the same stamp and pixel sizes as $\hat{H}^{\text{kn}}$ and $\hat{H}^{\text{RCA}}$.

Examples of $\hat{H}^{\text{kn}}$, $\hat{H}^{\text{RCA}}$ and $\hat{H}^{\text{PSFEx}}$, at the galaxy position corresponding to the simulated PSF on the left-hand side of Figure 3.10, are given in Figure 3.12 for the worst and best-case noise scenarii.

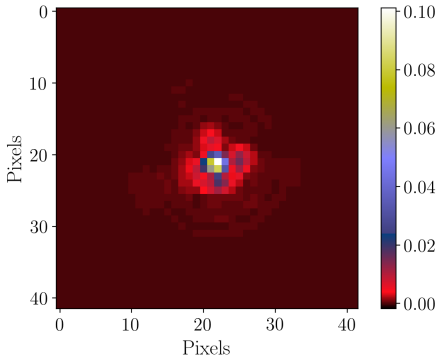
Results

In Section 2.4.1, we saw that Equation (2.109) was often used as a basis to study the quality of a PSF model. As discussed then, it is unclear whether this equation would hold true in the case of *Euclid*. Before answering this question, we first quantify the quality of our PSF by looking at the errors that appear in (2.109), i.e. $\delta(R^2_{\text{PSF}}), \delta e_i^{\text{PSF}}$, for both models, on the 297 test PSFs.

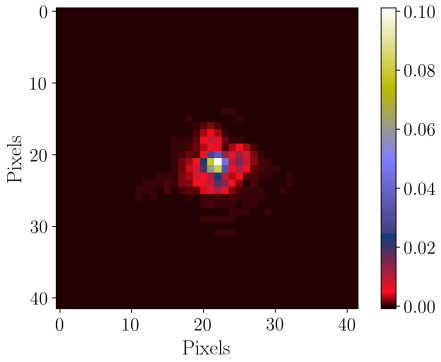
For the former, both models tend to overestimate the size of the PSF, likely because super-resolution is performed on a small sample of very narrow objects. This size error has a much stronger contribution to the quantities in (2.109) than the ellipticity error. RCA already reduces this bias in our current set-up, with an improvement of about 24% at all noise levels. This still leads to a RMS on the relative size $\delta(R^2_{\text{PSF}})/R^2_{\text{PSF}}$ that is about 10^4 times too high to match the requirements. Beyond the need to use more stars simultaneously to build the model, which also emerges from every other current shortcoming of our approach, this strong bias will already be greatly reduced in a more realistic *Euclid* scenario, since a broadband PSF is necessarily broader than the monochromatic PSF we are considering in this work, regardless of the target object's SED.

The values of the true PSF ellipticity at each "galaxy position" in our test set is shown in Figure 3.13. The corresponding ellipticity residuals for each PSF model, δe_i^{PSF}, and their distribution across all positions are shown respectively in Figures 3.14 and 3.15, when computed on stars with SNR 35. Noticeable residuals are present for both methods, though they are of lower amplitude in the case of RCA. Figure 3.13 shows a strong asymmetry between the two ellipticity components, with most objects showing mostly horizontally- or vertically-oriented sticks. This indicates the first ellipticity component has both higher values, and much stronger

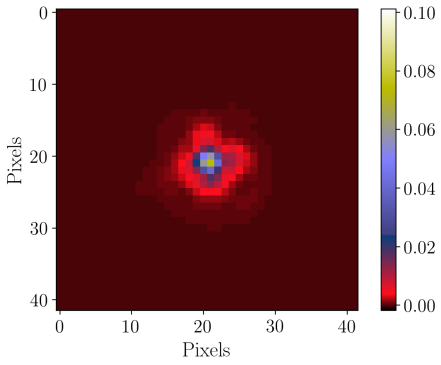
[4]The python script I used to perform this step, and others related to handling SExtractor and PSFEx with simulated PSFs can be found at https://github.com/MorganSchmitz/PySFEx.



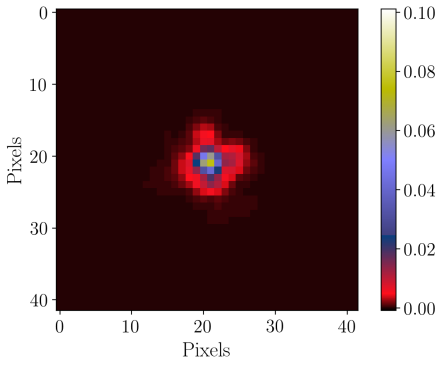
(a) $\hat{H}^{\text{kn}}$



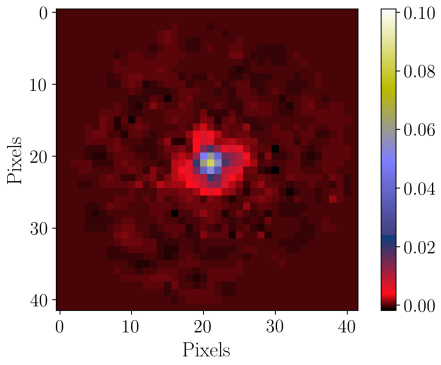
(b) $\hat{H}^{\text{kn}}$



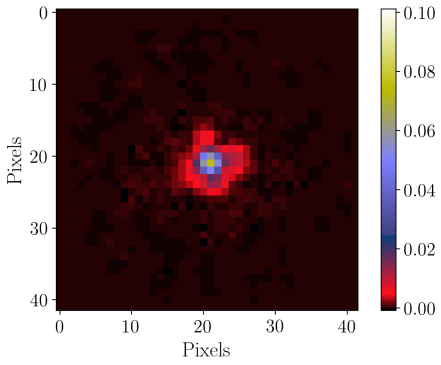
(c) $\hat{H}^{\text{RCA}}$ (SNR 10)



(d) $\hat{H}^{\text{RCA}}$ (SNR 50)



(e) $\hat{H}^{\text{PSFEx}}$ (SNR 10)



(f) $\hat{H}^{\text{PSFEx}}$ (SNR 50)

Figure 3.12: Examples from the three sets of estimated PSFs described in Section 3.3.4. Each stamp is approximately 2.1 arcsecond across. Note that the only difference between (a) and (b) is the color map, matched to be the same as that of $\hat{H}^{\text{RCA}}$ and $\hat{H}^{\text{PSFEx}}$ estimated from stars with SNR 10 and 50, respectively.

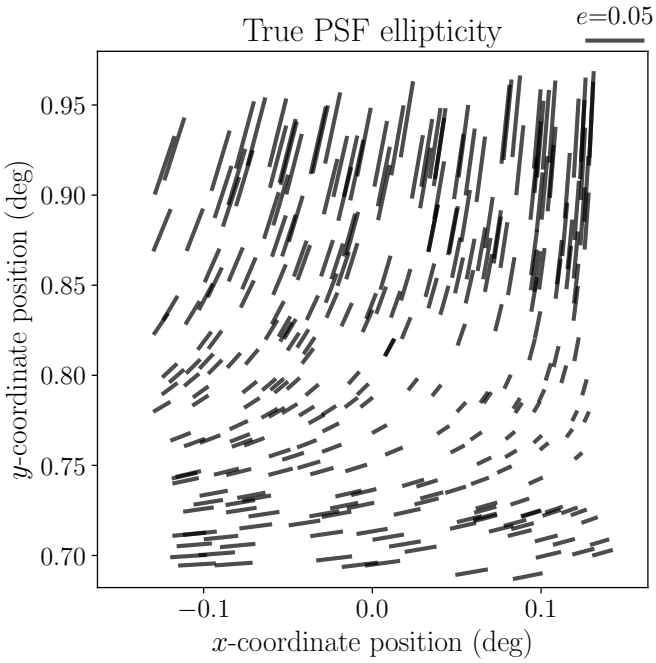


Figure 3.13: True PSF ellipticity as a function of position.

variations across the field than the second (which contributes to diagonal orientations). Residuals in Figures 3.14 and 3.15 are in turn similarly dominated by the first component. This is due to the pixel grid on which the input simulated PSFs were sampled. A simple rotation of the reference frame before sampling would reduce (or invert, through a $\pi/4$-rotation) the difference in amplitudes between the two ellipticity components.

As could already be glimpsed from Figure 3.14, Figure 3.15a shows PSFEx leads to a strong bias in the first ellipticity component that is systematically overestimated. This occurs at all SNRs and indicates that PSFEx, as it is, cannot capture the variations of the Euclid PSF model from undersampled stars.

The RMS error per star SNR level is shown in Figure 3.16. We observe the same overall behavior of both PSF models at all star SNRs, with RCA performing better at e_1 recovery, and worse at capturing the much smaller e_2 variations. As mentioned at the beginning of this section, *Euclid*'s requirements for weak lensing are that the RMS on both PSF ellipticity component should be lower than 5×10^{-5}. As expected, our purely non-parametric approach is far (at a factor of 100-300) from achieving these requirements on its own and with such few stars, though it already yields a significant improvement over PSFEx.

The RMS on the first ellipticity component gets increasingly worse for higher SNR values in the case of PSFEx, which might indicate the presence of spurious effects in the model that get attenuated by higher levels of noise. It might seem puzzling that the error we observe in the case of RCA is lower for a SNR of 35 than it is for one of 50. We observe the same effect when rerunning RCA on several different realizations of noise at those levels. A natural concern would be that this could indicate the quality of our PSF model gets worse with decreasing levels of noise; however, the pixel error between our RCA PSFs and the "known" ones does get smaller, as shown in Figure 3.17. These effects illustrate an important point: when building the PSF model, neither RCA nor PSFEx explicitly aim at matching the observed stars' shapes. It is therefore possible that a "better" model, as defined from the actual functionals both approaches aim at minimizing (in Equations 3.36 and 2.121, respectively), leads to a poorer ellipticity component. This is what we observe in Figure 3.16: the PSF model outputted by RCA, when run on given stars, varies smoothly as a function of their noise level. The overall quality of the model monotonically increases with SNR, as seen in Figure 3.17, eventually converging to the model that would be obtained if there were no noise in the input stars. The ellipticity of the model at any arbitrary position also varies smoothly, but there is no guarantee these variations monotonically tend to the true ellipticity. While the effect we observe here is much smaller (and is, in fact, not identifiable visually when

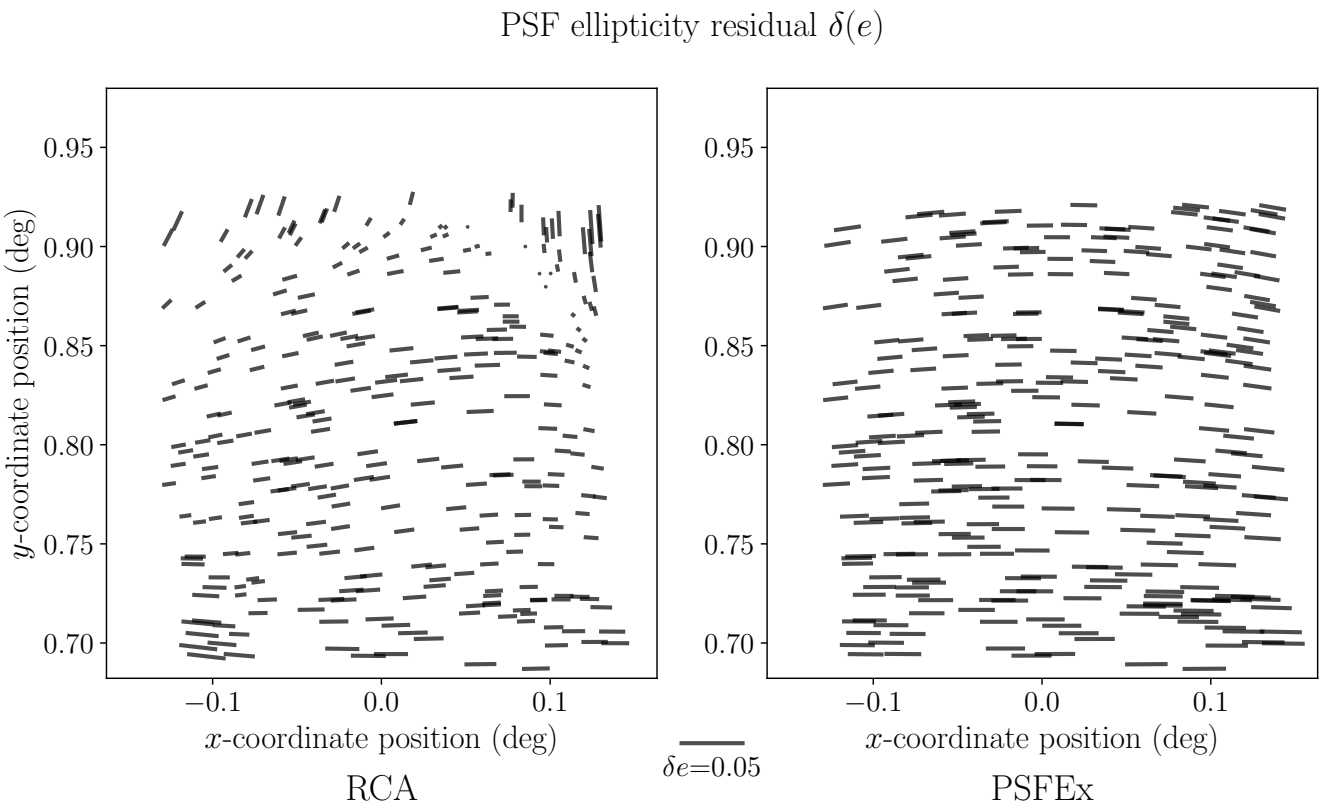
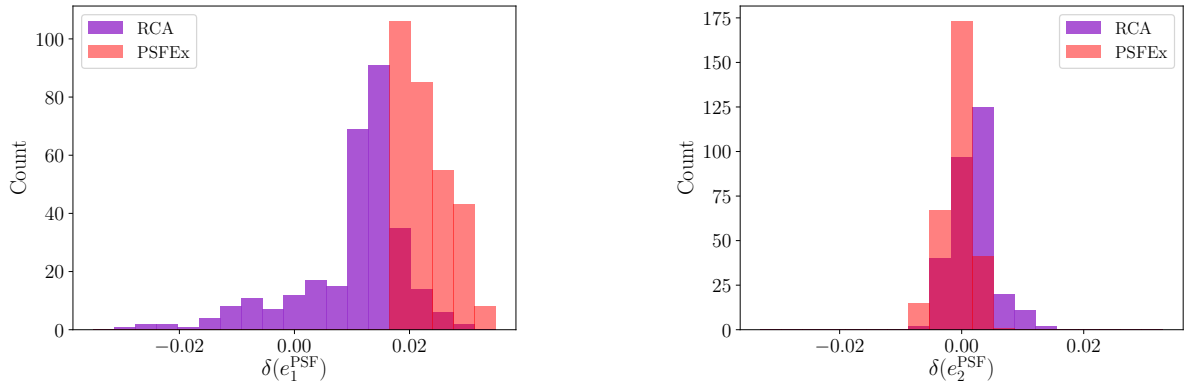


Figure 3.14: PSF ellipticity residuals as a function of position, for both PSF models. Left: RCA; right: `PSFEx`.



(a) 1st ellipticity component

(b) 2nd ellipticity component

Figure 3.15: Distribution of the ellipticity residuals for both PSF models. Measurements were made with star SNR 35.

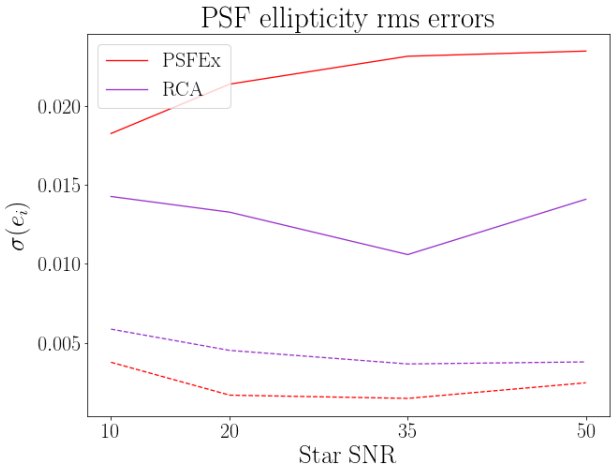


Figure 3.16: RMS error on each PSF ellipticity component for the two models, as a function of input star SNR. Continuous lines are for the first ellipticity component, dashed for the second.

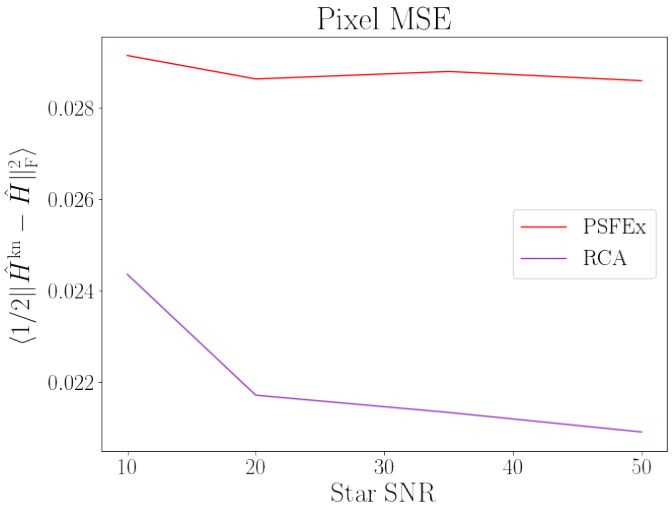


Figure 3.17: Average pixel error as a function of star SNR.

comparing the models obtained at SNRs of 35 and 50), as a crude illustration, consider a PSF with two outer rings: the first one having a dampening effect on the full PSF's first component ellipticity $\delta(e_1^{\mathrm{inner}}) < 0$, and the second leading to an increase $\delta(e_1^{\mathrm{outer}}) = -\delta(e_1^{\mathrm{inner}}) > 0$. For a given number of stars, suppose the best possible error achievable (with no noise) is $\delta(e_1^*) > 0$. Let us assume the quality of the reconstruction of the central part of the PSF is unchanged regardless of input noise levels, and both rings are completely lost to noise at low SNR. As we increase the SNR of input stars, the model would eventually capture the inner ring, while still completely missing the outer one. At this stage, the dampening effect of the first ring would counteract the overestimation of the central part's ellipticity, thus leading to a smaller ellipticity error $0 < \delta(e_1^*) + \delta(e_1^{\mathrm{inner}}) < \delta(e_1^*)$. If the SNR was to keep increasing, however, the outer ring would eventually be captured by the model, increasing once again the overestimation of the first component ellipticity.

3.3.5 Impact on galaxy shape measurement

At the end of Section 2.4.1, we discussed how the use of weighted moments introduced prefactors in the usual Paulin-Henriksson et al. (2008) formalism, leading to equation (2.112). For a Gaussian PSF, these prefactors

are exactly 1. To test whether this remains true in our present case of a *Euclid*-like PSF, in this section, we perform image simulations and galaxy shape measurement using each PSF model. In particular, we will apply both a moments-based shape measurement method, and one based on model-fitting. While each comes with its own method-dependent biases, we would expect the contribution of the PSF modelling errors to be the same in both cases if the assumption that the prefactors in (2.112) vanish held true.

Galaxies and observations

We perform galaxy image simulations using `GalSim` (Rowe et al., 2015). The galaxy parameters are identical to those used in several branches of GREAT3 (Mandelbaum et al., 2014), themselves based on fitting the COSMOS population. This gives us a population of $2\,040\,000$ galaxies that are either drawn from a single Sersic profile, or composed of both a bulge (following a de Vaucouleurs' profile) and a disk (with an exponential profile). We apply 204 different random shear values, each of them to a set of $10\,000$ different galaxies. These sets include the 90-degree rotated counterpart to each galaxy, so as to ensure intrinsic ellipticity truly averages to 0.

The main difference between our image simulations and those used in GREAT3 is, naturally, the PSF used. For our study, we randomly assign one of the 297 Euclid PSFs (at galaxy positions) to each of the galaxies, import them in `GalSim` and perform the convolution with the galaxy profile.

Our observations are then generated by sampling the resulting convolved profile on stamps of 42×42 pixels at half the nominal VIS pixel scale, to match our super-resolved PSFs. Note that in a real-life *Euclid* setting, the observed galaxies would also suffer from undersampling; however, we choose not to take it into account in this work in order to better isolate the effects of imperfect PSF modelling on shape measurement. Similarly, rather than matching the observations' noise level to those we used for the stars, we instead always add white Gaussian noise with $\sigma = 0.01$ (leading to an average SNR of about 50).

Shape measurement

With both the estimated PSFs and observed galaxies described in the previous sections, we can now perform the actual shape measurement step. For a given galaxy of intrinsic ellipticity $e^{\mathrm{int}} = (e_1^{\mathrm{int}}, e_2^{\mathrm{int}})$ and having undergone a shear (g_1, g_2), our shape measurement method yields

$$\hat{e}_i \approx e_i^{\mathrm{int}} + g_i. \tag{3.43}$$

The shear itself can then be obtained by averaging over sets of objects:

$$\hat{g} = \langle \hat{e} \rangle \approx \langle e^{\mathrm{int}} \rangle + \langle g \rangle = g. \tag{3.44}$$

In our case, we know $\langle e^{\mathrm{int}} \rangle$ is exactly 0. As discussed in Equation 2.3.7, numerous shape measurement methods that yield $\hat{e}$ (and thus $\hat{g}$) exist. However, they are known to be imperfect and introduce bias (see Section 2.3.8 for how this bias can be calibrated). Since we are interested in the impact of imperfect PSF models, in order to quantify the amount of error that is introduced by the shape measurement itself, we start by measuring the shape of each observed galaxy using the corresponding known PSF $\hat{H}^{\mathrm{kn}}$. Then, for each of our star noise levels, we repeat the measurement of the same object, both with the RCA-estimated $\hat{H}^{\mathrm{RCA}}$ and the `PSFEx` $\hat{H}^{\mathrm{PSFEx}}$.

As we saw in Equation 2.3.7, there are considerable differences between moments-based and model-fitting methods, especially with regards to how the PSF is taken into account, we perform our experiments with one method of each type. The most well-known moments-based approach is the KSB method, first introduced by Kaiser et al. (1995). In the present work, we use its implementation within the HSM (Hirata and Seljak, 2003; Mandelbaum et al., 2005) library of `GalSim`, where the size of the (circular) Gaussian window function

is matched to that of the observed galaxy. For model-fitting, we use the freely available `im3shape` package[5] described in Zuntz et al. (2013). In the results shown in Section 3.3.6, `im3shape` was ran with most parameters left to default, except for those related to the images stamp size, noise level, ranges for the estimation of the object's centroid and PSF handling (see Section C.1 in the Appendices for a complete list).

A particular consequence of this is that the model chosen for galaxies is a de Vaucouleurs bulge combined with an exponential disk, which in turn is the exact model used for generating some of our observations (though some others are composed of a Single Sersic profile with index $n \notin \{1, 4\}$). However, `im3shape` thus configured assumes the bulge and disk to have the exact same ellipticity, orientation and relative size, which is not necessarily the case for our simulated galaxies. Nonetheless, this means the actual galaxy profiles used by `im3shape` are fairly close to that of the observations, perhaps more so than what could be expected from real data. In other words, our model-fitting experiments may not suffer from so-called *model bias* quite as much as could be expected in a more realistic setting (Voigt and Bridle, 2010). However, our emphasis on the present work is on the effect of PSF modelling errors on galaxy shape measurement, and whether both approaches are similarly affected by them. For a study of the impact of model bias on shape measurement, see Pujol et al. (2017).

In some cases, the KSB implementation we used fails to compute the shapes of certain objects, or returns ellipticity estimates with an absolute value of more than 1. When this occurs with any of our three PSFs, we remove these objects from the analysis. This leads to about $72\,000$ objects being put aside. The exact amount of objects removed per SNR and PSF type are given in Section C.2 of the Appendices. Note the model-fitting approach always provides an estimate of the shape, and thus all $2\,040\,000$ objects are used.

3.3.6 Results

Ellipticity measurements

We first consider the measured shape of galaxies themselves. Regardless of the PSF model and shape measurement method applied, we obtain an estimate of the overall galaxy shape, which includes both its intrinsic ellipticity and the undergone shear as shown in (3.43). As discussed in Equation 2.3.7), both shape measurement approaches suffer from some form of model bias, either because of the window function or the choice of a model, though as mentioned above, in our current set-up, `im3shape` measurements are fairly exempt from model bias. Regardless, these potentially strong biases are due solely to the shape measurement methods themselves, and should be independent from the PSF modelling. Since the impact of the latter is our focus here, we thus study the *relative* ellipticity error of the various combinations of PSF models and shape measurements, that is,

$$\langle(\hat{e}_i^{\mathrm{kn}} - \hat{e}_i^{\mathrm{RCA}})^2\rangle, \langle(\hat{e}_i^{\mathrm{kn}} - \hat{e}_i^{\mathrm{PSFEx}})^2\rangle, \tag{3.45}$$

where the average is taken over all objects. The results are shown in Figure 3.18. Note the overall amplitude of errors is still related to the intrinsic biases of each shape measurement method, which could be alleviated by a proper calibration scheme. However, these results can still be used to inform us about the two PSF models and their impact on galaxy shape estimation.

When using KSB, there is a clear improvement of order $50 - 60\%$ in the shape error when using the proposed approach over `PSFEx`. Similarly to results shown on the PSF models themselves in Figure 3.3.4, this seems to indicate that both models yield significantly different PSFs, and that our RCA-based approach is more successful at reconstructing the true PSF.

Yet, interestingly, the observed difference is much smaller (of order $10 - 20\%$) between the two PSF models when shapes are measured through model-fitting. This would seem to imply that these methods are less sensitive to PSF modelling errors than moments-based methods, which was not an especially expected outcome: for

[5] https://bitbucket.org/joezuntz/im3shape-git

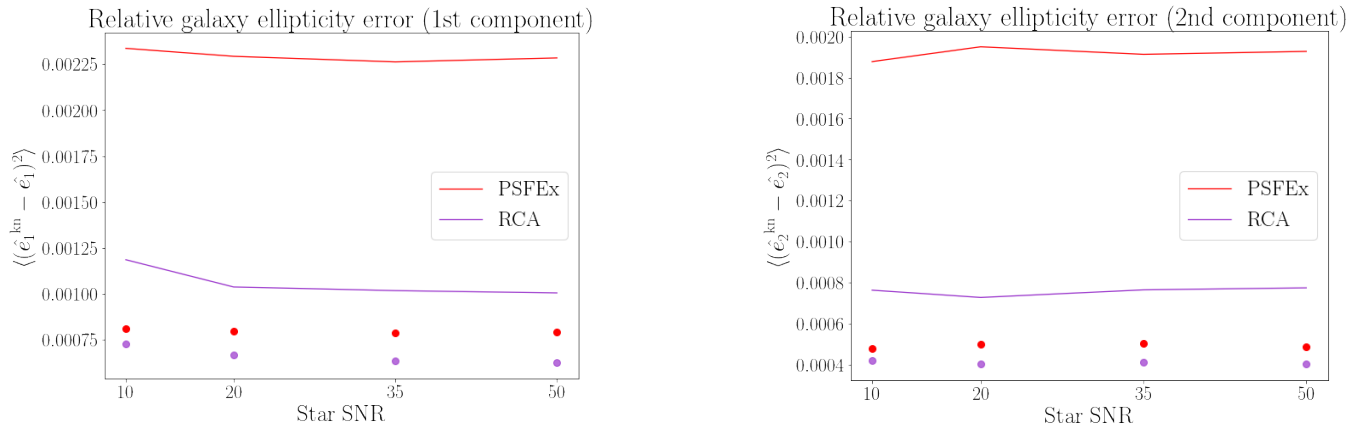


Figure 3.18: Relative ellipticity error on the measured galaxy shapes, for both PSF models, as a function of the star SNR at which they were fitted. The lines indicate errors made when applying KSB, the scattered points errors made when using model-fitting.

instance, Pujol et al. (2017) found no significant difference in sensitivity on various other potential factors when comparing methods of each type. This difference in behaviour when faced with imperfect PSF models could be related to effects due to mixing with higher-order moments discussed in Section 2.4.1. This will be further studied below.

Shear bias

A second way to study the impact of PSF models is to look at the actual inferred shear itself. In our case, since we know the intrinsic galaxy ellipticities average to 0 (and we can correct for it if not, e.g. if some objects were tagged as outliers and removed from the analysis), we only have to average across a set of $10\,000$ objects with the same applied shear to obtain our shear estimator, as shown in (3.44).

A common way to parametrize the bias made on shear measurement is to extend it to first order as in (2.104). We then consider m_i, c_i, the multiplicative and additive shear bias, respectively, for shear component $i \in \{1, 2\}$. As in the previous subsection, we compute the value of those two parameters for each combination of PSF and shape measurement technique. Once again, we emphasize that the goal of the present work is not to compare shape measurement methods per se, but rather how PSF model errors impact them. The focus should thus be on the differences of c and m values between different PSF models, rather than on the actual values themselves.

In the case of KSB, the distributions of the first component of measured and true shears, as well as the best fit linear regression yielding the two bias values, are shown in Figure 3.19 for the known PSF. The same figures for both our models, with star SNR of 35, are shown in Figure 3.20, with the line corresponding to shear bias in the ideal case shown in black for comparison.

PSF modelling induces a stronger shear bias in both cases, with over a factor of 2 gain in multiplicative bias compared to the ideal scenario. The RCA-based PSF leads to an improvement, in both components, of the multiplicative bias compared to `PSFEx`, as seen on Figure 3.21 (where, similarly to the ellipticities in Section 3.3.6, we show the relative biases $\Delta m_i, \Delta c_i$ after subtracting that measured using the known PSF).

Conversely, Figure 3.20 indicates our RCA-based PSFs lead to a higher additive bias than the `PSFEx` ones. This additive bias is present at every star SNR, as shown in Figure 3.22, though it undergoes strong variations. This higher RCA additive bias is especially noticeable on the first of the two shear components, despite both the bias and RMS error on the first component PSF ellipticity being smaller as shown in Figure 3.3.4. A common way to investigate the relationship between PSF and additive bias is to reparametrize (2.104) thus:

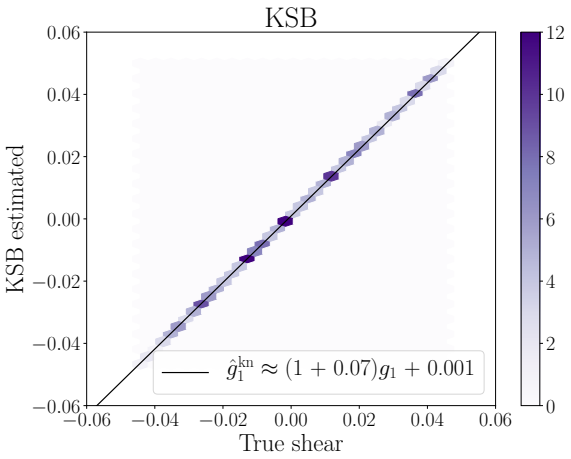
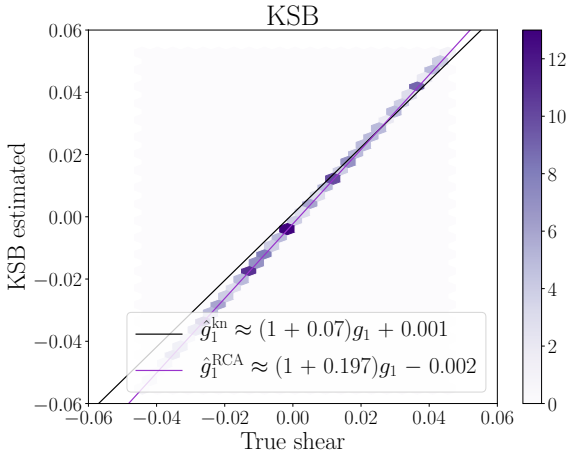
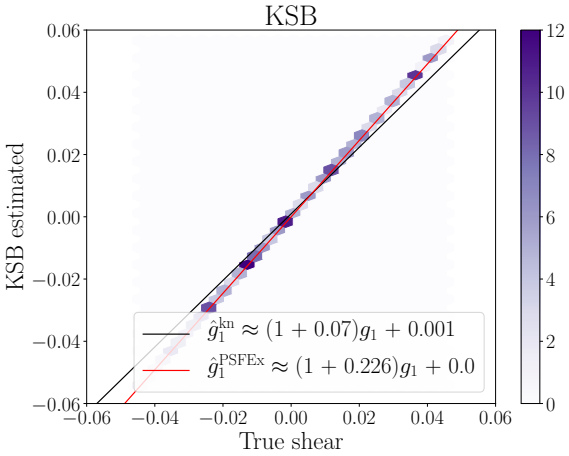


Figure 3.19: 2D density of true and measured shear; the colours correspond to the number of occurrences of measured shear values when using the "known" PSF, each from approximatively $10\,000$ galaxies, for the corresponding input shear. The line shows the best fit linear regression, yielding the bias values. Shapes were measured with KSB.



(a) RCA



(b) `PSFEx`

Figure 3.20: Similar to Figure 3.19, 2D density of true and measured shear, using the PSF models for the latter. The line corresponding to first order shear bias is shown for both the PSF models (in color) and the best case scenario (in black). Shape measurement is performed using KSB.

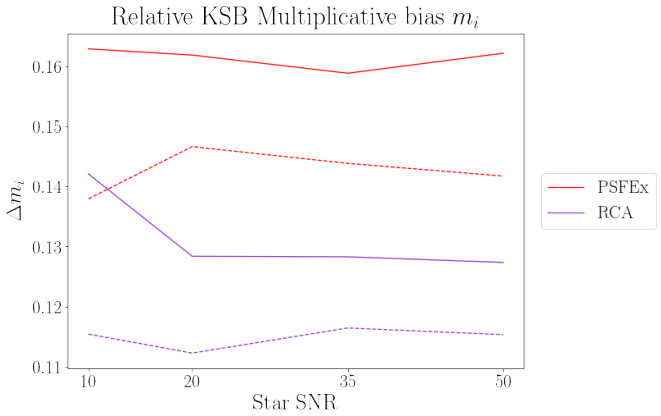


Figure 3.21: KSB-induced multiplicative shear bias m as a function of the SNR of stars on which the PSF models were fitted. Straight lines correspond to the first component, dashed lines to the second.

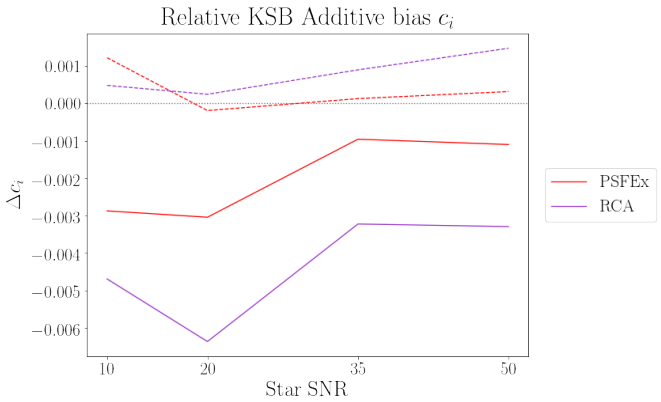


Figure 3.22: Same as Figure 3.21, for additive bias c.

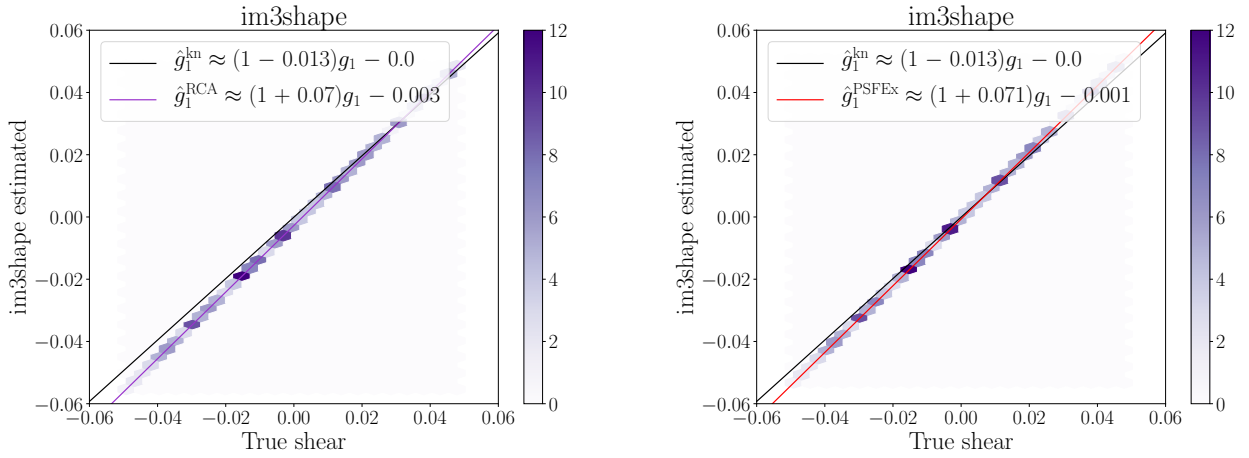
$$\hat{g}_i \approx (1 + m_i)g_i + c_i' + \alpha e_i^{\mathrm{PSF}}, \tag{3.46}$$

where α then quantifies the amount of *PSF leakage*. Note, however, that this quantity can contain both PSF effects that were not fully captured by the shape measurement step, and effects emanating from errors in the PSF model itself. It would therefore not be informative in our present case, where the additive bias appears stronger for the PSF model with the smallest errors despite the same shape measurement being applied in both cases.

A study of the shear biases obtained with our different PSF models when using `im3shape` also seems to indicate the presence of a slight additive bias when using the RCA PSF. This is illustrated in Figure 3.23, which features the same shear 2D densities and linear fit as Figure 3.20, also at star SNR 35, when the shape measurement is performed by model-fitting. In terms of multiplicative bias, the difference between the known and modeled PSFs is much smaller than it was with KSB, and insignificant in between models, which once again seems to indicate a lower sensitivity to PSF modelling errors of model-fitting methods.

Comparison to analytical predictions

These results are already at odds with those predicted by (2.109), since we observe different relative biases introduced by the same PSF model errors depending on the shape measurement used. To better illustrate this discrepancy, for any set of galaxy, true PSF, and PSF model, we compute the expected contribution to



(a) RCA (b) PSFEx

Figure 3.23: Same as Figure 3.20, when shape measurement is performed using model-fitting.

multiplicative bias, $m^{\rm PH}$, and to each component of the additive bias, $(c_i^{\rm PH})_{i\in\{1,2\}}$, defined in (2.110) and (2.111), respectively.

These are shown in Figures 3.24 and 3.25, respectively, for a range of galaxy sizes. Here and for the rest of this section, we use PSFs modelled at star SNR 35. The error bars correspond to the variations across our 297 estimated PSFs. Starting from our full experiment, we separate the pure-Sersic galaxies, split them by size, and recompute the shear biases we observe per galaxy size bin. As previously, we then compute the relative biases, $\Delta m, \Delta c_i$, by removing the bias measured with the "known" PSFs to those of both PSF models. These values are then overplotted for each galaxy size bin in Figures 3.24 and 3.25, and show strong deviations from the analytical predictions.

For instance, we showed in Figure 3.3.4 that the RCA model led to smaller errors in both the first PSF ellipticity component, $\delta e_1^{\rm PSF}$, and its size, $\delta\left(R^2_{\rm PSF}\right)$. It follows that we expect a lower relative (positive) additive bias when using the RCA PSFs, which is the opposite of what we observe with our full experiment. The worse performance in $e_2^{\rm PSF}$ recovery is compensated by RCA's smaller $\delta(R^2_{\rm PSF})$, which, as previously mentioned, largely outweighs the contribution of the ellipticity error term here. This leads to smaller $c_2^{\rm PH}$ values when compared with the prediction for `PSFEx`. However, we see that the biases we observe in practice are strongly dependent on the shape measurement method. With `im3shape`, the c_2 contribution is indeed smaller for RCA, though they were overestimated by the analytical prediction for both PSF models. With KSB, it is higher for RCA than it is for `PSFEx`, and while $c_2^{\rm PH} < 0$ for both PSF models, they lead to a positive contribution when propagated to KSB-measured shapes.

As discussed at the beginning of Section 3.3.5, we know the analytical predictions are exact when the prefactors in (2.112) vanish. In order to test whether these are the reason for the differences we observed, we generate a new set of simulations. The galaxies have the same properties (size, shape, applied shear) as those described above, but are drawn from a 2D Gaussian distribution. Similarly, the PSF applied have identical shape properties as our Euclid PSFs, but are also Gaussian. Lastly, we recreate a set of "RCA" and a set of "`PSFEx`" PSFs, Gaussian as well, but with the same shape errors $\delta e_i^{\rm PSF}, \delta\left(R^2_{\rm PSF}\right)$ as those measured on our actual star-fitted models. While we have access to the true galaxy sizes from the GREAT3 input catalogs, the PSF shapes have to be measured in all three cases. We once again used the HSM library of `GalSim`, which matches the size of the weighting function to the object being measured. We chose the size of our Gaussian PSFs to be the same as that of the matched window, which leads to a constant factor of two in their unweighted $R^2_{\rm PSF}$.

The results are shown in Figures 3.26 and 3.27 for the multiplicative and additive components, respectively,

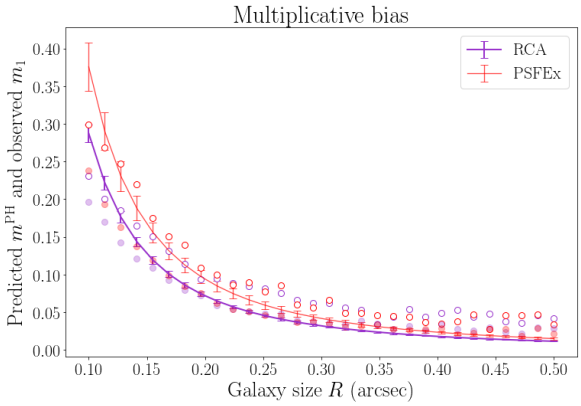
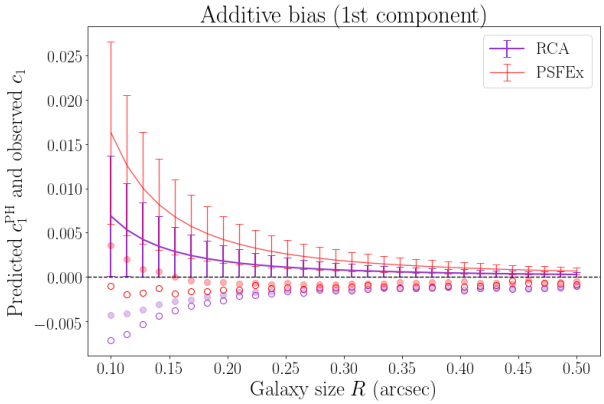
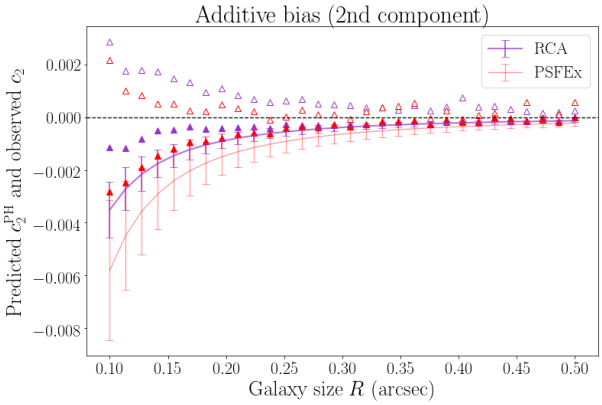


Figure 3.24: Multiplicative bias induced by the PSF models, as predicted from (2.110) (continuous line and error bars) and observed when measuring galaxy shapes with KSB (empty points) or `im3shape` (filled points).



(a) First ellipticity component.



(b) Second ellipticity component.

Figure 3.25: Similar to Figure 3.24, for the additive biases predicted from (2.111) (continuous line and error bars), and those observed with KSB (empty points) and `im3shape` (filled points). Note that in this case, the analytical predictions are different for each ellipticity component because of the left-hand term in (2.111).

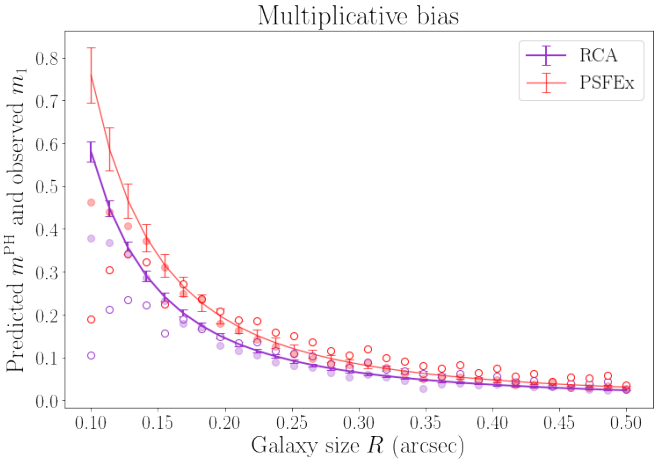
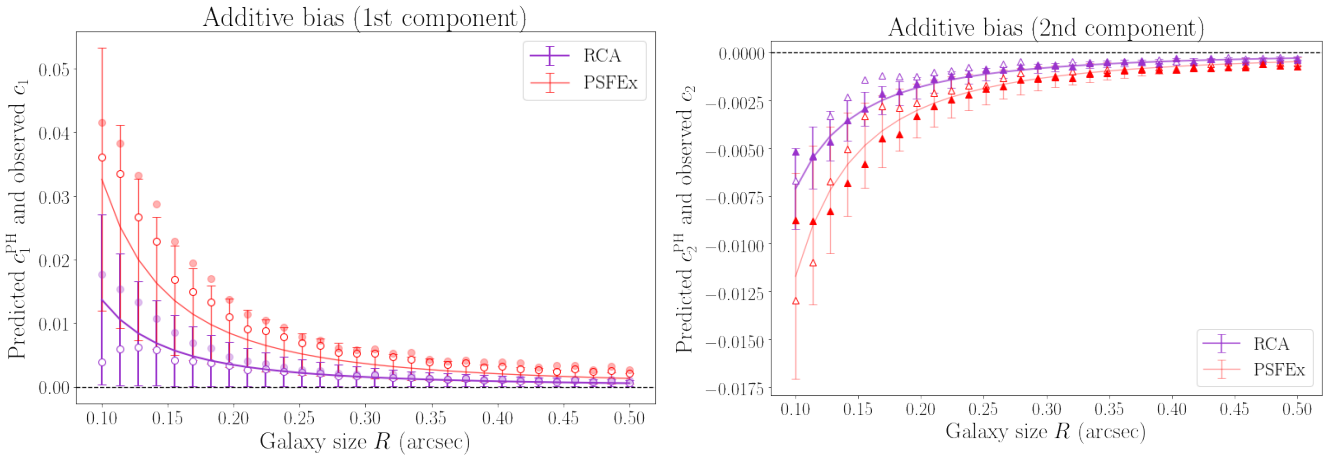


Figure 3.26: Same as Figure 3.24, when PSFs and galaxies are Gaussian.



(a) First ellipticity component.

(b) Second ellipticity component.

Figure 3.27: Same as Figure 3.25, when PSFs and galaxies are Gaussian.

and show good agreement with the predicted values. The first few galaxy size bins lead to smaller measured multiplicative biases, though these are only due to the small number of galaxies at these sizes.

3.3.7 Partial conclusion

In this section, we extended a previously proposed approach for PSF estimation, taking necessary steps toward a fully non-parametric approach applicable in the context of the upcoming *Euclid* survey. A study of the PSF models and their residuals shows our model outperforms the proven and widely used `PSFEx`. This could indicate a better handling of the super-resolution, as hints at potential issues with `PSFEx`'s super-resolution mode were recently observed in HSC (Bosch et al., 2017). Our method is still, however, far from achieving the *Euclid* requirements. As a non-parametric approach, its main limitation lies in the number of available stars, and a natural path of improvement is thus simultaneous use of stars from different exposures, that is, taking into account the temporal variability of the PSF. The other approach is that of a parametric PSF model, discussed in Figure 2.4.3. Ultimately, the combination of both approaches will likely do better than each taken separately, which warrants further study of non-parametric models, how to improve them, and make them capable of handling the specificities of *Euclid*'s PSF.

Despite the improvement in model quality, the use of our approach as the PSF in galaxy shape measurement unexpectedly led to stronger additive shear biases than when using `PSFEx`. Following this observation, as well as other observed discrepancies, this section also showed that in the case of *Euclid*, the way the PSF modelling errors impact shear measurement can be more complicated than previously thought and method-dependent. In particular, the Paulin-Henriksson et al. (2008) formalism no longer holds. Our experiments show this is likely coming from additional terms arising from the necessary addition of a window function to compute quadrupole moments. Similar effects could thus occur for any diffraction-limited telescope.

As mentioned at the beginning of this section, up to now we have considered a single monochromatic PSF. This will no longer be an acceptable assumption in the case of *Euclid* (Eriksen and Hoekstra, 2018), and aside from the need to use a greater number of stars, a necessary improvement to the presented method and focus of the next two chapters will be to take those chromatic variations into account. We first begin with a methodological chapter, where we present the field of numerical OT, and some developments we recently proposed. These will then be used to capture the chromatic variations of the Euclid PSF in Chapter 5.

Chapter 4

Interlude: Optimal Transport

Contents

In this chapter, we introduce the field of numerical OT. In particular, we will define and present approaches to compute the *Wasserstein distance* and the associated *Wasserstein barycenter*. For an overview of the field, see Peyré and Cuturi (2019).

In Sections 3.1.2 and 3.1.6, we have discussed matrix factorization and dictionary learning. As an illustration of the capabilities of OT, this chapter will introduce a new method, Wasserstein Dictionary Learning (WDL), where the linear relation of (3.5) is replaced by the Wasserstein barycenter operator.

As a general method of *representation learning*, it is applicable to many types of data. This is illustrated by our various applications in Section 4.5 and that of Xu et al. (2018), who used (and extended) our method to process hospital admission records. In the context of this thesis, the most relevant is the application of Section 4.5.2. There, we show that OT is particularly suited to capture the chromatic variations of the Euclid PSF. This is done using monochromatic PSFs as training data. In the context of *Euclid* observations, these will not be available, so this section has no direct practical use. However, it will serve as the central idea from which we will build the first polychromatic PSF model in Chapter 5, using the tools we develop in this chapter.

This method, and the rest of this chapter, is the result of work carried out with my collaborators, Matthieu Heitz, Nicolas Bonneel, Fred Ngolè, David Coeurjolly, Marco Cuturi, Gabriel Peyré and Jean-Luc Starck. It was published in Schmitz et al. (2018) (Paper I hereafter).

4.1 Partial Introduction

The idea of dimensionality reduction is as old as data analysis (Pearson, 1901). Dictionary learning (Lee and Seung, 1999), independent component analysis (Hyvärinen et al., 2004), sparse coding (Lee et al., 2007), autoencoders (Hinton and Salakhutdinov, 2006) or most simply PCA are all variations of the idea that each datapoint of a high-dimensional dataset can be efficiently encoded as a low-dimensional vector. Dimensionality reduction typically exploits a sufficient amount of data to produce an encoding map of datapoints into smaller vectors, coupled with a decoding map able to reconstruct an approximation of the original datapoints using such vectors. Algorithms to carry out the encoding and/or the decoding can rely on simple linear combinations of vectors, as is the case with PCA and NMF. They can also be highly nonlinear and employ kernel methods (Schölkopf et al., 1997) or neural networks for that purpose (Hinton and Salakhutdinov, 2006).

In this work, we consider a very specific type of encoding/decoding pair, which relies on OT geometry between probability measures. OT geometry, also known as Wasserstein or earth mover's, defines a distance between two probability measures μ, ν by computing the minimal effort required to morph measure μ into measure ν. The original interpretation of Monge (1781) was that μ would stand for a heap of sand, which should be used to fill in a hole in the ground of the shape of ν. The effort required to move the pile of sand is usually parametrized by a cost function to move one atom of sand from any location x in the support of μ to any location y in the support of ν (see Figure 4.1). Monge then considered the problem of finding the optimal (least costly) way to level the ground by transporting the heap into the hole. That cost defines a geometry between probability measures which has several attractive properties. In this paper we exploit the fact that shapes and, more generally, images can be cast as probability measures, and we propose several tools inherited from OT geometry, such as OT barycenters, to warp and average such images (Solomon et al., 2015). These tools can be exploited further to carry out non-linear inverse problems in a Wasserstein sense (Bonneel et al., 2016), and we propose in this work to extend this approach to carry out nonlinear dictionary learning on images using Wasserstein geometry.

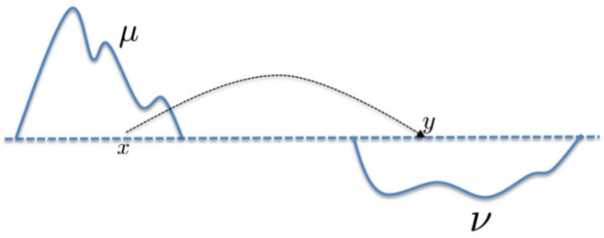


Figure 4.1: Graphical representation of the mass transportation problem. The minimal effort cost to transport one measure into the other defines the OT distance between μ and ν.

4.1.1 Previous works

Linear dictionary learning

Several dimensionality reduction approaches rely on using a predefined orthogonal basis upon which datapoints can be projected. Such basis are usually defined without even looking at data, as is the case for Fourier transforms or wavelet-based dictionaries (Mallat, 1999). Dictionary learning methods instead underline the idea that dictionaries should be *customized* to fit a particular dataset in an optimal way. Suppose that the M datapoints of interest can be stored in a matrix $X = (x_1, \dots, x_M) \in \mathbb{R}^{N \times M}$. The aim of (linear) dictionary learning is to factorize the data matrix X using two matrices: a dictionary, D, whose elements (the atoms) have the same dimension N as those of X, and a list of codes Λ used to relate the two: $X \approx D\Lambda$.

When no constraints on D or Λ are given, and one simply seeks to minimize the Fröbenius norm of the difference of X and $D\Lambda$, the problem amounts to computing the singular value decomposition of X or, equivalently, the diagonalization of the variance matrix of X. In practical situations, one may wish to enforce certain properties of that factorization, which can be done in practice by adding a prior or a constraint on the dictionary D, the codes Λ, or both. For instance, an l_0 or l_1 norm penalty on the codes yields a sparse representation of data (Aharon et al., 2006; Mairal et al., 2010). The sparsity constraint might instead be imposed upon the new components (or atoms), as is the case for sparse PCA (d'Aspremont et al., 2007). Properties other than sparsity might be desired, for example, statistical independence between the components, yielding independent component analysis (ICA, Hyvärinen et al., 2004), or positivity of both the dictionary entries and the codes, yielding NMF (Lee and Seung, 1999). A third possible modification of the dictionary learning problem is to change the *fitting loss* function that measures the discrepancy between a datapoint and its reconstruction. When data lies in the nonnegative orthant, Lee and Seung (1999) have shown, for instance, the interest of considering the Kullback-Leibler (KL) divergence to compute such a loss or, more recently, the Wasserstein distance (Rolet et al., 2016), as detailed later in this section. More advanced fitting losses can also be derived using probabilistic graphical models, such as those considered in the topic modelling literature (Blei and Lafferty, 2009).

Nonlinear dictionary learning

The methods described above are linear in the sense that they attempt to reconstruct each datapoint x_i by a linear combination of a few dictionary elements. Nonlinear dictionary learning techniques involve reconstructing such datapoints using nonlinear operations instead. Autoencoders (Hinton and Salakhutdinov, 2006) propose using neural networks and to use their versatility to encode datapoints into low-dimensional vectors and later decode them with another network to form a reconstruction. The main motivation behind principal geodesic analysis (Fletcher et al., 2004) is to build such nonlinear operations using geometry, namely by replacing linear interpolations with geodesic interpolations. Of particular relevance to our paper is the body of work that relies on Wasserstein geometry to compute geodesic components (Bigot et al., 2017; Wang et al., 2013; Boissard et al., 2015; Seguy and Cuturi, 2015, see Section 4.5.1).

More generally, when data lies on a Riemannian manifold for which Riemannian exponential and logarithmic maps are known, Ho et al. (2013) propose a generalization of both sparse coding and dictionary learning. Nonlinear dictionary learning can also be performed by relying on the "kernel trick", which allows one to learn dictionary atoms that lie in some feature space of higher, or even infinite, dimension (Van Nguyen et al., 2013; Harandi and Salzmann, 2015; Liu et al., 2015). Equiangular kernel dictionary learning, proposed by Quan et al. (2016), further enforces stability of the learned sparse codes. Several problems where data is known to belong to a specific manifold are well studied within this framework, e.g. sparse coding and dictionary learning for Grassmann manifolds (Harandi et al., 2013), or for positive definite matrices (Harandi et al., 2012), and methods to find appropriate kernels and make full use of the associated manifold's geometry have been proposed for the latter (Li et al., 2013). Kernel dictionary learning has also been studied for the (nonlinear) adaptive filtering framework, where Gao et al. (2014) propose an online approach that discards obsolete dictionary elements as new inputs are acquired. These methods rely on the choice of a particular feature space and an associated kernel and achieve nonlinearity through the use of the latter. The learned dictionary atoms then lie in that feature space. Conversely, our proposed approach requires no choice of kernel. Moreover, the training data and the atoms we learn belong to the same probability simplex, which allows for easy representation and interpretation; e.g. our learned atoms can (depending on the chosen fitting loss) capture the extreme states of a transformation undergone by the data. This is opposed to kernel dictionary atoms, which cannot be naturally represented in the same space as datapoints because of their belonging to the chosen high-dimensional feature space.

Computational optimal transport

Optimal transport has seen significant interest from mathematicians in recent decades (Talagrand, 1996; Rachev and Rüschendorf, 1998; Villani, 2003). For many years, that theory was, however, of limited practical use and

mostly restricted to the comparison of small histograms or point clouds, since typical algorithms used to compute them, such as the auction algorithm (Bertsekas, 1988) or the Hungarian algorithm (Kuhn, 1955), were intractable beyond a few hundred bins or points. Recent approaches (Shirdhonkar and Jacobs, 2008; Rabin et al., 2011) have ignited interest for fast yet faithful approximations of OT distances. Of particular interest to this work is the entropic regularization scheme proposed by Cuturi (2013), which finds its roots in the gravity model used in transportation theory (Erlander and Stewart, 1990). This regularization can also be tied to the relation between OT and Schrödinger's problem (Schrödinger, 1931) (as explored by Léonard, 2014). Whereas the original OT problem is a *linear* problem, regularizing it with an entropic regularization term results in a strictly convex problem with a unique solution which can be solved with Sinkhorn's fixed-point algorithm (Sinkhorn, 1967), also known as block coordinate ascent in the dual regularized OT problem. That iterative fixed-point scheme yields a numerical approach relying only on elementwise operations on vectors and matrix-vector products. The latter can in many cases be replaced by a separable convolution operator (Solomon et al., 2015), forgoing the need to manipulate a full cost matrix of prohibitive dimensions in some use cases of interest (e.g. when input measures are large images).

Wasserstein barycenters

Agueh and Carlier (2011) introduced the idea of a Wasserstein barycenter in the space of probability measures, namely Fréchet means (Fréchet, 1948) computed with the Wasserstein metric. Such barycenters are the basic building block of our proposal of a nonlinear dictionary learning scheme with Wasserstein geometry. Agueh and Carlier studied several properties of Wasserstein barycenters and showed very importantly that their exact computation for empirical measures involves solving a multimarginal optimal transport problem, namely a linear program with the size growing exponentially with the size of the support of the considered measures.

Since that work, several algorithms have been proposed to efficiently compute these barycenters (Rabin et al., 2011; Carlier et al., 2015; Solomon et al., 2014; Bonneel et al., 2015; Ye et al., 2017). The computation of such barycenters using regularized distances (Cuturi and Doucet, 2014) is of particular interest to this work. Cuturi and Peyré (2016) use entropic regularization and duality to cast a wide range of problems involving Wasserstein distances (including the computation of Wasserstein barycenters) as simple convex programs with closed form derivatives. They also illustrate the fact that the smoothness introduced by the addition of the entropic penalty can be *desirable*, beyond its computational gains, in the case of the Wasserstein barycenter problem. Indeed, when the discretization grid is small, its true optimum can be highly unstable, which is counteracted by the smoothing introduced by the entropy (Cuturi and Peyré, 2016, Section 3.4). The idea of performing iterative Bregman projections to compute approximate Wasserstein distances can be extended to the barycenter problem, allowing its direct computation using a generalized form of the Sinkhorn algorithm (Benamou et al., 2015). Chizat et al. (2018) recently proposed a unifying framework for solving unbalanced optimal transport problems, including computing a generalization of the Wasserstein barycenter.

Wasserstein barycentric coordinates

An approach to solving the inverse problem associated with Wasserstein barycenters was recently proposed (Bonneel et al., 2016): Given a database of S histograms, a vector of S weights can be associated to any new input histogram, such that the barycenter of that database with those weights approximates as closely as possible the input histogram. These weights are obtained by automatic differentiation (with respect to the weights) of the generalized Sinkhorn algorithm that outputs the approximate Wasserstein barycenter. This step can be seen as an analogy of, given a dictionary D, finding the best vector of weights Λ that can help reconstruct a new datapoint using the atoms in the dictionary. That work can be seen as a precursor for our proposal, whose aim is to learn both weights *and* dictionary atoms.

Applications to image processing

OT was introduced into the computer graphics community by Rubner et al. (2000) to retrieve images from their color distribution, by considering images as distributions of pixels within a 3D color space. Color processing has remained a recurring application of OT, for instance to color grade an input image using a photograph of a desired color style (Pitié et al., 2005), or using a database of photographs (Bonneel et al., 2016), or to harmonize multiple images' colors (Bonneel et al., 2015). Another approach considers grayscale images as 2D histograms. OT then allows one to find a transport-based warping between images (Haker et al., 2004; Mérigot, 2011). Further image processing applications are reviewed in the habilitation dissertation of Papadakis (2015).

Wasserstein loss and fidelity

Several recent papers have investigated the use of OT distances as fitting losses that have desirable properties that KL or Euclidean distances cannot offer. We have already mentioned generalizations of PCA to the set of probability measures via the use of OT distances (Bigot et al., 2017; Seguy and Cuturi, 2015). Sandler and Lindenbaum (2009) first considered the NMF problem with a Wasserstein loss. Their computational approach was, however, of limited practical use. More scalable algorithms for Wasserstein NMF and (linear) dictionary learning were subsequently proposed (Rolet et al., 2016). The Wasserstein distance was also used as a loss function with desirable robustness properties to address multilabel supervised learning problems (Frogner et al., 2015).

Using the Wasserstein distance to quantify the fit between data (an empirical measure) and a parametric family of densities, or a generative model defined using a parameterized push-forward map of a base measure, has also received ample attention in the recent literature. Theoretical properties of such estimators were established by Bassetti et al. (2006) and Bassetti and Regazzini (2006), and additional results by Bernton et al. (2017). Entropic smoothing has facilitated the translation of these ideas into practical algorithms, as illustrated in the work by Montavon et al. (2016), who proposed to estimate the parameters of restricted Boltzmann machines using the Wasserstein distance instead of the KL divergence. Purely generative models, namely, degenerate probability measures defined as the push-forward of a measure supported on a low-dimensional space into a high-dimensional space using a parameterized function, have also been fitted to observations using a Wasserstein loss (Bernton et al., 2017), allowing for density fitting without having to choose summary statistics (as is often the case with usual methods). The Wasserstein distance has also been used in the context of generative adversarial networks (GANs, Arjovsky et al., 2017). In that work, the authors use a proxy to approximate the 1-Wasserstein distance. Instead of computing the 1-Wasserstein distance using 1-Lipschitz functions, a classic result from Kantorovich's dual formulation of OT (see Theorem 1.14 in Villani, 2003), the authors restrict that set to multilayer networks with rectified linear units and boundedness constraints on weights, which allows them to enforce some form of Lipschitzness of their networks. Unlike the entropic smoothing used in this paper, that approximation requires solving a nonconvex problem whose optimum, even if attained, would be arbitrarily far from the true Wassertein distance. More recently, Genevay et al. (2018) introduced a general scheme for using OT distances as the loss in generative models, which relies on both the entropic penalty and automatic differentiation of the Sinkhorn algorithm. Our work shares some similarities with that paper, since we also propose automatically differentiating the Sinkhorn iterations used in Wasserstein barycenter computations.

4.1.2 Contributions

In this paper, we introduce a new method for carrying out nonlinear dictionary learning for probability histograms using OT geometry. Nonlinearity comes from the fact that we replace the usual linear combination of dictionary atoms by Wasserstein barycenters. Our goal is to reconstruct datapoints using the closest (according to any arbitrary fitting loss on the simplex, not necessarily the Wasserstein distance) Wasserstein barycenter to that point using the dictionary atoms. Namely, instead of considering linear reconstructions for $X \approx D\Lambda$, our aim is to approximate columns of $X \approx \mathbf{P}(D, \Lambda)$ using the $\mathbf{P}$ operator which maps atoms D with lists of weights Λ to their respective barycenters.

Similar to many traditional dictionary learning approaches, this is achieved by finding local minima of a nonconvex energy function. To do so, we propose using automatic differentiation of the iterative scheme used to compute Wasserstein barycenters. We can thus obtain gradients with respect to both the dictionary atoms and the weights that can then be used within one's solver of choice (in this work, we chose to use an off-the-shelf quasi-Newton approach and perform both dictionary and code updates simultaneously).

Our nonlinear dictionary learning approach makes full use of the Wasserstein space's properties, as illustrated in Figure 4.2: two atoms are learned from a dataset made up of five discretized Gaussian distributions in 1D, each slightly translated on the grid. Despite the simplicity of the transformation (translation), linear generative models fail to capture the changes of the geometrical space, as opposed to our OT approach. Moreover, the atoms we learn are also discrete measures, unlike the PCA and NMF components.

We also offer several variants and improvements to our method:

- Arbitrarily sharp reconstructions can be reached by performing the barycenter computation in the log-domain;
- We offer a general method to make use of the separability of the kernel involved and greatly alleviate the computational cost of this log-domain stabilization;
- Our representation is learned from the differentiation of an iterative, Sinkhorn-like algorithm, whose convergence can be accelerated by using information from previous Sinkhorn loops at each initialization (warm start), or adding a momentum term to the Sinkhorn iterations (heavyball);
- We expand our method to the unbalanced transport framework.

Part of this work was previously presented as a conference proceedings (Schmitz et al., 2017), featuring an initial version of our method, without any of the above improvements and variants, and in the case where we were only interested in learning two different atoms.

Additional background on OT is given in Section 4.2. The method itself and an efficient implementation are presented in Section 4.3. We highlight other extensions in Section 4.4. We showcase its use in several image processing applications in Section 4.5.

4.1.3 Notations used in this chapter

We denote Σ_d the simplex of $\mathbb{R}^d$, that is,

$$\Sigma_d := \left\{ u \in \mathbb{R}_+^d, \sum_{i=1}^{d} u_i = 1 \right\}. \tag{4.1}$$

For any positive matrix T, we define its negative entropy as

$$H(T) := \sum_{i,j} T_{ij} \log(T_{ij} - 1). \tag{4.2}$$

$\odot$ denotes the Hadamard product between matrices or vectors. Throughout this chapter, when applied to matrices, $\prod, \div$, and $\exp$ notations refer to elementwise operators. The scalar product between two matrices denotes the usual inner product, that is,

$$\langle A, B \rangle := \mathrm{Tr}(A^\top B) = \sum_{i,j} A_{ij} B_{ij}, \tag{4.3}$$

where $A^\top$ is the transpose of A. For $(p, q) \in \Sigma_N^2$, we denote their set of couplings as

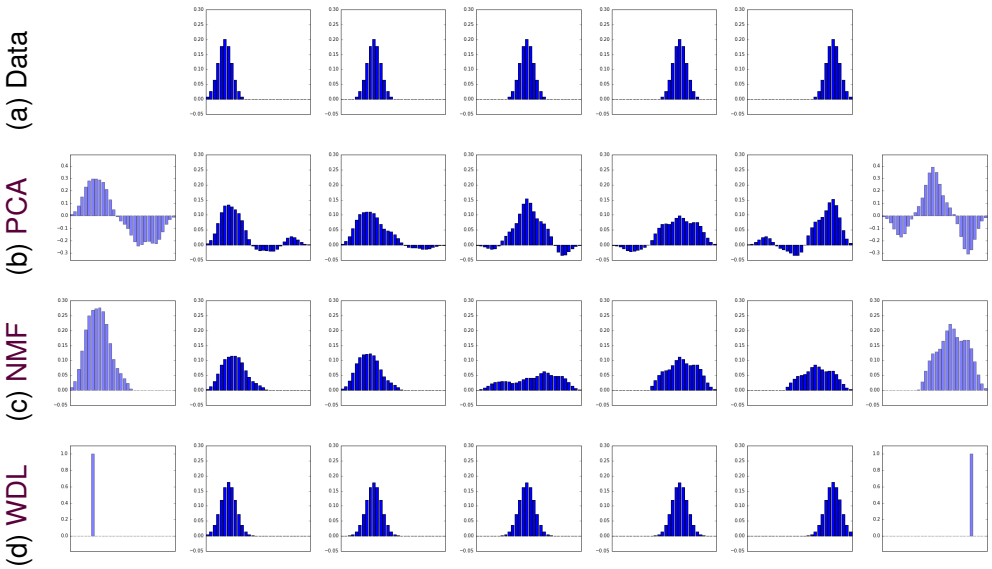


Figure 4.2: Top row: data points. Bottom three rows: On the far sides, in purple, are the two atoms learned by PCA, NMF and our method (WDL), respectively. In between the two atoms are the reconstructions of the five datapoints for each method. The latter two were relaunched a few times with randomized initializations and the best local minimum was kept. As discussed in Section 4.2, the addition of an entropy penalty to the usual OT program causes a blur in the reconstructions. When the parameter associated with the entropy is high, our method yields atoms that are sharper than the dataset on which it was trained, as is observed here where the atoms are Dirac despite the dataset consisting of discretized Gaussians. See Section 4.4.1 for a method to reach arbitrarily low values of the entropy parameter and counteract the blurring effect.

$$\Pi(p,q) := \left\{T \in \mathbb{R}_+^{N\times N}, T\mathbf{1}_N = p, T^\top \mathbf{1}_N = q\right\}, \tag{4.4}$$

where $\mathbf{1}_N = (1,\dots,1)^\top \in \mathbb{R}^N$. Δ denotes the diag operator, such that if $u \in \mathbb{R}^N$, then

$$\Delta(u) := \begin{pmatrix} u_1 & & \\ & \ddots & \\ & & u_N \end{pmatrix} \in \mathbb{R}^{N\times N}. \tag{4.5}$$

For two vectors u, v, $\mathrm{KL}(.|.)$ is their KL divergence, defined here as

$$\mathrm{KL}(u|v) = \sum_i u_i \log\left(\frac{u_i}{v_i}\right) - u_i + v_i. \tag{4.6}$$

For a concatenated family of vectors $t = \left[t_1^\top, \dots, t_S^\top\right]^\top \in \mathbb{R}^{NS}$, we write the ith element of t_s as $[t_s]_i$. We denote the rows of matrix M as $M_{i.}$ and its columns as $M_{.j}$. I_N and $\mathbf{0}_{N\times N}$ are the $N \times N$ identity and zero matrices, respectively.

4.2 Optimal transport

4.2.1 OT distances

In the present work, we restrict ourselves to the discrete setting, i.e. our measures of interest will be histograms, discretized on a fixed grid of size N (Eulerian discretization), and represented as vectors in Σ_N. In this case, the cost function is represented as a cost matrix $C \in \mathbb{R}^{N\times N}$, containing the costs of transportation between any two locations in the discretization grid. The OT distance between two histograms $(p,q) \in \Sigma_N^2$ is the solution to the discretized Monge–Kantorovich problem:

$$W(p,q) := \min_{T\in\Pi(p,q)} \langle T, C\rangle. \tag{4.7}$$

As defined in (4.4), $\Pi(p,q)$ is the set of admissible couplings between p and q, that is, the set of matrices with rows summing to p and columns to q. A solution, $T^* \in \mathbb{R}^{N\times N}$, is an optimal transport plan.

Villani's books give extended theoretical overviews of OT (Villani, 2003, 2008) and, in particular, several properties of such distances. The particular case where the cost matrix is derived from a metric on the chosen discretization grid yields the so-called Wasserstein distance (sometimes called the earth mover's distance). For example, if $C_{ij} = \|x_i - x_j\|_2^2$ (where x_i, x_j are the positions on the grid), the above formulation yields the squared 2-Wasserstein distance, the square-root of which is indeed a distance in the mathematical sense. Despite its intuitive formulation, the computation of Wasserstein distances grows supercubicly in N, making them impractical as dimensions reach the order of one thousand grid points. This issue has motivated the recent introduction of several approximations that can be obtained at a lower computational cost (see Section 4.1.1). Among such approximations, the entropic regularization of OT distances (Cuturi, 2013) relies on the addition of a penalty term as follows:

$$W_\gamma(p,q) := \min_{T\in\Pi(p,q)} \langle T, C\rangle + \gamma H(T), \tag{4.8}$$

where $\gamma > 0$ is a hyperparameter. As $\gamma \to 0$, W_γ converges to the original Wasserstein distance, while higher values of γ promote more diffuse transport matrices. The addition of a negentropic penalty makes the problem

γ-strongly convex; first-order conditions show that the problem can be analyzed as a matrix-scaling problem which can be solved using Sinkhorn's algorithm (Sinkhorn, 1967), also known as the iterative proportional fitting procedure (Deming and Stephan, 1940). The Sinkhorn algorithm can be interpreted in several ways: for instance, it can be thought of as an alternate projection scheme under a KL divergence for couplings (Benamou et al., 2015) or as a block-coordinate ascent on a dual problem (Cuturi and Doucet, 2014). The Sinkhorn algorithm consists in using the following iterations for $l \geq 1$, starting with $b^{(0)} = \mathbf{1}_N$:

$$\begin{aligned} a^{(l)} &= \frac{q}{K^\top b^{(l-1)}}, \\ b^{(l)} &= \frac{p}{K a^{(l)}}, \end{aligned} \tag{4.9}$$

where $K := \exp(\frac{-C}{\gamma})$ is the elementwise exponential of the negative of the rescaled cost matrix. Note that when γ gets close to 0, some values of K become negligible, and values within the scaling vectors, $a^{(l)}$ and $b^{(l)}$, can also result in numerical instability in practice (we will study workarounds for that issue in Section 4.4.1). Application of the matrix K can often be closely approximated by a separable operation (Solomon et al., 2015, see Equation 4.4.1 for separability even in the log-domain). In the case where the histograms are defined on a uniform grid and the cost matrix is the squared Euclidean distance, the convolution kernel is simply Gaussian with standard deviation $\sqrt{\gamma/2}$. The two vectors $a^{(l)}, b^{(l)}$ converge linearly towards the optimal scalings (Franklin and Lorenz, 1989) corresponding to the optimal solution of (4.8). Notice finally that the Sinkhorn algorithm at each iteration $l \geq 1$ results in an approximate optimal transport matrix $T^{(l)} = \Delta(b^{(l)}) K \Delta(a^{(l)})$.

4.2.2 Wasserstein barycenter

Analogous to the usual Euclidean barycenter, the Wasserstein barycenter of a family of measures is defined as the minimizer of the (weighted) sum of squared Wasserstein distances from the variable to each of the measures in that family (Agueh and Carlier, 2011). For measures with the same discrete support, we define, using entropic regularization, the barycenter of histograms $(d_1, \ldots, d_S) \in (\Sigma_N)^S$ with barycentric weights $\lambda = (\lambda_1, \ldots, \lambda_S) \in \Sigma_S$ as

$$P(D, \lambda) := \operatorname*{argmin}_{u \in \Sigma_N} \sum_{s=1}^{S} \lambda_s W_\gamma(d_s, u), \tag{4.10}$$

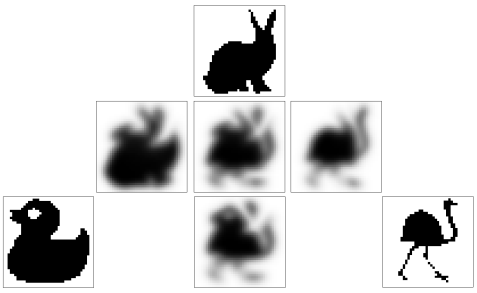
where $D := (d_1^\top, \ldots, d_S^\top)^\top \in \mathbb{R}^{NS}$. The addition of the entropy term ensures strict convexity and thus that the Wasserstein barycenter is uniquely defined. It also yields a simple and efficient iterative scheme to compute approximate Wasserstein barycenters, which can be seen as a particular case of the unbalanced OT setting (Chizat et al., 2018). This scheme, a generalization of the Sinkhorn algorithm, once again relies on two scaling vectors:

$$a_s^{(l)} = \frac{d_s}{K b_s^{(l-1)}}, \tag{4.11}$$

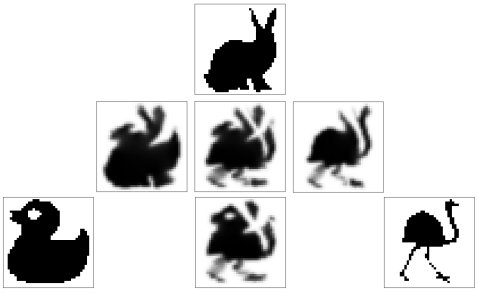
$$P^{(l)}(D, \lambda) = \prod_{s=1}^{S} \left(K^\top a_s^{(l)} \right)^{\lambda_s}, \tag{4.12}$$

$$b_s^{(l)} = \frac{P^{(l)}(D, \lambda)}{K^\top a_s^{(l)}}, \tag{4.13}$$

where, as before, $K = \exp(\frac{-C}{\gamma})$. In this case, however, the scaling vectors are of size NS, such that $a^{(l)} = \left(a_1^{(l)\top}, \ldots, a_S^{(l)\top}\right)^\top$, $b^{(l)} = \left(b_1^{(l)\top}, \ldots, b_S^{(l)\top}\right)^\top$ and $b^{(0)} = \mathbf{1}_{NS}$. Note that one can perform both scaling vector updates



(a) Wasserstein simplex; $\gamma = 8$



(b) Wasserstein simplex; $\gamma = 1$

Figure 4.3: Wasserstein simplices: barycenters of the three images in the corners with varying barycentric weights. Middle row: $\lambda = \left[\frac{1}{2}, \frac{1}{2}, 0\right], \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right], \left[0, \frac{1}{2}, \frac{1}{2}\right]$. Bottom row, center: $\left[\frac{1}{2}, 0, \frac{1}{2}\right]$.

at once (and avoid storing both) by plugging one of (4.11), (4.13) into the other. An illustration of the Wasserstein barycenter, as well as the impact of the γ parameter, is given in Figure 4.3.

4.3 Wasserstein dictionary learning

4.3.1 Overview

Given data $X \in \mathbb{R}^{N \times M}$ in the form of histograms, i.e., each column $x_i \in \Sigma_N$ (for instance a list of M images with normalized pixel intensities), and the desired number of atoms S, we aim to learn a dictionary D made up of histograms $(d_1, \ldots, d_S) \in (\Sigma_N)^S$ and a list of barycentric weights $\Lambda = (\lambda_1, \ldots, \lambda_M) \in (\Sigma_S)^M$ so that for each input, $P(D, \lambda_i)$ is the best approximation of x_i according to some criterion $\mathcal{L}$ (see Table 4.1 for examples). Namely, our representation is obtained by solving the problem

$$\min_{D,\Lambda} \mathcal{E}(D, \Lambda) := \sum_{i=1}^{M} \mathcal{L}\left(P(D, \lambda_i), x_i\right). \tag{4.14}$$

Note the similarity between the usual dictionary learning formulation (see Section 4.1.1) and the one above. In our case, however, the reconstruction of the original data happens via the nonlinear Wasserstein barycenter operator, $\mathbf{P}(D, \Lambda) = (P(D, \lambda_i))_i$, instead of the (linear) matrix product $D\Lambda$.

Differentiation of (4.14) relies in part on the computation of the Wasserstein barycenter operator's Jacobians with regard to either the barycentric weights or the atoms. While it is possible to obtain their analytical formulae, for example by using the fact that Sinkhorn updates (4.12–4.13) become fixed-point equations when convergence is reached, they rely on solving a linear system of prohibitive dimensionality for our settings of interest where N is typically large (Bonneel et al., 2016, derived the expression with regard to barycentric weights and discussed the issue in their Section 4.1). Moreover, in practice, the true Wasserstein barycenters with entropic penalty $P(D, \lambda_i)$ are unknown and approximated by sufficient Sinkhorn iterations (4.12–4.13). As is now common practice in some machine learning methods (a typical example being backward propagation for neural nets), and following recent works (Bonneel et al., 2016), we instead take an approach in the vein of automatic differentiation (Griewank and Walther, 2008). That is, we recursively differentiate the iterative scheme yielding our algorithm instead of the analytical formula of our Wasserstein barycenter. In our case, this is the generalization of the Sinkhorn algorithm for barycenters. Instead of (4.14), we thus aim to minimize

$$\min_{D,\Lambda} \mathcal{E}_L(D, \Lambda) := \sum_{i=1}^{M} \mathcal{L}\left(P^{(L)}(D, \lambda_i), x_i\right), \tag{4.15}$$

where $P^{(L)}$ is the approximate barycenter after L iterations, defined as in (4.12). Even when using an entropy penalty term, we have no guarantee on the convexity of the above problem, whether jointly in D and Λ or for each separately, contrary to the case of OT distance computation in (4.8). We thus aim to reach a local minimum of energy landscape $\mathcal{E}_L$ by computing its gradients and applying a descent method. By additivity of $\mathcal{E}_L$ and without loss of generality, we will focus on the derivations of such gradients for a single datapoint $x \in \Sigma_N$ (in which case Λ only comprises one list of weights $\lambda \in \Sigma_S$). Differentiation of (4.15) yields

$$\nabla_D \mathcal{E}_L(D, \Lambda) = \left[\partial_D P^{(L)}(D, \lambda)\right]^\top \nabla \mathcal{L}(P^{(L)}(D, \lambda), x), \tag{4.16}$$

$$\nabla_\lambda \mathcal{E}_L(D, \Lambda) = \left[\partial_\lambda P^{(L)}(D, \lambda)\right]^\top \nabla \mathcal{L}(P^{(L)}(D, \lambda), x). \tag{4.17}$$

The right-hand term in both cases is the gradient of the loss which is typically readily computable (see Table 4.1) and depends on the choice of fitting loss. The left-hand terms are the Jacobians of the Wasserstein barycenter operator with regard to either the weights or the dictionary. These can be obtained either by performing the analytical differentiation of the $P^{(l)}$ operator, as is done in Section 4.3.2 (and Appendix D), or by using an automatic differentiation library such as Theano (Theano Development Team, 2016). The latter approach ensures that the complexity of the backward loop is the same as that of the forward, but it can lead to memory problems due to the storing of all objects being part of the gradient computation graph (as can be the case, for instance, when performing the forward Sinkhorn loop in the log-domain as in Section 4.4.1; for this specific case, an alternative is given in Equation 4.4.1). The resulting numerical scheme relies only on elementwise operations and on the application of the matrix K (or its transpose) which often amounts to applying a separable convolution (Solomon et al., 2015, see Equation 4.4.1). The resulting algorithm is given in Algorithm 4.1. At first, a "forward" loop is performed, which amounts to the exact same operations as those used to compute the approximate Wasserstein barycenter using updates (4.12–4.13) (the barycenter for current weights and atoms is thus computed as a by-product). Two additional vectors of size SNL are stored and then used in the recursive backward differentiation loops that compute the gradients with regard to the dictionary and the weights.

Name	$\mathcal{L}(p, q)$	$\nabla\mathcal{L}$
Total variation	$\|p - q\|_1$	$\operatorname{sign}(p - q)$
Quadratic	$\|p - q\|_2^2$	$2(p - q)$
Kullback-Leibler	$\mathrm{KL}(p\|q)$	$\log(p/q) - 1$
Wasserstein[1]	$W_\gamma^{(L)}(p, q)$	$\gamma \log(a^{(L)})$

Table 4.1: Examples of similarity criteria and their gradient in p. See Figure 4.14 for the atoms yielded by our method for these various fitting losses.

Using the above scheme to compute gradients, or its automatically computed counterpart from an automatic differentiation tool, one can find a local minimum of the energy landscape (4.15), and thus the eventual representation Λ and dictionary D, by applying any appropriate optimization method under the constraints that both the atoms and the weights belong to their respective simplices Σ_N, Σ_S.

For the applications shown in Section 4.5, we chose to enforce these constraints through the following change of variables

$$\forall i, d_i := F_N(\alpha_i) := \frac{e^{\alpha_i}}{\sum_{j=1}^N e^{[\alpha_i]_j}}, \qquad \lambda := F_S(\beta) := \frac{e^{\beta}}{\sum_{j=1}^S e^{\beta_j}}. \tag{4.18}$$

The energy to minimize (with regard to α, β) then reads as

[1] In this case, the loss is computed iteratively as explained in Section 4.2.1, and $a^{(L)}$ in the gradient's expression is obtained after L iterations as in (4.9).

Inputs: Data $x \in \Sigma_N$, atoms $d_1, \dots, d_S \in \Sigma_N$, current weights $\lambda \in \Sigma_S$

comment: Sinkhorn loop

$\forall s, b_s^{(0)} := \mathbf{1}_N$

for $l = 1$ to L step 1 do

$\quad \forall s, \varphi_s^{(l)} := K^\top \frac{d_s}{K b_s^{(l-1)}}$

$\quad p := \prod_s \left(\varphi_s^{(l)}\right)^{\lambda_s}$

$\quad \forall s, b_s^{(l)} := \frac{p}{\varphi_s^{(l)}}$

od

comment: Backward loop - weights

$w := \mathbf{0}_S$

$r := \mathbf{0}_{S \times N}$

$g := \nabla \mathcal{L}(p, x) \odot p$

for $l = L$ to 1 step -1 do

$\quad \forall s, w_s := w_s + \langle \log \varphi_s^{(l)}, g \rangle$

$\quad \forall s, r_s := -K^\top \left(K \left(\frac{\lambda_s g - r_s}{\varphi_s^{(l)}} \right) \odot \frac{d_s}{(K b_s^{(l-1)})^2} \right) \odot b_s^{(l-1)}$

$\quad g := \sum_s r_s$

od

comment: Backward loop - dictionary

$y := \mathbf{0}_{S \times N}$

$z := \mathbf{0}_{S \times N}$

$n := \nabla \mathcal{L}(p, x)$

for $l = L$ to 1 step -1 do

$\quad \forall s, c_s := K((\lambda_s n - z_s) \odot b_s^{(l)})$

$\quad \forall s, y_s := y_s + \frac{c_s}{K b_s^{(l-1)}}$

$\quad \forall s, z_s := -\frac{\mathbf{1}_N}{\varphi_s^{(l-1)}} \odot K^\top \frac{d_s \odot c_s}{(K b_s^{(l-1)})^2}$

$\quad n := \sum_s z_s$

od

Outputs: $P^{(L)}(D, \lambda) := p, \nabla_D \mathcal{E}^{(L)} := y, \nabla_\lambda \mathcal{E}^{(L)} := w$

Algorithm 4.1: `SinkhornGrads`: Computation of dictionary and barycentric weights gradients.

$$\mathcal{G}_L(\alpha,\beta) := \mathcal{E}_L(F(\alpha), F_S(\beta)), \tag{4.19}$$

where $F(\alpha) := (F_N(\alpha_1), \dots, F_N(\alpha_S)) = D$. Differentiating (4.19) yields

$$\nabla_\alpha \mathcal{G}_L(\alpha,\beta) = [\partial F(\alpha)]^\top \nabla_D \mathcal{E}_L(F(\alpha), F_S(\beta)) = [\partial F(\alpha)]^\top \nabla_D \mathcal{E}_L(D,\Lambda), \tag{4.20}$$
$$\nabla_\beta \mathcal{G}_L(\alpha,\beta) = [\partial F_S(\beta)]^\top \nabla_\lambda \mathcal{E}_L(F(\alpha), F_S(\beta)) = [\partial F_S(\beta)]^\top \nabla_\lambda \mathcal{E}_L(D,\Lambda), \tag{4.21}$$

where $\left[\partial F_p(u)\right]^\top = \partial F_p(u) = \left(I_p - F_p(u)\mathbf{1}_p^\top\right)\Delta\left(F_p(u)\right)$, p being either N or S for each atom or the weights, respectively, and both derivatives of $\mathcal{E}_L$ are computed using either automatic differentiation or as given in (4.16), (4.17) with Algorithm 4.1 (see Section 4.3.2). The optimization can then be performed with no constraints over α, β.

Since the resulting problem is one where the function to minimize is differentiable and we are left with no constraints, in this work we chose to use a quasi-Newton method (though our approach can be used with any appropriate solver); that is, at each iteration t, an approximation of the inverse Hessian matrix of the objective function, $B^{(t)}$, is updated, and the logistic variables for the atoms and weights are updated as

$$\alpha^{(t+1)} := \alpha^{(t)} - \rho_\alpha^{(t)} B_\alpha^{(t)} \nabla_\alpha \mathcal{G}_L(\alpha,\beta), \qquad \beta^{(t+1)} := \beta^{(t)} - \rho_\beta^{(t)} B_\beta^{(t)} \nabla_\beta \mathcal{G}_L(\alpha,\beta), \tag{4.22}$$

where the $\rho^{(t)}$ are step sizes. An overall algorithm yielding our representation in this particular setup of quasi-Newton after a logistic change of variables is given in Algorithm 4.2.

In the applications of Section 4.5, $B^{(t)}$ and $\rho^{(t)}$ were chosen using an off-the-shelf L-BFGS solver (Morales and Nocedal, 2011). We chose to perform updates to atoms and weights simultaneously. Note that in this case, both are fed to the solver of choice as a concatenated vector. It is then beneficial to add a "variable scale" hyperparameter ζ and to multiply all gradient entries related to the weights by that value. Otherwise, the solver might reach its convergence criterion when approaching a local minimum with regards to either dictionary atoms or weights, even if convergence is not yet achieved in the other. Setting either a low or high value of ζ bypasses the problem by forcing the solver to keep optimizing with regard to one of the two variables in particular. In practice, and as expected, we have observed that relaunching the optimization with different ζ values upon convergence can increase the quality of the learned representation. While analogous to the usual alternated optimization scheme often used in dictionary learning problems, this approach avoids having to compute two different forward Sinkhorn loops to obtain the derivatives in both variables.

4.3.2 Backward recursive differentiation

To differentiate $P^{(L)}(D,\Lambda)$, we first rewrite its definition (4.12) by introducing the following notations:

$$P^{(l)}(D,\lambda) = \Psi(b^{(l-1)}(D,\lambda), D, \lambda), \tag{4.23}$$
$$b^{(l)}(D,\lambda) = \Phi(b^{(l-1)}(D,\lambda), D, \lambda), \tag{4.24}$$

where

$$\Psi(b,D,\lambda) := \prod_s \left(K^\top \frac{d_s}{K b_s}\right)^{\lambda_s}, \tag{4.25}$$

$$\Phi(b,D,\lambda) := \left[\left(\frac{\Psi(b,D,\lambda)}{K^\top \frac{d_1}{K b_1}}\right)^\top, \dots, \left(\frac{\Psi(b,D,\lambda)}{K^\top \frac{d_S}{K b_S}}\right)^\top\right]^\top. \tag{4.26}$$

Inputs: Data $X \in \mathbb{R}^{N \times M}$, initial guesses $\alpha^{(0)}, \beta_1^{(0)}, \dots, \beta_M^{(0)}$, convergence criterion
$t := 0$
while convergence not achieved **do**
  $D^{(t)} := F(\alpha^{(t)})$
  $\alpha^{(t+1)} := \alpha^{(t)}$
  for $i = 1$ **to** M **step** 1 **do**
    $\lambda_i^{(t)} := F_S(\beta_i^{(t)})$
    $p_i, g_i^D, g_i^\lambda :=$ `SinkhornGrads`$(x_i, D^{(t)}, \lambda_i^{(t)})$
    Select $\rho_\alpha^{(t)}, \rho_i^{(t)} B_\alpha^{(t)}, B_i^{(t)}$ (`L-BFGS`)
    $\alpha^{(t+1)} := \alpha^{(t+1)} - \rho_\alpha^{(t)} B_\alpha^{(t)} \partial F(\alpha^{(t)}) g_i^D$
    $\beta_i^{(t+1)} := \beta_i^{(t)} - \rho_i^{(t)} B_i^{(t)} \partial F_S(\beta_i^{(t)}) g_i^\lambda$
  od
  $t := t + 1$
od
Outputs: $D = F\left(\alpha^{(t)}\right), \Lambda = \left(F_S\left(\beta_1^{(t)}\right), \dots, F_S\left(\beta_S^{(t)}\right)\right)$

Algorithm 4.2: Quasi-Newton implementation of the WDL algorithm.

Finally, we introduce the following notations for readability:

$$\xi_y^{(l)} := \left[\partial_y \xi(b^{(l)}, D, \lambda)\right]^\top, \qquad B_y^{(l)} := \left[\partial_y b^{(l)}(D, \lambda)\right]^\top, \tag{4.27}$$

where ξ can be Ψ or Φ, and y can be D or λ.

We then have

$$\nabla_D \mathcal{E}_L(D, \lambda) = \Psi_D^{(L-1)}\left(\nabla \mathcal{L}(P^{(L)}(D, \lambda), x)\right) + \sum_{l=0}^{L-2} \Phi_D^{(l)}\left(v^{(l+1)}\right), \tag{4.28}$$

$$\nabla_\lambda \mathcal{E}_L(D, \lambda) = \Psi_\lambda^{(L-1)}\left(\nabla \mathcal{L}(P^{(L)}(D, \lambda), x)\right) + \sum_{l=0}^{L-2} \Phi_\lambda^{(l)}\left(v^{(l+1)}\right), \tag{4.29}$$

where:

$$v^{(L-1)} := \Psi_b^{(L-1)}\left(\nabla L(P^{(L)}(D, \lambda), x)\right), \tag{4.30}$$

$$\forall l < L - 1, v^{(l-1)} := \Phi_b^{(l-1)}\left(v^{(l)}\right). \tag{4.31}$$

See Appendix D for proof.

4.4 Extensions

4.4.1 Log-domain stabilization

Stabilization

In its most general framework, representation learning aims at finding a useful representation of data, rather than one allowing for perfect reconstruction. In some particular cases, however, it might also be desirable to achieve a very low reconstruction error, for instance if the representation is to be used for compression of data rather than a task such as classification. In the case of our method, the quality of the reconstruction is directly

linked to the selected value of the entropy parameter γ, as it introduces a blur in the reconstructed images as illustrated in Figure 4.3. In the case where sharp features in the reconstructed images are desired, we need to take extremely low values of γ, which can lead to numerical problems, e.g. because values within the scaling vectors a and b can then tend to infinity. As suggested by Chizat et al. (2018) and Schmitzer (2019), we can instead perform the generalized Sinkhorn updates (4.12–4.13) in the log-domain. Indeed, noting $u_s^{(l)}, v_s^{(l)}$ as the dual scaling variables, that is,

$$a_s^{(l)} := \exp\left(\frac{u_s^{(l)}}{\gamma}\right), \qquad b_s^{(l)} := \exp\left(\frac{v_s^{(l)}}{\gamma}\right), \tag{4.32}$$

the quantity $-c_{ij} + u_i + v_j$ is known to be bounded and thus remains numerically stable. We can then introduce the stabilized kernel $\tilde{K}(u, v)$ defined as

$$\tilde{K}(u, v) := \exp\left(\frac{-C + u\mathbf{1}^\top + \mathbf{1}v^\top}{\gamma}\right), \tag{4.33}$$

and notice that we then have

$$u_s^{(l)} = \gamma\left[\log(d_s) - \log(Kb_s^{(l-1)})\right], \tag{4.34}$$

$$\left[\log(Kb_s^{(l-1)})\right]_i = \log\left[\sum_j \exp\left(\frac{-c_{ij} + v_j^{(l-1)}}{\gamma}\right)\right] \tag{4.35}$$

$$= \log\left(\sum_j \tilde{K}(u_s^{(l-1)}, v_s^{(l-1)})_{.j}\right) - \frac{\left[u_s^{(l-1)}\right]_i}{\gamma}. \tag{4.36}$$

With similar computations for the v_s updates, we can then reformulate the Sinkhorn updates in the stabilized domain as

$$u_s^{(l)} := \gamma\left[\log(d_s) - \log\left(\sum_j \tilde{K}(u_s^{(l-1)}, v_s^{(l-1)})_{.j}\right)\right] + u_s^{(l-1)}, \tag{4.37}$$

$$v_s^{(l)} := \gamma\left[\log(P^{(l)}) - \log\left(\sum_i \tilde{K}(u_s^{(l)}, v_s^{(l-1)})_{i.}\right)\right] + v_s^{(l-1)}. \tag{4.38}$$

This provides a forward scheme for computing Wasserstein barycenters with arbitrarily low values of γ, which could be expanded to the backward loop of our method either by applying an automatic differentiation tool to the stabilized forward barycenter algorithm or by changing the steps in the backward loop of Algorithm 4.1 to make them rely solely on stable quantities. However, this would imply computing a great number of stabilized kernels as in (4.33), which relies on nonseparable operations. Each of those kernels would also have to either be stored in memory or recomputed when performing the backward loop. In both cases, the cost in memory or number of operations, respectively, can easily be too high in large scale settings.

Separable log kernel

These issues can be avoided by noticing that when the application of the kernel K is separable, this operation can be performed at a much lower cost. For a d-dimensional histogram of $N = n^d$ bins, applying a separable kernel amounts to performing a sequence of d steps, where each step computes n operations per bin. It results

in a $O(n^{d+1}) = O(N^{\frac{d+1}{d}})$ cost instead of $O(N^2)$. As mentioned previously, the stabilized kernel (4.33) is not separable, prompting us to introduce a new stable and separable kernel suitable for log-domain processing. We illustrate this process using 2D kernels without loss of generality. Let $\mathcal{X}$ be a 2D domain discretized as an $n \times n$ grid. Applying a kernel of the form $K = \exp(-\frac{C}{\gamma})$ to a 2D image $b \in \mathcal{X}$ is performed as such:

$$R(i,j) := \sum_{k=1}^{n} \sum_{l=1}^{n} \exp\left(-\frac{C((i,j),(k,l))}{\gamma}\right) b(k,l)\,, \tag{4.39}$$

where $C((i,j),(k,l))$ denotes the cost to transport mass between the points (i,j) and (k,l).

Assuming a separable cost such that $C((i,j),(k,l)) := C_y(i,k) + C_x(j,l)$, it amounts to performing two sets of 1D kernel applications:

$$A(k,j) = \sum_{l=1}^{n} \exp\left(\frac{C_x(j,l)}{\gamma}\right) b(k,l), \tag{4.40}$$

$$R(i,j) = \sum_{k=1}^{n} \exp\left(\frac{C_y(i,k)}{\gamma}\right) A(k,j)\,. \tag{4.41}$$

In order to stabilize the computation and avoid reaching representation limits, we transfer it to the log-domain ($v := \log(b)$). Moreover, we shift the input values by their maximum and add it at the end. The final process can be written as the operator $K_{LS} : \log(b) \to \log(K(b))$ with K a separable kernel, and is described in Algorithm 4.3.

Inputs: Cost matrix $C \in \mathbb{R}^{N \times N}$, image in log-domain $v \in \mathbb{R}^{n \times n}$

$\forall k,j,\ x_l(k,j) := \frac{C_x(j,l)}{\gamma} + v(k,l)$

$\forall k,j,\ A'(k,j) := \log\left(\sum_l^n \exp(x_l - \max_l x_l)\right) + \max_l x_l$

$\forall i,j,\ y_k(i,j) := \frac{C_y(i,k)}{\gamma} + A'(k,j)$

$\forall i,j,\ R'(i,j) := \log\left(\sum_k^n \exp(y_k - \max_k y_k)\right) + \max_k y_k$

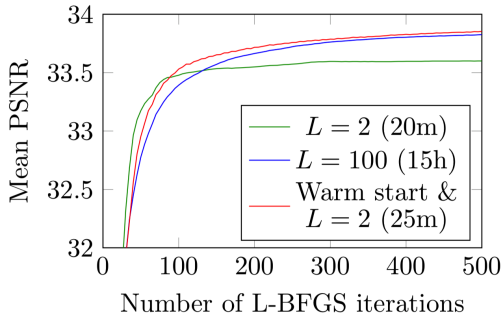
Outputs: Image in log-domain $K_{LS}(v) = R'$

Algorithm 4.3: `LogSeparableKer` K_{LS}: Application of a 2D separable kernel in log-domain.

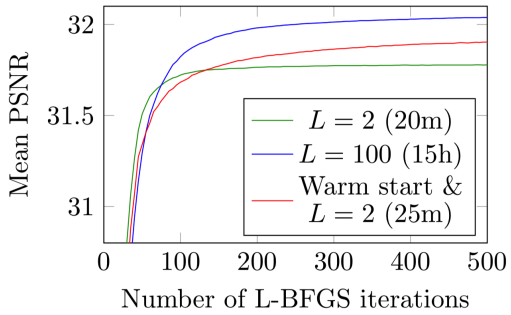
This operator can be used directly in the forward loop, as seen in Algorithm E.1. For backward loops, intermediate values can be negative and real-valued logarithms are not suited. While complex-valued logarithms solve this problem, they come at a prohibitive computational cost. Instead, we store the sign of the input values and compute logarithms of absolute values. When exponentiating, the stored sign is used to recover the correct value.

4.4.2 Warm start

Warm start, often used in optimization problems, consists in using the solution of a previous optimization problem, close to the current one, as the initialization point in order to speed up the convergence. Our method relies on performing an iterative optimization process (for example, L-BFGS in the following experiments) which, at each iteration, calls upon *another* iterative scheme: the forward Sinkhorn loop to compute the barycenter and its automatic differentiation to obtain gradients. As described in Section 4.2.2, this second, nested iterative scheme is usually initialized with constant scaling vectors. However, in our case, since each iteration of our descent method performs a new Sinkhorn loop, the scaling vectors of the previous iteration can be used to set the values of $b^{(0)}$ instead of the usual $\mathbf{1}_{NS}$, thus "warm-starting" the barycenter computation. In the remainder of this subsection, for illustrative purposes, we will focus on our particular case where the chosen descent method is L-BFGS, though the idea of applying warm start to the generalized Sinkhorn algorithm should be directly applicable with any other optimization scheme.



(a) MUG dataset: woman



(b) MUG dataset: man

Figure 4.4: Evolution of the mean PSNR of the reconstructions per L-BFGS iteration, for three configurations, on two datasets. The KL loss was used for this experiment. We see that the warm start yields better reconstructions with the same number of Sinkhorn iterations (L) in roughly the same time.

As an example, in our case, instead of a single L-BFGS step after $L = 500$ Sinkhorn iterations, we perform an L-BFGS step every $L = 10$ iterations, initializing the scaling vectors as the ones reached at the end of the previous 10. This technique accumulates the Sinkhorn iterations as we accumulate L-BFGS steps. This has several consequences: a gain in precision and time, a potential increase in the instability of the scaling vectors, and changes in the energy we minimize.

First, the last scaling vectors of the previous overall iteration are closer to that of the current one than a vector of constant value. Therefore, the Sinkhorn algorithm converges more rapidly, and the final barycenters computed at each iteration gain accuracy compared to the classical version of the algorithm.

Second, as mentioned in Section 4.4.1, the scaling vectors may become unstable when computing a large number of iterations of the Sinkhorn algorithm. When using a warm start strategy, Sinkhorn iterations tend to accumulate, which may consequently degrade the stability of the scaling vectors. For example, using 20 Sinkhorn iterations running through 50 L-BFGS steps, a warm start would lead to barycenters computed using scaling vectors comparable to those obtained after 1000 Sinkhorn iterations. When instabilities become an issue, we couple the warm start approach with our log-domain stabilization. The reduced speed of log-domain computations is largely compensated by the fact that our warm start allows the computation of fewer Sinkhorn iterations for an equivalent or better result.

Third, when differentiating (4.15), we consider the initial, warm-started (as opposed to initializing $b^{(0)}$ to $\mathbf{1}_{NS}$) values given to the scaling vectors to be constant and independent of weights and atoms. This amounts to considering a different energy to minimize at each L-BFGS step.

We demonstrate the benefits of the warm start in Figure 4.4. We plot the evolution of the mean Peak SNR (PSNR) of the reconstructions throughout the L-BFGS iterations for different settings for the two datasets used in Section 4.5.4. For these examples, we used the KL loss (since it gave the best reconstructions overall), we did not have to use the log-domain stabilization, and we restarted L-BFGS every 10 iterations. At an equal number of Sinkhorn iterations (L), enabling the warm start always yields better reconstructions after a certain number of iterations. It comes at a small overhead cost in time (around 25%) because the L-BFGS line search routine requires more evaluations at the start. For the example in Figure 4.4a, the computation times are 20 minutes for $L = 2$, 25 minutes for the warm restart and $L = 2$, and 15 hours for $L = 100$. In this particular case, enabling the warm start with two Sinkhorn iterations yields even better results than having 100 Sinkhorn iterations without a warm start and with a 36 gain factor in time. For the second dataset (Figure 4.4b), enabling the warm start does not yield results as good as when running 100 Sinkhorn iterations. However, it would require considerably more than two Sinkhorn iterations, and hence a lot more time, to achieve the same result without it. The computation times in all three cases are similar to the previous example.

4.4.3 Sinkhorn heavyball

As part of a generalization of the Sinkhorn algorithm for solving OT between tensor fields, Peyré et al. (2016) introduced relaxation variables. In the particular case of scalar OT (our framework in the present work), these relaxation variables amount to an averaging step in the Sinkhorn updates; for instance, in the case of the barycenter scaling updates (4.11), (4.13),

$$\tilde{a}_s^{(l)} = \frac{d_s}{K b_s^{(l-1)}}, \tag{4.42}$$

$$a_s^{(l)} = \left(a_s^{(l-1)}\right)^{\tau} \left(\tilde{a}_s^{(l)}\right)^{1-\tau},$$

$$\tilde{b}_s^{(l)} = \frac{P^{(l)}(D, \lambda)}{K^{\top} a_s^{(l)}}, \tag{4.43}$$

$$b_s^{(l)} = \left(b_s^{(l-1)}\right)^{\tau} \left(\tilde{b}_s^{(l)}\right)^{1-\tau}.$$

$\tau = 0$ yields the usual Sinkhorn iterations, but it has been shown that negative values of τ produce extrapolation and can lead to a considerable increase in the rate of convergence of the Sinkhorn algorithm (Peyré et al., 2016, Remark 6). This effect can be thought of in the same way as the heavyball method (Zavriev and Kostyuk, 1993; Nesterov, 2018), often used in optimization problems and dating back to Polyak (1964), i.e. as the addition of a momentum term (e.g., $(a_s^{(l-1)}/\tilde{a}_s^{(l)})^{\tau}$, which amounts to $\tau(u_s^{(l-1)} - \tilde{u}_s^{(l)})$ in the log-domain) to the usual Sinkhorn updates. This acceleration scheme can be used within our method by applying an automatic differentiation tool (Theano Development Team, 2016) to the forward Sinkhorn loop yielding the barycenter (shown in Algorithm G.1 in the Appendices) and feeding the gradients to Algorithm 4.2.

4.4.4 Unbalanced

In (4.4), we defined the set of admissible transport plans $\Pi(p, q)$ as the set of matrices whose marginals are equal to the two input measures, that is, with rows summing to p and columns summing to q. Equivalently, we can reformulate the definition of the approximate Wasserstein distance (4.8) as

$$W_\gamma(p, q) := \min_{T \in \mathbb{R}_+^{N \times N}} \langle T, C \rangle + \gamma H(T) + \iota_{\{p\}}(T\mathbf{1}_N) + \iota_{\{q\}}(T^{\top}\mathbf{1}_N), \tag{4.44}$$

where ι is the indicator function defined in (3.17). Chizat et al. (2018) introduce the notion of unbalanced transport problems, wherein this equality constraint between the marginals of the OT plan and the input measures is replaced by some other similarity criterion. Using entropic regularization, they introduce matrix scaling algorithms generalizing the Sinkhorn algorithm to compute, among others, unbalanced barycenters. This generalizes the notion of approximate Wasserstein barycenters that we have focused on thus far.

In particular, using the KL divergence between the transport plan's marginals and the input measures allows for less stringent constraints on mass conservation, which can in turn yield barycenters which maintain more of the structure seen in the input measures. This amounts to using the following definition of W_γ in the barycenter formulation (4.10):

$$W_\gamma(p, q) := \min_{T \in \mathbb{R}_+^{N \times N}} \langle T, C \rangle + \gamma H(T) + \rho \left(\mathrm{KL}(T\mathbf{1}_N | p) + \mathrm{KL}(T^{\top}\mathbf{1}_N | q)\right), \tag{4.45}$$

where $\rho > 0$ is the parameter determining how far from the balanced OT case we can stray, with $\rho = \infty$ yielding the usual OT formulation. In this case, the iterative matrix scaling updates (4.12–4.13) read, respectively (Chizat et al., 2018), as

$$P^{(l)}(D,\lambda) = \left(\sum_{s=1}^{S} \lambda_s \left(K^\top a_s^{(l)}\right)^{\frac{\gamma}{\rho+\gamma}}\right)^{\frac{\rho+\gamma}{\gamma}}, \tag{4.46}$$

$$a_s^{(l)} = \left(\tilde{a}_s^{(l)}\right)^{\frac{\rho}{\rho+\gamma}}, b_s^{(l)} = \left(\tilde{b}_s^{(l)}\right)^{\frac{\rho}{\rho+\gamma}}, \tag{4.47}$$

where $\tilde{a}_s^{(l)}, \tilde{b}_s^{(l)}$ are obtained from the usual Sinkhorn updates as in (4.42), (4.43).

Algorithm G.2, given in the Appendices, performs the barycenter computation (forward loop) including both the unbalanced formulation and the acceleration scheme shown in Section 4.4.3. Automatic differentiation can then be performed using an appropriate library (Theano Development Team, 2016) to obtain the dictionary and weights gradients, which can then be plugged into Algorithm 4.2 to obtain a representation relying on unbalanced barycenters.

4.5 Applications

4.5.1 Comparison with Wasserstein principal geodesics

As mentioned in Section 4.1.1, an approach to generalize PCA to the set of probability measures on some space, endowed with the Wasserstein distance, has recently been proposed (Seguy and Cuturi, 2015). Given a set of input measures, an approximation of their Wasserstein Principal Geodesics (WPGs) can be computed, namely geodesics that pass through their isobarycenter (in the Wasserstein sense) and are close to all input measures. Because of the close link between Wasserstein geodesics and the Wasserstein barycenter, it would stand to reason that the set of barycenters of $S = 2$ atoms learned using our method could be fairly close to the first WPG. In order to test this, and to compare both approaches, we reproduce the setting of the WPG paper (Seguy and Cuturi, 2015) experiment on the Modified National Institute of Standards and Technology database (MNIST) within our framework.

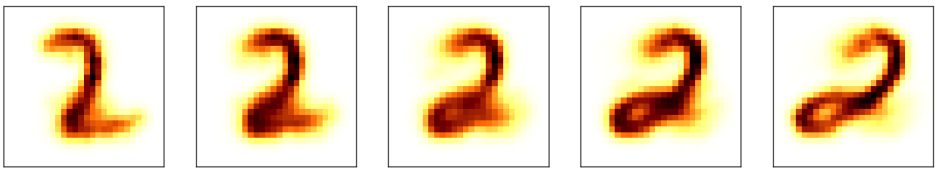


Figure 4.5: Span of our 2-atom dictionary for weights $(1-t,t), t \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$, when trained on images of digit 2.

We first run our method to learn two atoms on samples of 1000 images for each of the first four nonzero digits, with parameters $\gamma = 2.5, L = 30$, and compare the geodesic that runs in between the two learned atoms with the first WPG. An example of the former is shown in Figure 4.5. Interestingly, in this case, as with the 3's and 4's, the two appear visually extremely close (see the first columns of Seguy and Cuturi, 2015, Figure 5, for the first WPG). It appears our method *can* thus capture WPGs. We do not seem to recover the first WPG when running on the dataset made up of 1's, however. This is not unexpected, as several factors can cause the representation we learn to vary from this computation of the first WPG:

- In our case, there is no guarantee the isobarycenter of all input measures lies within the span of the learned dictionary.
- Even when it does, since we minimize a non-convex function, the algorithm might converge toward another local minimum.

- In this experiment, the WPGs are computed using several approximations (Seguy and Cuturi, 2015), including some for the computation of the geodesics themselves, which we are not required to make in order to learn our representation.

Note that in the case of this particular experiment (on a subsample of MNIST 1's), we tried relaunching our method several times with different random initializations and never observed a span similar to the first WPG computed using these approximations.

Our approach further enables us to combine, in a straightforward way, each of the captured variations when learning more than two atoms. This is illustrated in Figure 4.6, where we run our method with $S = 3$. Warpings similar to those captured when learning only $S = 2$ atoms (the appearance of a loop within the 2) are also captured, along with others (shrinking of the vertical size of the digit toward the right). Intermediate values of the weight given to each of the three atoms allow our representation to cover the whole simplex, thus arbitrarily combining any of these captured warpings (e.g., vertically shrinked, loopless 2 in the middle of the bottom row).

Figures similar to Figure 4.5 and 4.6 for all other digits are given in the Appendices, Section H.1.

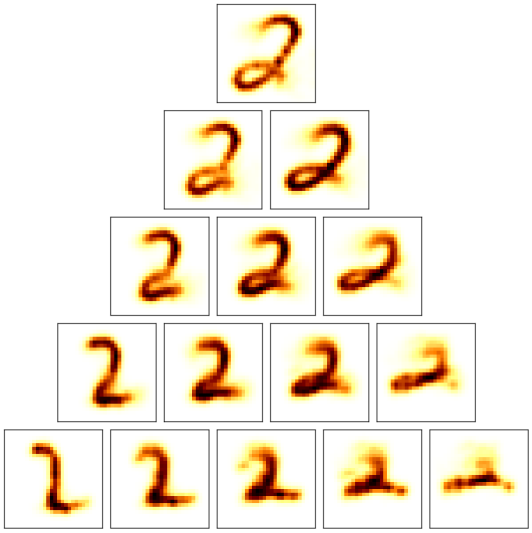
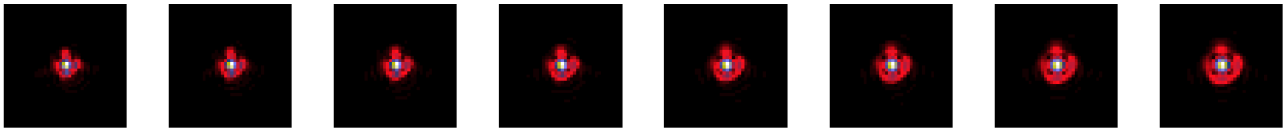


Figure 4.6: Span of a 3-atom dictionary learned on a set of 2's. Weights along each edge are the same as in Figure 4.5 for the two extreme vertices and 0 for the other, while the three center barycenters have a weight of $\frac{1}{2}$ for the atom corresponding to the closest vertex and $\frac{1}{4}$ for each of the other two.

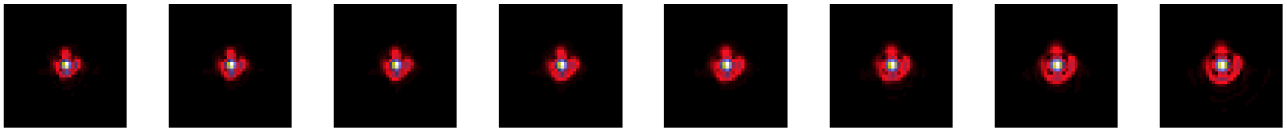
4.5.2 Point spread functions

As with every optical system, observations from astrophysical telescopes suffer from a blurring related to the instrument's optics and various other effects (such as the telescope's jitter for space-based instruments). The blurring function, or PSF, can vary spatially (across the instrument's FOV), temporally and chromatically (with the incoming light's wavelength). In order to reach its scientific goals, ESA's upcoming *Euclid* space mission (Laureijs et al., 2011) will need to measure the shape of billions of galaxies extremely accurately, and therefore correcting the PSF effects is of paramount importance. The use of OT for PSF modelling has been investigated by Irace and Batatia (2013) and Ngolè and Starck (2017), both with the aim of capturing the spatial variation of the PSF. For any given position in the FOV, the transformations undergone by the PSF depending on the incoming light's wavelength are also known to contain strong geometrical information, as illustrated in Figure 4.7. It is therefore tempting to express these variations as the intermediary steps in the optimal transportation between the PSFs at the two extreme wavelengths. This succession of intermediary steps, the *displacement interpolation* (also known as McCann's interpolation McCann, 1997) between two measures,



(a) 550nm (b) 600nm (c) 650nm (d) 700nm (e) 750nm (f) 800nm (g) 850nm (h) 900nm

Figure 4.7: Simulated *Euclid*-like PSF variation at a fixed position in the FOV for varying incoming wavelengths.



(a) 550nm (b) 600nm (c) 650nm (d) 700nm (e) 750nm (f) 800nm (g) 850nm (h) 900nm

Figure 4.8: Polychromatic variations of PSFs by displacement interpolation.

can be computed (in the case of the 2-Wasserstein distance) as their Wasserstein barycenters with weights $\lambda = (1-t, t), t \in [0, 1]$ (Agueh and Carlier, 2011).

We thus ran our method on a dataset of simulated, *Euclid*-like PSFs (described in Section 3.3.4 and Ngolè et al., 2015, Section 4.1) at various wavelengths and learned only two atoms. The weights were initialized as a projection of the wavelengths into $[0, 1]$ but allowed to vary. The atoms were initialized without using any prior information as two uniform images with all pixels set at $1/N$, N being the number of pixels (in this case 40^2). The fitting loss was quadratic, the entropic parameter γ set to a value of 0.5 to allow for sharp reconstructions, and the number of Sinkhorn iterations set at 120, with a heavyball parameter $\tau = -0.1$.

The learned atoms, as well as the actual PSFs at both ends of the spectrum, are shown in Figure 4.9. Our method does indeed learn atoms that are extremely close visually to the two extremal PSFs. The reconstructed PSFs at the same wavelength as those of Figure 4.7 are shown in Figure 4.8 (the corresponding final barycentric weights are shown in Figure 4.11b). This shows that OT, and in particular displacement interpolation, does indeed capture the geometry of the polychromatic transformations undergone by the PSF. On the other hand, when one learns only two components using a PCA, they have no direct interpretation (see Figure 4.10), and the weights given to the second principal component appear to have no direct link to the PSF's wavelength, as shown in Figure 4.11a.

Note that while adding constraints can also make linear generative methods yield two atoms that are visually close to the extreme PSFs, for instance by using NMF instead of PCA (see Figure H.5 in the Appendices), our method yields lower reconstruction error, with an average normalized mean square error of 1.71×10^{-3} across the whole dataset, as opposed to 2.62×10^{-3} for NMF. As expected, this difference in reconstruction error is particularly noticeable for datapoints corresponding to wavelengths in the middle of the spectrum, as the NMF reconstruction then simply corresponds to a weighted sum of the two atoms, while our method captures more complex warping between them. This shows that the OT representation allows us to better capture the nonlinear geometrical variations due to the optical characteristics of the telescope.

4.5.3 Cardiac sequences

We tested our dictionary learning algorithm on a reconstructed Magnetic Resonance Imaging (MRI) sequence of a beating heart. The goal was to learn a dictionary of four atoms, representing the key frames of the sequence.

An advantageous side effect of the weights learned by our method lying in the simplex is that it provides a natural way to visualize them: by associating each atom d_i with a fiducial position $(x_i, y_i) \in \mathbb{R}^2$, each set of weights can be represented as one point placed at the position of the Euclidean barycenter of these positions,

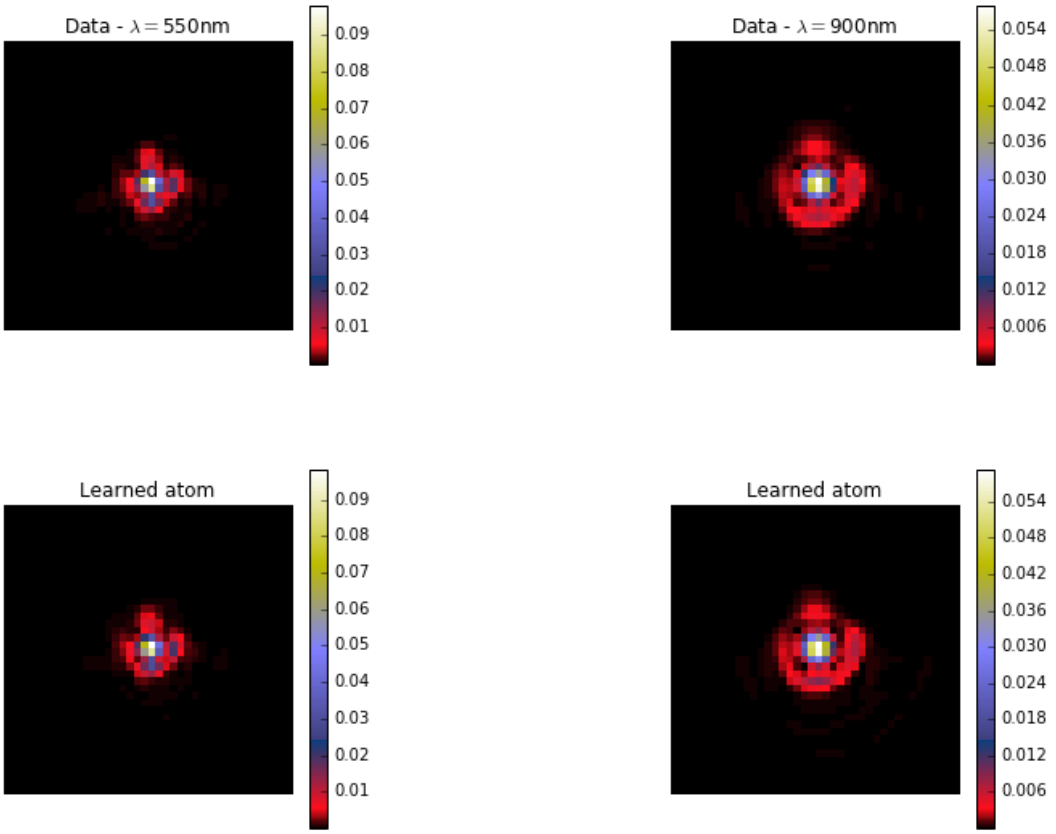


Figure 4.9: Extreme wavelength PSFs in the dataset and the atoms making up the learned dictionary.

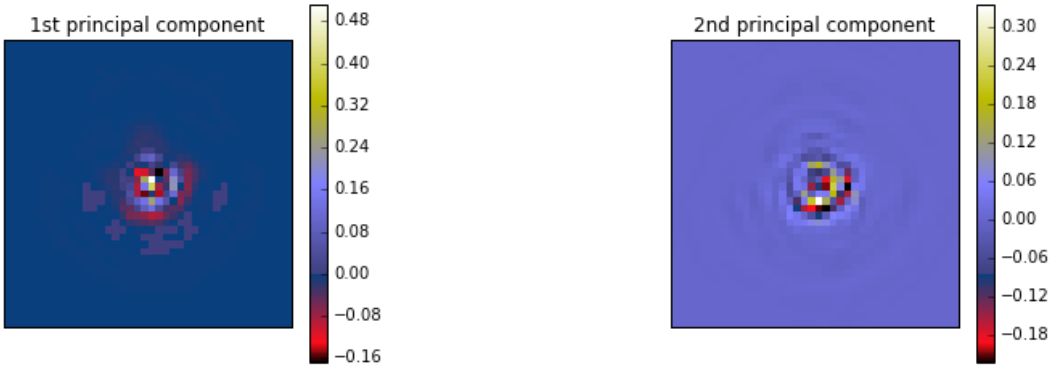
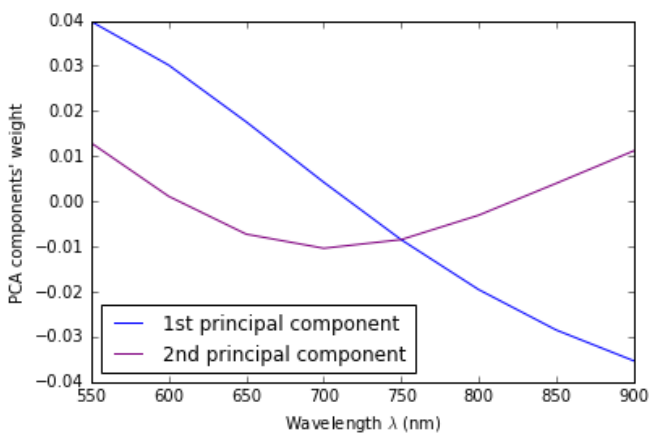
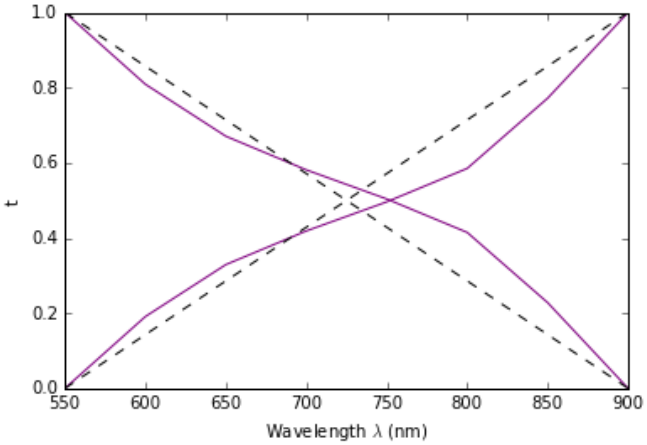


Figure 4.10: PCA-learned components.



(a) Weights for the first two principal components learned by PCA.



(b) Barycentric weights learned by our method. The dashed lines are the initialization.

Figure 4.11: Evolution of representation coefficients by wavelength.

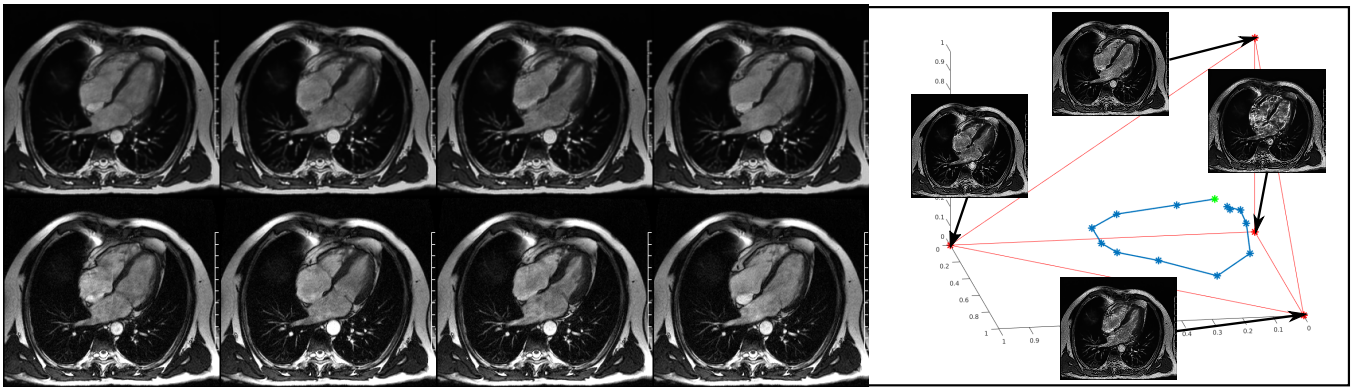


Figure 4.12: Left: Comparison between four frames (out of 13) of the measures (lower row) and the same reconstructed frames (upper row). Right: plot of the reconstructed frames (blue points) by their barycentric coordinates in the 4-atom basis, with each atom (red points) at the vertices of the tetrahedra. The green point is the first frame.

with individual weights given to the corresponding atom. Up to rotations and inverse ordering, there are only as many such representations as there are possible orderings of the atoms. In the present case of $S = 4$, we can further use the fact that any of the four weights λ_i is perfectly known through the other three as $1 - \sum_{j \neq i} \lambda_j$. By giving atoms fiducial positions in $\mathbb{R}^3$ and ignoring one of them or, equivalently, assigning it the $(0, 0, 0)$ position, we thus obtain a unique representation of the weights as seen in Figure 4.12. The "barycentric path" (polyline of the barycentric points) is a cycle, which means the algorithm is successful at finding those key frames that, when interpolated, can represent the whole dataset. This is confirmed by the similarity between the reconstructions and the input measures.

For this application, we used 13 frames of 272×240, a regularization $\gamma = 2$, and a scale between weights and atoms of $\zeta = N/(100 * M)$, $N = 272 \times 240$, $M = 13$ frames. Initialization was random for the weights, and constant for the atoms. We used a quadratic loss because it provided the best results in terms of reconstruction and representation. We found 25 iterations for the Sinkhorn algorithm to be a good trade-off between computation time and precision.

4.5.4 Wasserstein faces

It has been shown that images of faces, when properly aligned, span a low-dimensional space that can be obtained via PCA. These principal components, called Eigenfaces, have been used for face recognition (Turk and Pentland, 1991). We show that, with the right setting, our dictionary learning algorithm can produce atoms that can be interpreted more easily than their linear counterparts, and can be used to edit a human face's appearance.

We illustrate this application on the Multimedia Understanding Groups database (MUG) of facial expression (Aifanti et al., 2010). From the raw images of the MUG database, we isolated faces and converted the images to grayscale. The resulting images are in the top rows of Figure 4.13. We can optionally invert the colors and apply a power factor α similarly to a gamma-correction. We used a total of 20 (224×224) images of a single person performing five facial expressions and learned dictionaries of five atoms using PCA, NMF, a K-SVD implementation (Rubinstein et al., 2008), and our proposed method. For the last, we set the number of Sinkhorn iterations to 100 and the maximum number of L-BFGS iterations to 450. The weights were randomly initialized, and the atoms were initialized as constant.

We performed a cross validation using two datasets, four loss functions, four values for α $(1, 2.2, 3, 5)$, and colors either inverted or not. We found that none of the α values we tested gave significantly better results (in terms of reconstruction errors). Interestingly, however, inverting colors improved the result for our method in most cases. We can conclude that when dealing with faces, it is better to transport the thin and dark zones (eyebrows, mouth, creases) than the large and bright ones (cheeks, forehead, chin).

As illustrated by Figure 4.13 (and H.7 in the Appendices), our method reaches similarly successful reconstructions given the low number of atoms, with a slightly higher mean PSNR of 33.8 compared to PSNRs of 33.6, 33.5 and 33.6 for PCA, NMF and K-SVD respectively.

We show in Figure 4.14 (and H.6 in the Appendices) the atoms obtained when using different loss functions. This shows how sensible the learned atoms are to the chosen fitting loss, which highlights the necessity for its careful selection if atoms' interpretability is important for the application at hand.

Finally, we showcase an appealing feature of our method: the atoms that it computes allow for facial editing. We demonstrate this application in Figure 4.15. Starting from the isobarycenter of the atoms, by interpolating weights towards a particular atom, we add some of the corresponding emotion to the face.

4.5.5 Literature learning

We use our algorithm to represent literary work. To this end, we use a bag-of-words representation (Salton and McGill, 1983), where each book is represented by a histogram of its words. In this particular application, the cost matrix C (distance between each word) is computed exhaustively and stored. We use a semantic distance between words. These distances were computed from the Euclidian embedding provided by the GloVe database (Global Vectors for Word Representation, Pennington et al., 2014).

Our learning algorithm is unsupervised and considers similarity between books based on their lexical fields. Consequently we expect it to sort books by either author, writing style, or genre.

To demonstrate our algorithm's performance, we created a database of 20 books by five different authors. In order to keep the problem size reasonable we only considered words that are between seven and eight letters long. In our case, it is better to deal with long words because they have a higher chance of holding discriminative information than shorter ones.

The results can be seen in Figure 4.16. Our algorithm is able to group the novels by author, recognizing the proximity of lexical fields across the different books. Atom 0 seems to be representing Charlotte Brontë's style, atoms 1 and 4 that of Mark Twain, atom 2 that of Arthur Conan Doyle, and atom 3 that of Jane Austen. Charles Dickens appears to share an extended amount of vocabulary with the other authors without it differing enough to be represented by its own atom, like others are.

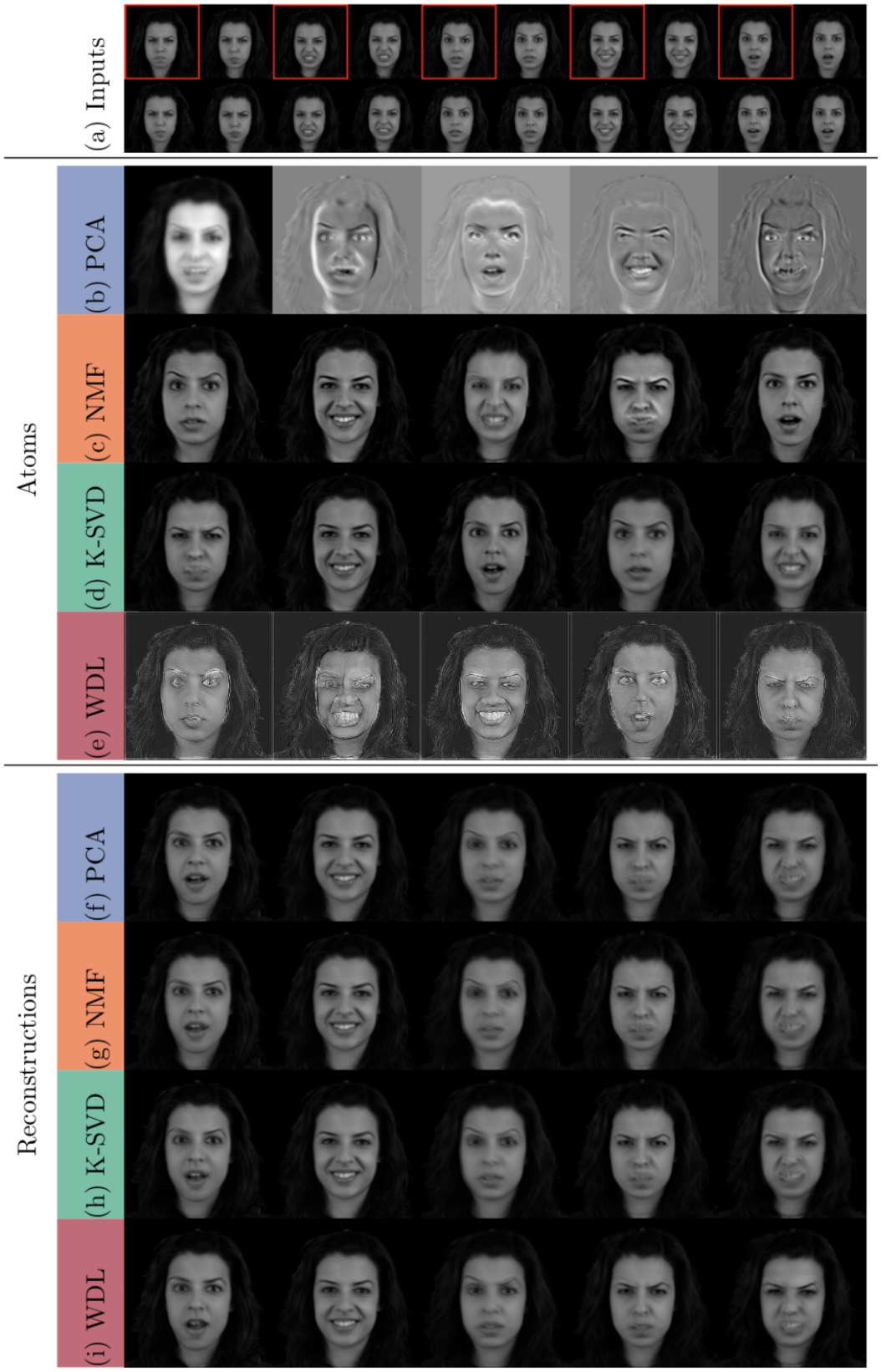


Figure 4.13: We compare our method with Eigenfaces (Turk and Pentland, 1991), NMF and K-SVD (Rubinstein et al., 2008) as a tool to represent faces on a low-dimensional space. Given a dataset of 20 images of the same person from the MUG (Aifanti et al., 2010) performing five facial expressions four times (row (a) illustrates each expression), we project the dataset on the first five Eigenfaces (row (b)). The reconstructed faces corresponding to the highlighted input images are shown in row (f). Rows (c) and (d), respectively, show atoms obtained using NMF and K-SVD and rows (g) and (h) their respective reconstructions. Using our method, we obtain five atoms shown in row (e) that produce the reconstructions in row (i).

Figure 4.14: We compare the atoms (columns 1 to 5) obtained using different loss functions, ordered by the fidelity of the reconstructions to the input measures (using the mean PSNR), from best to worst: the KL divergence (a) $\overline{PSNR} = 32.03$, the quadratic loss (b) $\overline{PSNR} = 31.93$, the total variation loss (c) $\overline{PSNR} = 31.41$, and the Wasserstein loss (d) $\overline{PSNR} = 30.33$. In the last column, we show the reconstruction of the same input image for each loss. We notice that from (a) to (d), the atoms' visual appearance seems to increase even though the reconstruction quality decreases.

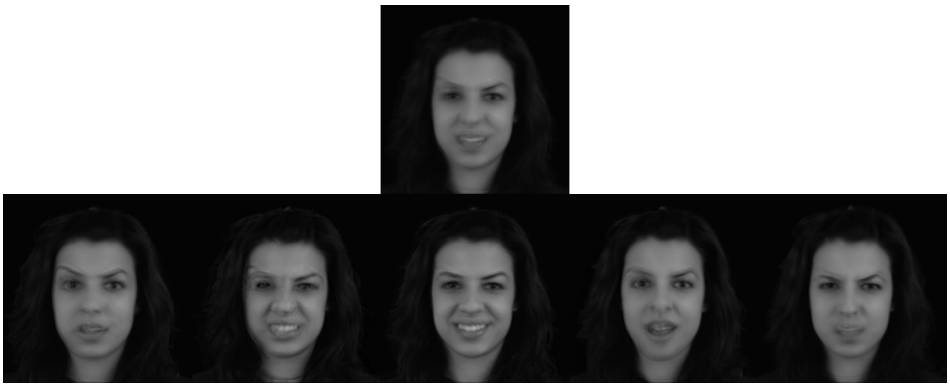


Figure 4.15: Face editing : Using the atoms shown in row (a) of Figure H.6, we interpolate between the atoms' isobarycenter (top image) and each one of the atoms (giving it a relative contribution of 70%). This allows us to emphasize each emotion (bottom images) when starting from a neutral face.

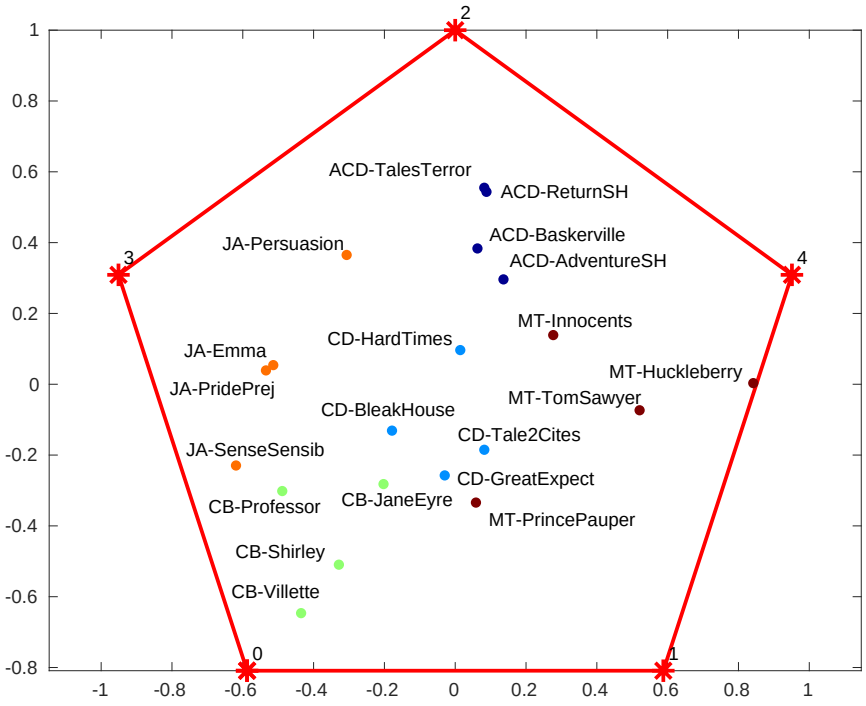


Figure 4.16: Using our algorithm, we look at word histograms of novels and learn five atoms in a sample of 20 books by five authors. Each book is plotted according to its barycentric coordinates with regard to the learned atoms, as explained in Section 4.5.3.

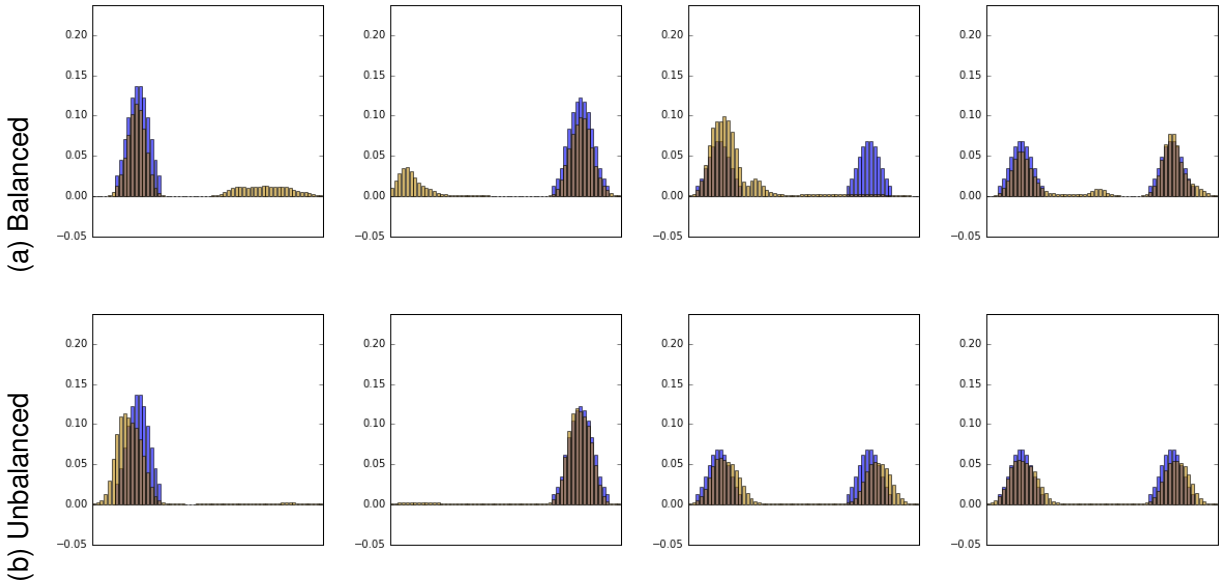


Figure 4.17: Four different original datapoints (in blue) and their reconstructions (in yellow) from our method in both the balanced (top row) and unbalanced (bottom row) settings. In the balanced case, we see the appearance of spurious modes where there was no mass in the original data or a lack of mass where there was a mode originally (the third example). Conversely, in the unbalanced case, our approach always places mass at the right positions on the grid.

4.5.6 Multimodal distributions

It is a well-known limitation of the regular OT-based Wasserstein barycenters that when there are several distinct areas containing mass, the supports of which are disjoint on the grid, the barycenter operator will still produce barycenters with mass in between them. To illustrate the advantages of using the unbalanced version our method introduced in Section 4.4.4 and the use cases where it might be preferable to do so, we place ourselves in such a setting.

We generate a dataset as follows: A 1D grid is separated into three equal parts, and while the center part is left empty, we place two discretized and truncated 1D Gaussians with the same standard deviation, their mean randomly drawn from every other appropriate position on the grid. We draw 40 such datapoints, yielding several distributions with either one (if the same mean is drawn twice) or two modes in one of the two extreme parts of the grid or one mode in each.

We then run our method in both the balanced and the unbalanced settings. In both cases, γ is set to 7, 100 Sinkhorn iterations are performed, the loss is quadratic, and the learned dictionary is made up of three atoms. In the unbalanced case, the KL-regularization parameter is set as $\rho = 20$.

Figure 4.17 shows examples of the input data and its reconstructions in both settings. In the unbalanced case, our method always yields the right number of modes in the right parts of the grid. Running our method with balanced Wasserstein barycenters, however, leads to reconstructions featuring mass in parts of the grid where there was none in the original datapoint (the two left-most examples). Parts of the grid where the datapoint featured a mode can also be reconstructed as empty (the third example). Lastly, we observe mass in areas of the grid that were empty for *all datapoints* (the fourth example).

4.6 Conclusion to Chapter 4

This chapter introduces a nonlinear dictionary learning approach that uses OT geometry by fitting data to Wasserstein barycenters of a list of learned atoms. We offer schemes to compute this representation based on the addition of an entropic penalty to the definition of OT distances, as well as several variants and extensions of our method. We illustrate the representation our approach yields on several different applications.

Some very recent works present a faster Sinkhorn routine, such as the Greenkhorn algorithm (Altschuler et al., 2017) or a multiscale approach (Schmitzer, 2019). These could be integrated into our method along with automatic differentiation in order to speed up the algorithm.

Chapter 5

Polychromatic PSF field estimation

Contents

Space telescopes like *Euclid* are particularly adapted to WL studies because of their small and very stable PSF. However, a space observatory comes with its own complications. In Section 2.4.1 and Chapter 3, for example, we showed the complex profile of a diffraction-limited PSF leads to higher-order effects in its impact on galaxy shape measurement.

Another issue that will need to be addressed with *Euclid* is the chromatic variations of the PSF (Cypriano et al., 2010; Eriksen and Hoekstra, 2018). The PSF is not just a function of the position within the focal plane where the objects lie, but also of the wavelength of incoming light. This effect is especially significant in the case of the VIS instrument because of its very wide band. The effective PSF applied to each object is therefore a function of its SED. In order to correctly measure a galaxy's shape, we must find the PSF that corresponds to this specific SED. This means we need to recover chromatic variations of the PSF from stellar observations. This information can then be used to estimate the PSF not just at the position of the galaxy, but also with its own SED.

Note that while the issue we discuss here is the chromatic variation of the *Optical PSF*, Meyers and Burchat (2015) showed that seeing was also wavelength-dependent. This means that ground-based PSF can also vary with wavelength. Carlsten et al. (2018) study this effect in the case of the HSC survey, and propose a correction scheme to apply to the size of the PSF model. Both of these studies conclude that tnderstanding the chromatic variations of the PSF will also be crucial for WL studies with the upcoming, ground-based LSST.

The main aim of this thesis is the development of non-parametric methods to model the VIS PSF. No existing method (see Section 2.4.3) can recover chromatic variations. In this chapter, we propose the first non-parametric PSF model capable of doing so. We start from the approach proposed in Chapter 3 and Paper II, and expand it using the OT concepts developed in Chapter 4 and Paper I. This work was carried out in collaboration with Jean-Luc Starck, Fred Ngolè, and graduate student Rebeca Araripe, who contributed

significantly to improving the method, and to both identifying and addressing some of its limitations during her internship at CEA.

5.1 Chromatic variations of the PSF

5.1.1 Overview

We start by describing the formalism of a wavelength-dependent PSF. Following Section 3.3.1, let $\mathcal{H}$ denote the true PSF. We now consider it to be a function of both position u and wavelength λ. Consider an object (star or galaxy) i, at position u_i and with SED $\mathcal{S}_i$. The PSF affecting it is

$$\mathcal{H}_i = \int_B \mathcal{S}_i(\lambda)\mathcal{H}(u_i, \lambda)\mathrm{d}\lambda, \tag{5.1}$$

where B is the band of the instrument. In our case of *Euclid*'s VIS instrument, we have $B := [550, 900]\mathrm{nm}$. $\mathcal{H}_i$ is a 2D function of image coordinates, x_1, x_2. Using (2.88)'s notations, let I denote the true brightness profile of object i. Its PSF-convolved profile is

$$\begin{aligned} I^{\mathrm{obs}}(x_1, x_2) &= \int_B I(x_1, x_2)\mathcal{S}_i(\lambda) * \left[\mathcal{H}(u_i, \lambda)\right](x_1, x_2)\mathrm{d}\lambda & (5.2)\\ &= I(x_1, x_2) * \int_B \mathcal{S}_i(\lambda)\left[\mathcal{H}(u_i, \lambda)\right](x_1, x_2)\mathrm{d}\lambda & (5.3)\\ I^{\mathrm{obs}}(x_1, x_2) &= I(x_1, x_2) * \mathcal{H}_i(x_1, x_2). & (5.4) \end{aligned}$$

(5.4) is the polychromatic (and continuous) equivalent to the general PSF convolution setting of (2.103). All wavelength-dependent quantities are contained in $\mathcal{H}_i$. This required, to go from (5.2) to (5.3), the assumption that I does not vary with λ. In the case of an unresolved star, $I(x_1, x_2) := \mathbb{1}_{\{(x_1,x_2)=(x_1^c,x_2^c)\}}$, and this is true. For a galaxy, however, the true brightness profile is likely to be a function of wavelength. For instance, its bulge and disk are likely to have different spectra. In the general case, the effect of a chromatically varying brightness profile, $I(x_1, x_2, \lambda)$ (or, equivalently, of a spatially varying SED $\mathcal{S}_i(\lambda, x_1, x_2)$), is called *color gradient* (Semboloni et al., 2013). It can be a significant source of bias in shear measurements. Er et al. (2018) showed that it is possible to calibrate it to sufficient accuracy for *Euclid*, using HST data. This correction also requires an accurate knowledge of the PSF chromaticity, which needs to be acquired from stellar images. In this chapter, we focus on this task, and will therefore always assume (5.4) to be a valid assumption. Color gradient correction could then be carried out a posteriori, using our chromatic model as an input.

Similarly to the simplified setting of Section 3.3.1, our aim is to build a PSF model, $\hat{H}(u_{\mathrm{gal}}, \mathcal{S}_{\mathrm{gal}})$, for any set of galaxy position and SED, $(u_{\mathrm{gal}}, \mathcal{S}_{\mathrm{gal}})$. Available observations are star images, $Y_1, \ldots, Y_{n_{\mathrm{stars}}}$, with

$$Y_i := \mathcal{F}(\mathcal{H}_i) + N_i, \tag{5.5}$$

where, as in (3.29), $\mathcal{F}$ and N_i are the degradation operator (discrete sampling, subpixel shift and further undersampling) and Gaussian noise, respectively. $\mathcal{H}_i)$ is the true PSF affecting star i.

Since we want our PSF model to be computable for any SED $\mathcal{S}$, we will build it as a function of λ. In other words, we want to break down integrated, degraded observations into *monochromatic* PSF components $\left(\hat{H}(u, \lambda)\right)_\lambda$. We can then reconstruct the PSF that affected the object with

$$\hat{H}(u, \mathcal{S}) = \sum_\lambda \mathcal{S}(\lambda)\hat{H}(u, \lambda), \tag{5.6}$$

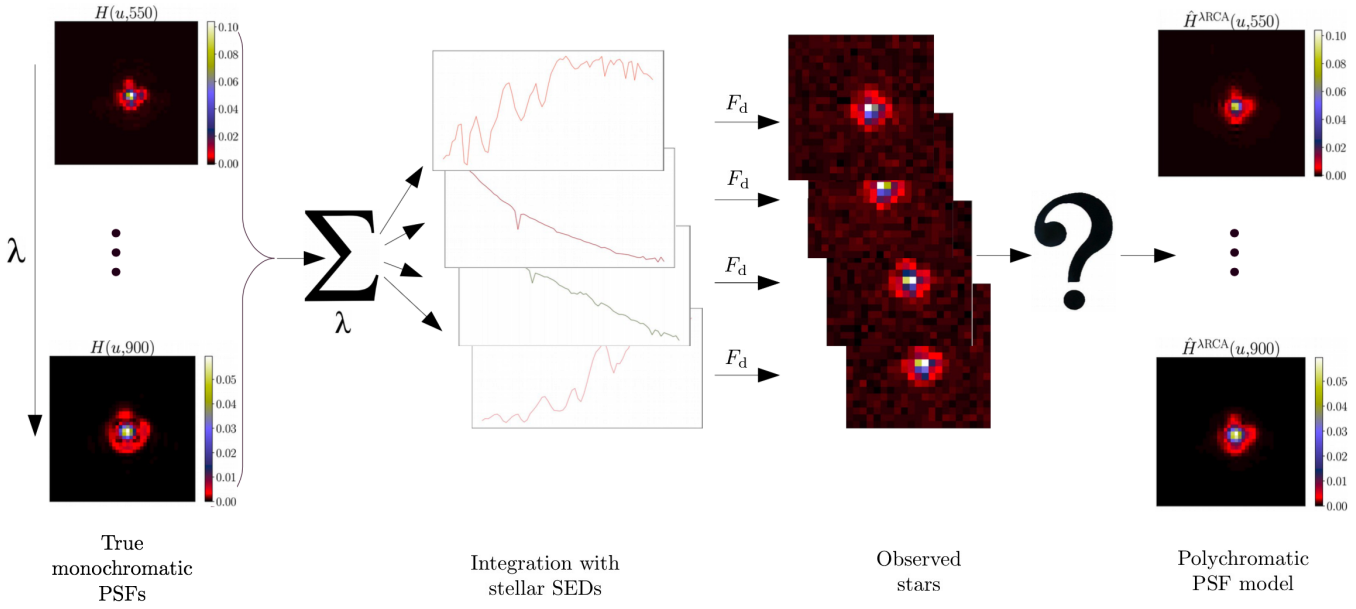


Figure 5.1: Illustration of the imaging process in the case of a chromatically varying PSF, and the associated polychromatic PSF estimation problem (indicated by the question mark).

where the summation is carried out over a discretization of B. This is an ill-posed inverse problem, as infinitely many $\hat{H}(u_i, \lambda)$ can reproduce the integrated star observation Y_i. The most trivial example is

$$\hat{H}(u_i, \lambda) = \frac{\bar{\mathcal{H}}_i}{\mathcal{S}(\lambda)}, \tag{5.7}$$

with $\bar{\mathcal{H}}_i$ the average integrated PSF across the whole band B. Such a model would entirely fail at recovering chromatic information.

This example shows the complexity of this problem; this is further illustrated in Figure 5.1. Even if we assume the stars' SED to be known (see Section 5.1.2 for a discussion of this issue), the chromatic information is highly degraded by the integration over λ. The only way to recover it is through the use of prior knowledge and regularization, as discussed in Section 3.1.2. We will build these using OT in Section 5.2.

5.1.2 SEDs

Knowledge of both stars' and galaxies' spectra are required. The former to build a chromatic PSF model, the latter to compute it for galaxies using (5.6). Photometric measurements, for each galaxy, from both *Euclid*'s own NISP instrument and supporting ground-based surveys will be made available for photo-z. These can also be used to obtain an estimate of $\mathcal{S}$ per object. This, however, raises the concern that the same measurements would be used for both PSF and redshift estimation. This in turn leads to correlations between errors in the two. Eriksen and Hoekstra (2018) explore this issue, and find that existing template-fitting methods would need to be modified to achieve *Euclid* accuracy goals. They propose a machine learning approach that reaches requirements, and corrects for potential correlations between the integrated galaxy PSF models and estimated redshifts. Interestingly, since their approach allows for a mapping between objects' colors and PSF size that is independent from the main PSF model, it can also be used as an extra validation of the latter. For their approach to be successful in recovering the effective galaxy PSFs (i.e. evaluating Equation 5.6), they do require accurate

knowledge of the chromatic variations of the PSF. Similar to color gradient corrections, the model proposed in this chapter would thus be fitted on stellar objects first, then provided as input.

In order to build a polychromatic PSF model, we also require per-object spectral information about stars. Kuntzer et al. (2016) showed that they could successfully classify individual stars into stellar classes using *only* their observation with VIS. This in turn allows us to obtain at least an SED template per star. More accurate spectral information will also be available for some objects, including from CSS-OS (Gong et al., 2019) and Gaia (Perryman et al., 2001). In the experiments of Section 5.3, we will only use stellar template SEDs, though we note our proposed model can only be improved by more precise spectral information on the observed stars. It can also, in principle, be fitted using stars with discrete SEDs of various accuracy simultaneously.

In what follows, we will always assume that for each observed star Y_i, we also have access to some measurement of its SED, $\mathcal{S}_i$. Moreover, in our experiments, the SED associated to each star is always exactly correct; the impact of a misclassification of stellar classes and/or inaccurate SED measurements on our method is left for further study.

5.2 λRCA

5.2.1 Rationale

In Section 4.5.2, we showed that the Wasserstein barycenter operator was a good choice to capture the chromatic variations of the VIS PSF. Our proposed polychromatic PSF model, λRCA, uses this information as a prior on the chromatic variations in order to expand the method of Chapter 3. In this model, the PSF was decomposed through the matrix factorization $\hat{H}_i := SA_i$, with S a set of eigenPSFs and A_i the weights for observation i. Star images could then be recovered by applying the degradation operator,

$$Y_i \approx F_{\mathrm{d}}(SA_i) = F_d\left(\sum_j a_{ij} S_j\right). \tag{5.8}$$

The central idea behind λRCA is to replace the set of eigenPSFs $(S_j)_j$ with a set of 2-atoms dictionaries $(D_j)_j := (d_j^-, d_j^+)_j$ containing, similar to those learned in Section 4.5.2, two extreme wavelengths eigenPSFs. Each monochromatic eigenPSF at any arbitrary wavelength λ can then be recovered as the Wasserstein barycenter of D_j with weights $T_\lambda := (t_\lambda, 1 - t_\lambda)$, as shown in Figure 4.8. Replacing S_j in (5.8), we get

$$Y_i \approx F_d\left(\sum_j a_{ij} \sum_\lambda \mathcal{S}_i(\lambda) P(D_j, T_\lambda)\right) = F_d\left(\sum_\lambda \mathcal{S}_i(\lambda) \sum_j a_{ij} P(D_j, T_\lambda)\right), \tag{5.9}$$

where P is the Wasserstein barycenter operator defined in (4.10). We will discuss the choice of T_λ in Section 5.2.3.

By learning both $\mathbf{D} := (D_j)_j$ and $A = (A_i)_{i\in\{1,\dots,n_{\mathrm{stars}}\}}$ from observed stars, we can then reconstruct the PSF at any position and for any SED. Spatial interpolation is carried out over A in the same way as in RCA (see Section 3.3.3), yielding $A_{\mathrm{gal}} = (a_{\mathrm{gal},j})_j$. "Spectral interpolation" is performed by using, in (5.6),

$$\hat{H}(u_{\mathrm{gal}}, \lambda) := \sum_j a_{\mathrm{gal},j} P(D_j, T_\lambda). \tag{5.10}$$

5.2.2 Learning $\mathbf{D}$

We have shown in Chapter 4 that the Wasserstein barycenter operator can recover the chromatic variations of the PSF. However, in that experiment, our WDL method was trained on monochromatic PSFs. The learned

atoms, d^-, d^+, were thus extremely close to the true extreme wavelengths PSFs (Figure 4.9). In the real-life setting of Figure 5.1, however, monochromatic PSFs are unavailable. It is therefore impossible to directly apply the idea presented above.

To illustrate this, let us start by simply replacing the eigenPSFs S of RCA by a set of Wasserstein barycenters $\mathbf{D}$, and the observation model (5.8) by (5.9). We would then fit our model by solving a minimization problem similar to (3.36) in the case of RCA, i.e.

$$\min_{\mathbf{D},A}\left(\frac{1}{2}\left\|Y - F_{\mathrm{d}}\left(\sum_{\lambda,j}\mathcal{S}_i(\lambda)a_{ij}P(D_j,T_\lambda)\right)\right\|_{\mathrm{F}}^2 + g(\mathbf{D},A)\right), \tag{5.11}$$

with g enforcing a set of constraints. If these were chosen to be the same as those of RCA (where applicable, see Section 5.2.4), with no extra constraints on the set of 2-atoms of $\mathbf{D}$, it would lead to convergence toward

$$\forall j, d_j^- = d_j^+ = S_j, \tag{5.12}$$

where S_j is the eigenPSF that would be obtained by running RCA on the same data. Since $P((S_j, S_j), \Lambda) = S_j$ for any set of weights Λ, we would revert to the same PSF model as that of monochromatic RCA. This is the exact same degeneracy as that illustrated by (5.7).

We thus need further constraints to ensure our prior knowledge (that the Wasserstein barycenter recovers chromatic variations) is active. We achieve this through two means: an adequate first guess, and a constraint on the *spread* of the atoms. Both come from the observation that, as λ increases, chromatic variations resemble a "zoom-in", as seen in Figure 4.7.

Since we will obtain our final model by alternate minimization (Section 3.1.6) over $\mathbf{D}$ and A, we expect to converge toward a local minimum. By selecting a first guess $\mathbf{D}^{(0)}$ motivated by the actual variations in our data, we can influence *which* local minimum will be reached. We build the first guess atoms, $\left(d_j^-\right)^{(0)}, \left(d_j^+\right)^{(0)}$, as follows:

1. we split the available stars into r groups (where r is the number of sets of dictionaries, or polychromatic eigenPSFs);
2. we perform a crude super-resolution step to obtain one superresolved star, Y_j^+, per group;
3. we apply a zoom-out operation to each Y_j^+, yielding Y_j^-;
4. we set $\left(d_j^-\right)^{(0)} := Y_j^+, \left(d_j^+\right)^{(0)} := Y_j^+$.

We tried several different ways to allocate stars to groups at step 1, and found it to have little influence. In practice, the stars are thus randomly allocated to groups, with replacement, as the number of stars used impacts upon the quality of the superresolution step. Step 2 was performed using the shift-and-add approach (Elad and Hel-Or, 2001). Lastly, the zoom-out operation of step 3 was carried out using an off-the-shelf method within `SciPy`[1] (Jones et al., 2001). An example of the resulting first guess atoms is shown in Figure 5.2.

We want to ensure the property of d^+ being more spread out than d^- is conserved throughout the iterations. While we could set one to be the zoom-out of the other at every iteration, and only optimize over the set of $(d_j^+)_j$, this would be too stringent a constraint. The true extreme wavelength PSFs clearly exhibit more intricate variations, as can be seen in Figure 4.9. Instead, if we impose the support of $\tilde{d}^-$, the Fourier transform of d^-, to be smaller than that of d^+, it will lead to the latter being more spread out in direct space. In practice, this is achieved through the addition of the following *spread constraint*:

[1] `https://docs.scipy.org/doc/scipy/reference/generated/scipy.ndimage.zoom.html`

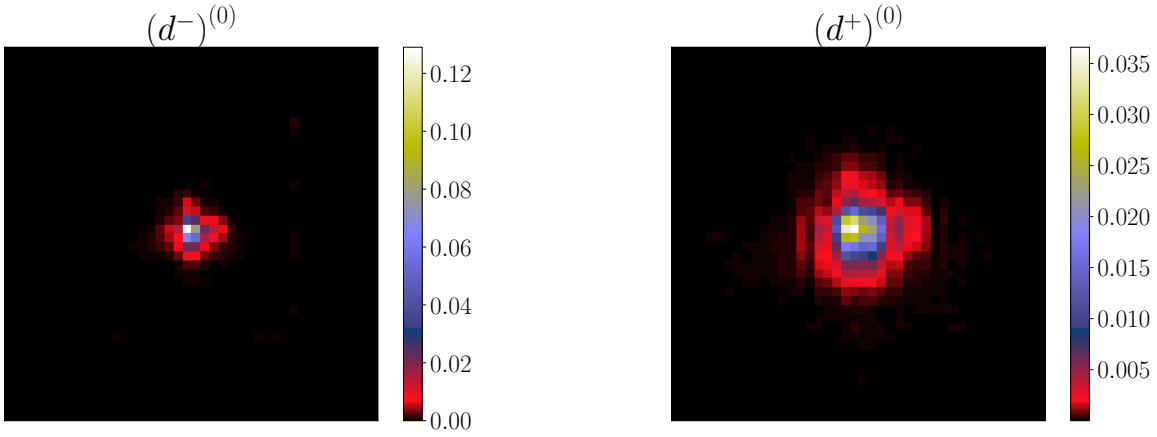


Figure 5.2: First-guess dictionary atoms for an arbitrary component j. Each of the other $r-1$ initializations appear very similar.

$$g^{\text{spread}}(\mathbf{D}) := \beta^{\text{spread}} \sum_j \|w_j^{\text{spread}} \odot \tilde{d_j^-}\|_1, \tag{5.13}$$

where β^{spread} is a trade-off parameter weighing the relative importance of this constraint compared to others and to the data fidelity term (similar to λ in Equation 3.11), and w_j^{spread} are weights that determine the amount of sparsity desired.

These are two new hyperparameters that could in principle be set to any value. We can, however, provide the following empirical considerations. In practice, it is beneficial to have β^{spread} vanish with increasing iterations. Indeed, as $\mathbf{D}$ approaches convergence, the spread constraint becomes less needed. Lowering its relative importance further allows for our final $\tilde{d}^-$ to include low-frequency variations that would have otherwise been removed by thresholding. Since we want the Fourier transform of $\tilde{d}^-$ to be sparser than that of $\tilde{d}^+$, w_j^{spread} should be related to $\|\tilde{d}^+\|_1$. We found

$$w_j^{\text{spread}} := \frac{1}{2p^2}\|\tilde{d}^+\|_1, \tag{5.14}$$

where p^2 is the number of pixels in the superresolved PSF images, to work well in practice for the VIS PSF. The amount of extra spread of the largest wavelength PSF, when comparing it to the smallest, is related to how the size of the PSF changes with wavelength. This relationship is often modelled as a power law. It is reasonable to assume that some knowledge about its parameter should be available for any instrument's PSF (see, e.g. Carlsten et al., 2018; Eriksen and Hoekstra, 2018, for that of HSC and VIS, respectively), in turn giving an empirical mean to select w_j^{spread} for any survey.

We note that both of these choices lead to a change in the cost function that is minimized. However, since fitting λRCA amounts to an alternated minimization, this is acceptable as long as overall convergence is achieved in practice. Indeed, each minimization over $\mathbf{D}$ is already carried out with a different cost function regardless, since A is updated in between each of them.

5.2.3 Choosing t_λ

The other new parameters arising from the introduction of OT in our PSF modelling are the barycentric weights, T_λ (that were called Λ in the generic case of Chapter 4). Since we only consider 2-atoms dictionaries in the

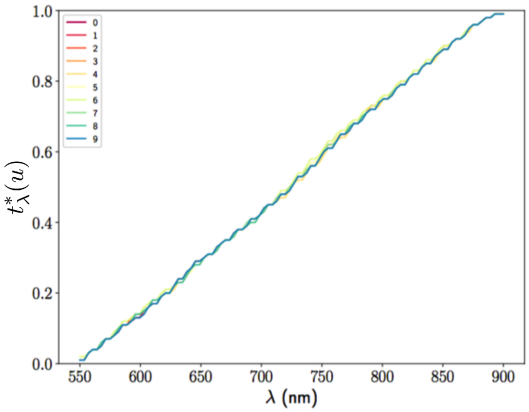


Figure 5.3: Optimal barycentric weight $t^*_\lambda(u)$, for 10 positions u, each shown with a different color.

present chapter, there is only one scalar paramater, t_λ, to choose. The other barycentric weight will then necessarily be $1 - t_\lambda$. t_λ corresponds to the relative weight given to d^- (compared to d^+) at any wavelength λ. For the extreme wavelengths of our band B, the choice is naturally

$$t_{\lambda_{\min}} := 1, \quad t_{\lambda_{\max}} := 0. \tag{5.15}$$

For all intermediary wavelengths, t_λ can either be added as a free parameter to optimize over when fitting λRCA to stars, or as a predetermined function of λ.

The second option is preferable as it reduces the number of free parameters. Our experiment of Section 4.5.2 provided us with some early insight into the relationship between λ and the Wasserstein barycentric weights between two "extreme wavelength" atoms. There, t_λ was initialized as a projection of $B = [\lambda_{\min}, \lambda_{\max}]$ into $[0, 1]$. While still allowed to vary, it remained fairly close to linear upon convergence of WDL, as shown in Figure 4.11b. The present setting is different, since we are not running WDL on a set of true monochromatic PSFs, though this result is an indication that a linear relationship between λ and t_λ might be a good choice.

If our model succeeds at capturing chromatic variations, our set of polychromatic eigenPSFs should reproduce the true extreme wavelength PSFs, $H^{\text{ext}}(u) := (H(u, \lambda_{\min}), H(u, \lambda_{\max}))$. To explore the relationship between λ and an optimal t_λ, we thus compute a set of Wasserstein barycenters,

$$P_t(u) := P(H^{\text{ext}}(u), (t, 1 - t)), \tag{5.16}$$

on a finely sampled grid $t \in [0, 1]$, for 10 different positions u in the FOV. For each intermediary wavelength λ, we select the optimal t as

$$t^*_\lambda(u) := \underset{t}{\operatorname{argmin}} \, \|H(u, \lambda) - P_t(u)\|^2_{\text{F}}. \tag{5.17}$$

That is, the barycenter of the true extreme wavelength PSFs that achieves the lowest pixel reconstruction error when compared with the true monochromatic PSF. These are plotted in Figure 5.3, and show excellent agreement with a linear relationship. We will therefore always set

$$t_\lambda := \frac{\lambda - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}} \tag{5.18}$$

for all components j.

5.2.4 Other constraints and optimization problem

Now that we have addressed the new variables and parameters of λRCA, we turn to the rest of the constraints of RCA described in Section 3.3.2. The "low rank", or dimensionality reduction aspect is preserved through the choice of r, the number of components. The only difference is these are now a set of dictionaries (D_j), instead of a set of single eigenPSF images $(S_j)_j$. Similarly, we still want to enforce sparsity of these components in an appropriate representation space (Starlets, in our case). This is achieved by preserving the sparsity constraint, but applying it to both elements of each D_j instead of S_j.

The graph constraint described in Sections 3.3.2 and 3.3.3 is directly applicable to λRCA and remains unchanged. We build the set of PSF graphs at initialization as in (3.33), factorize A by the last few eigenvectors of their "Laplacian matrices", and optimize over the resulting α coefficients instead of A. The graph constraint is still enforced through a ι_Ω term in the cost function, with Ω defined in (3.38).

As was the case for the dictionary of WDL in Section 4.2, we will now require all (d_j^+, d_j^-) to belong to the probability simplex Σ_{p^2}, defined in (4.1). This can be achieved by two means: either using the logistic change of variable (as we did in Chapter 4), or by adding $\iota_{\Sigma_{p^2}}$ terms for each d_j^+ and d_j^- to our cost function. Note that Σ_{p^2} is a convex set, and $\iota_{\Sigma_{p^2}}$ is thus proximable. The experiments shown in this chapter used the latter option. Either way, a side effect is that all elements of $\mathbf{D}$ are positive. This does not guarantee the positivity of the reconstructed PSFs, as A can still contain negative values. However, in practice, we found that the new simplex constraint was sufficient to obtain a positive-valued PSF model. We thus remove the positivity constraint that was present in regular RCA (enforced through the $\iota_+(S\alpha V^T)$ term in Equation 3.36). Note that a similar observation was already made in both Ngolè et al. (2016) and Section 3.3: when performing the alternated optimization of RCA, it was sufficient to consider the positivity constraint when optimizing over S. It was already discarded when optimizing over α.

Altogether, these considerations give the following optimization problem:

$$\begin{aligned}\min_{\mathbf{D},\alpha} \Bigg(& \frac{1}{2}\left\| Y - F_{\mathrm{d}}\left(\sum_{\lambda,j} \mathcal{S}_i(\lambda) a_{ij} P^{(L)}(D_j, T_\lambda)\right)\right\|_{\mathrm{F}}^2 \\ & + \sum_{j=1}^{r}\left[\|w_j^- \odot \Phi d_j^-\|_1 + \|w_j^+ \odot \Phi d_j^+\|_1\right] + \sum_j \left[\iota_{\Sigma_{p^2}}(d_j^+) + \iota_{\Sigma_{p^2}}(d_j^-)\right] + g^{\mathrm{spread}}(\mathbf{D}) \\ & + \iota_\Omega(\alpha)\Bigg).\end{aligned} \tag{5.19}$$

The first line is the data fidelity term. The second contains all constraints applied to $\mathbf{D}$: sparsity in Starlet domain, need to belong to the simplex, and spread constraint (defined in Equation 5.13). The third is the graph constraint applied to α.

Notice we replaced the Wasserstein barycenter operator P with its entropy penalized approximation $P^{(L)}$, as in Section 4.2.2. As in the WDL method of Chapter 4, this will allow us to compute the gradient of our cost function with regard to $\mathbf{D}$ through automatic differentiation.

As with regular RCA, fitting λRCA to a star field amounts to solving (5.19). This is achieved through alternatively optimizing over $\mathbf{D}$ and α using proximal algorithms (Condat, 2013; Raguet et al., 2013). A python library to perform this fitting is currently under development[2]. It includes a wrapper of C functions[3], originally developed by my collaborator Matthieu Heitz for the log-stabilized implementation of WDL described in Section 4.4.1, heavily modified by Rebeca Araripe to fit the needs of λRCA.

[2] https://github.com/CosmoStat/LambdaRCA
[3] https://github.com/matthieuheitz/WassersteinDictionaryLearning

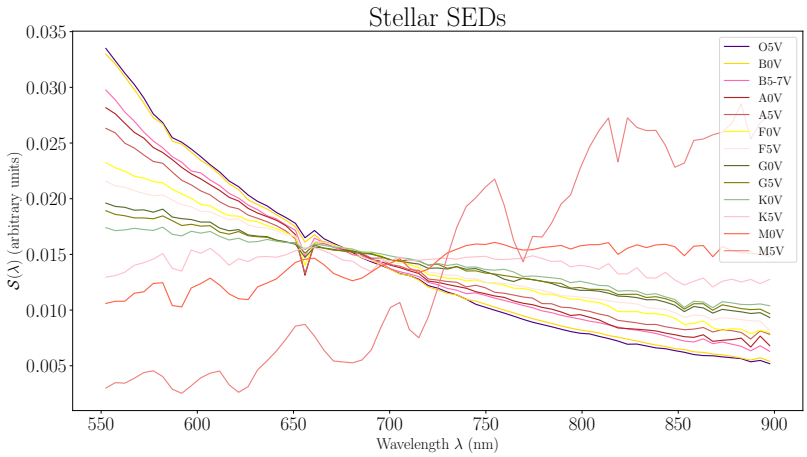


Figure 5.4: Stellar SEDs used in our experiments.

5.3 Experiments

In order to test our proposed approach, we compare its performance with that of regular RCA on the experiment described in this section.

The PSFs used as ground truth are the same VIS simulations as those used in Section 3.3 and the PSF experiment of Section 4.5.2, described in more detail in Ngolè et al. (2015). Future experiments in the non-parametric Euclid PSF modelling efforts will, however, use a new set of simulations provided by Pierre-Antoine Frugier. These were custom made for this purpose, are based on more up-to-date specifications of the flight instrument, include modelling of polishing errors that lead to high-frequency variation, and the impact of CCD misalignment. Because of this last addition, the underlying PSF model is discontinuous at the border of each CCD. Conversely, no CCD misalignment is present in the dataset used in this thesis. The experimental set-up of Section 3.3 used $n_{\text{stars}} = 300$, which is higher than the expected number of usable PSF stars per CCD, but lower than the number across the whole field. In preparation for the new dataset, and future efforts to add multi-CCD handling to our non-parametric models, we will use $n_{\text{stars}} = 80$ in the present chapter. This number corresponds to the average expected number of PSF stars, per VIS CCD, on science exposures.

For each position considered, we save a set of "true" PSFs, $(H(u, \lambda))_\lambda$, sampled at the target pixel scale. As in Section 3.3, that target is half of *Euclid*'s detector pixel scale. Resulting PSF images are then truncated to stamps of 42×42 pixels. These correspond, similar to H^{kn} in Chapter 3, to the best reconstruction achievable by our models, and will be used for comparison and to compute PSF model quality criteria. The PSF simulations contain 100 monochromatic PSFs, with a spectral sampling of $\Delta\lambda = 5\text{nm}$. These, however, contain some out-of-band simulations that we will not consider in the present work. Restraining ourselves to the nominal VIS band of $B = [\lambda_{\min}, \lambda_{\max}] = [550, 900]\text{nm}$, we obtain 71 monochromatic PSFs per position.

5.3.1 Observed stars

80 positions $\mathcal{U}_{\text{stars}}$ are randomly drawn across the FOV. One of 13 stellar SED is assigned to each of these. Following Kuntzer et al. (2016), these are chosen to be templates of main sequence stars, spanning the whole Morgan-Keenan range (Morgan and Keenan, 1973), from the Pickles (1998) library. These are shown in Figure 5.4. The SED assigned to any position is uniformly drawn.

The monochromatic PSFs are then summed at the smallest pixel scale available, weighted by the normalized selected stellar SED. These are sampled on the VIS band with a spectral sampling step of $\Delta\lambda = 5\text{nm}$, to match that of the PSF dataset. The resulting integrated stellar PSFs are then sampled at the target pixel scale, and the degradation operator F_{d} containing both a sub-pixel shift (as illustrated in Figure 2.13) and the

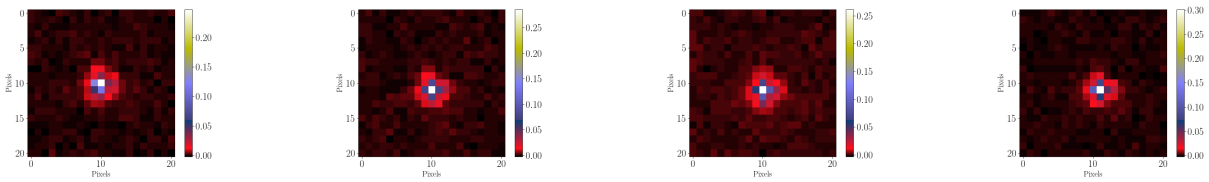


Figure 5.5: Example observed stars. Each corresponds to both a different position in the FOV *and* a different stellar SED.

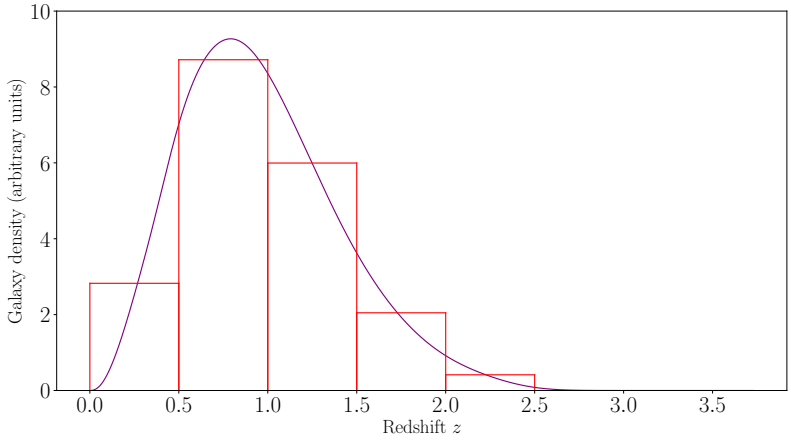


Figure 5.6: *Euclid* lensing $n(z)$ (purple line) and the corresponding PDF for each of our 5 redshift bins (red bars).

further degradation to *Euclid* pixel scale is applied. Stamps of $p^2 := 21 \times 21$ pixels are then extracted, and white Gaussian noise is added, yielding 80 star observations $(Y_i)_{i\in\{1,\dots,80\}}$. Since we saw no significant trend with noise levels in the experiments of Chapter 3, we consider a single noise level of $\sigma = 0.03$, applied after normalization of the star stamps. Examples of a few resulting star images are shown in Figure 5.5.

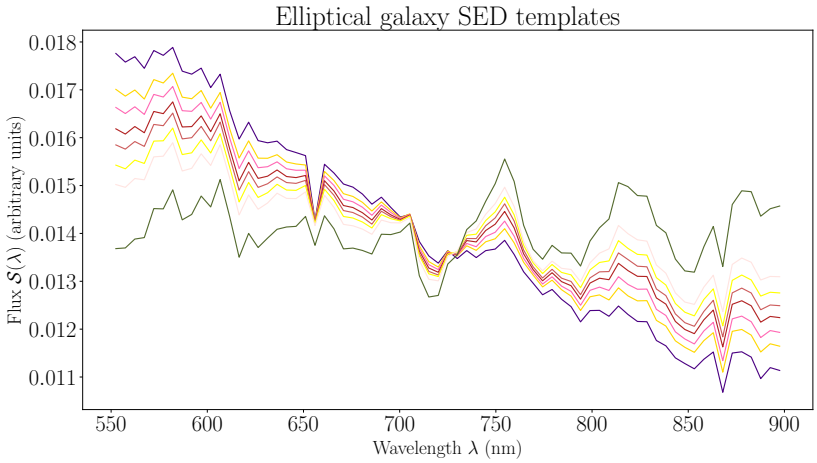
5.3.2 Test/galaxy PSFs

80 other positions are drawn from those with available PSF simulations. We will refer to these as "test" or galaxy positions.

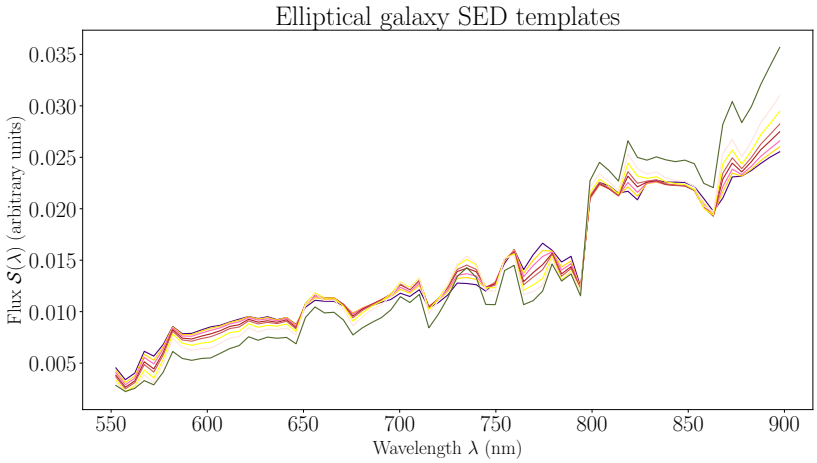
The galaxy SEDs we use are those of Ilbert et al. (2008) and were provided by my cosupervisor, Martin Kilbinger, following exchanges with Francisco Castander[4]. These contain 31 different template SEDs, corresponding to 7 elliptical, 1 lenticular, 11 spiral, and 12 starburst galaxy types. Each of these include extinction by the galaxy itself, selected from 5 different possible extinction laws, but no extinction by the Milky Way.

For each of $31 \times 80 = 2480$ unique combinations of position and galactic SED template, we draw a redshift value from 5 very wide bins of $\Delta z = 0.5$, with probabilities matching the expected lensing $n(z)$ of the Euclid Wide Survey (Euclid Collaboration et al., 2019) as shown in Figure 5.6. Each SED is redshifted accordingly, and sampled with the correct $\Delta\lambda$ within the VIS band. Examples for two redshift bins and the elliptical galaxy SED templates are shown in Figure 5.7. The final, integrated test PSF are then computed by summation at the finest available pixel scale, then sampled at our target pixel scale and stamp size of 42 pixels. This results in a set of 2480 test PSFs $\left(H(u, \mathcal{S}_j)\right)_{u\in\mathcal{U}_{\text{gal}}, j\in\{1,\dots,31\}}$.

[4]Readers with access to CEA's *Direction de la Recherche Fondamentale* private repositories can find them at `https://drf-gitlab.cea.fr/mkilbing/SED`.



(a) Redshift bin centered on $z = 0.25$.



(b) Redshift bin centered on $z = 1.25$.

Figure 5.7: Example galaxy SEDs for Elliptical galaxies for two different redshift bins. Each color corresponds to one of 7 templates for elliptical galaxies. Similar plots for all considered galaxy SEDs are shown in Appendix I.

5.3.3 PSF models

We run both RCA and λRCA on the observed stars. Since RCA does not handle chromatic variations, its PSF model is simply interpolated to each galaxy position in $\mathcal{U}_{\text{gal}}$, yielding 80 different $\hat{H}^{\text{RCA}}$ estimated PSF. Each will be used to compute diagnostics by comparing it with each of the 31 different integrated test PSFs corresponding to different SEDs.

λRCA uses the true individual stellar SEDs, $(\mathcal{S}_i)_i$, along with the corresponding observed stars and positions, $(Y_i, u_i)_i$, as inputs. The spectral sampling rate $\Delta\lambda$ of these is the same as that of our simulated PSFs. However, in a real-life setting, this sampling rate would depend on the source of our SED information. The number of individual wavelengths at which the Wasserstein barycenters of each of our D_j are computed is another free parameter. In other words, this parameter determines the set over which the summation in λ is carried out in (5.9). The higher this number is, the more chromatic information is used, up to the limit imposed by the quality of our SED estimates. However, it also comes at increasingly higher computational costs. In the results shown in this chapter, we chose to compute barycenters at 12 different, evenly spaced individual wavelengths. This value was chosen because of the small improvement in reconstruction error obtained when adding extra wavelengths values.

For both methods, we learn $r = 4$ components. These are either eigenPSFs S_j for RCA, or 2-atoms dictionaries D_j for λRCA. Other hyperparameters are set as suggested in Chapter 3 and Section 5.2. The addition of the entropic penalty in the computation of the Wasserstein barycenter operator in (5.19) adds γ and L as hyperparameters to λRCA. These are the entropic penalty trade-off parameter and number of Sinkhorn iterations, respectively. The role of these two is described in Chapter 4, with a graphical representation of the influence of γ in Figure 4.3.

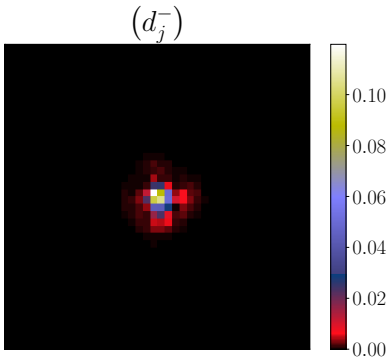
Once again, empirical ways to determine adequate values for these two parameters can easily be proposed. While high γ values can lead to a denoising effect (Cuturi and Peyré, 2016), in our case, the noise is already handled by the sparsity constraint. On the other hand, we want our PSF model to recover sharp features, which requires low values of γ. The exact value can be chosen empirically by computing Wasserstein barycenters of images expected to contain features of the same sharpness as the desired PSF model, and ensuring these are preserved within the barycenters. In our case, γ could be chosen by computing Wasserstein barycenters of the first guess, lowest wavelengths dictionary atoms, $(d_j^-)^{(0)}$, and ensuring they do not exhibit excessive smoothing. Once γ is chosen, L should simply be chosen to be the smallest possible, while still allowing for convergent Sinkhorn iterations. In our case, these empirical recipes led to our choosing $\gamma = 0.3$ and $L = 10$. These values caused a need for the log-domain stabilization of Sinkhorn iterations we introduced in Section 4.4.1.

Every element of $\mathbf{D}$ upon convergence of the λRCA fitting process are shown in Figure 5.8. The features they exhibit clearly show that the algorithm was successful in capturing chromatic variations. We emphasize that, unlike the experiment of Section 4.5.2, these atoms were obtained using only *undersampled*, *single-band* images of stars, like those shown in Figure 5.5.

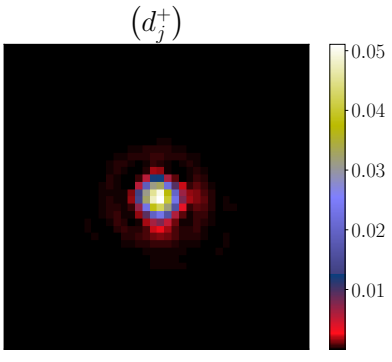
5.4 Results

5.4.1 Monochromatic PSFs

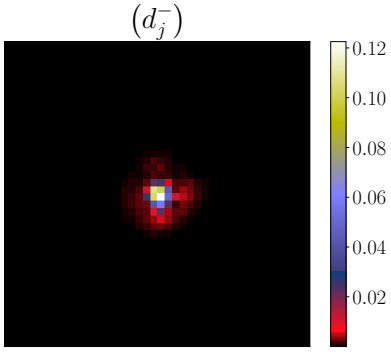
We first study the recovery of individual monochromatic PSFs by λRCA. No comparison with RCA (or any other non-parametric PSF model) is possible at this stage, since λRCA is the only one able to recover these. Figure 5.9 show the true and reconstructed PSF, for a range of evenly spaced λ values, for a single randomly drawn position. The normalized pixel mean square error (MSE) per wavelength is shown in Figure 5.10. Monochromatic profiles are recovered to few-percent accuracy, with a decrease in quality, and wider spread of errors across positions, in the few lowest wavelengths.



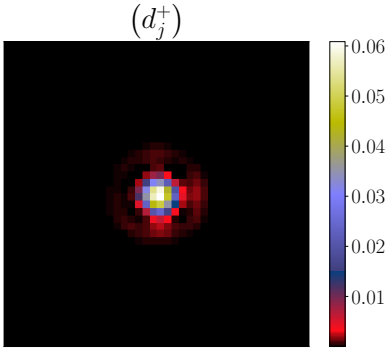
(a) $j = 1$.



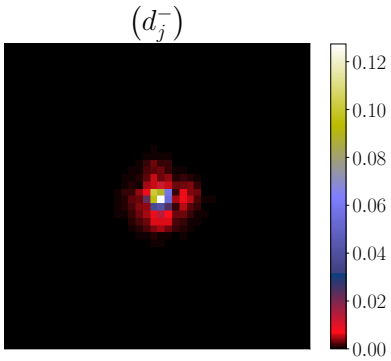
(b) $j = 1$.



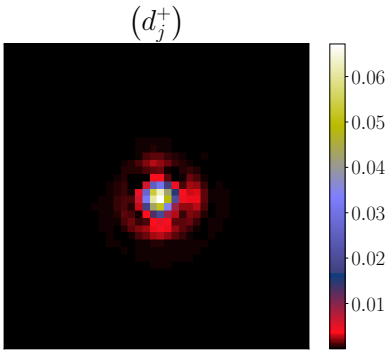
(c) $j = 2$.



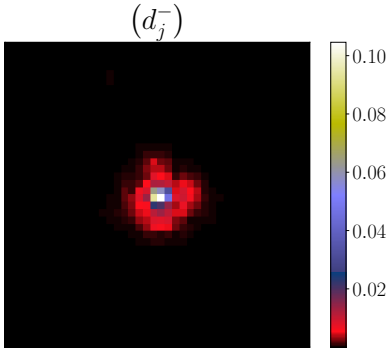
(d) $j = 2$.



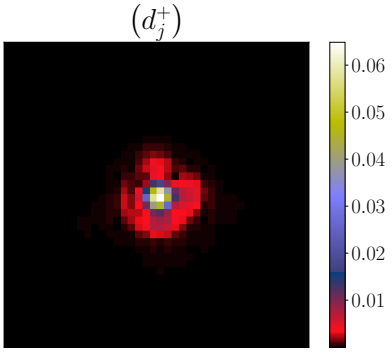
(e) $j = 3$.



(f) $j = 3$.

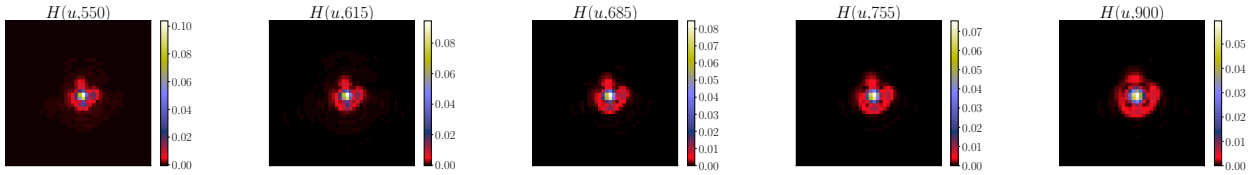


(g) $j = 4$.

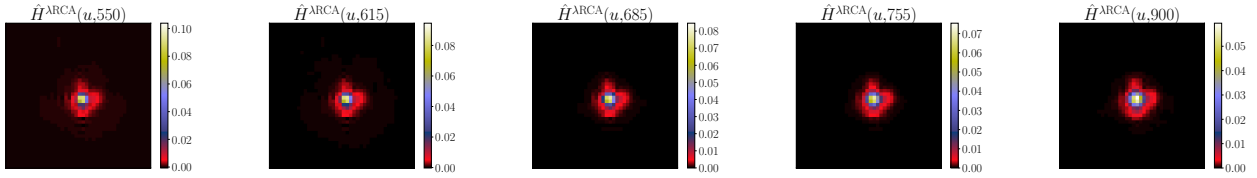


(h) $j = 4$.

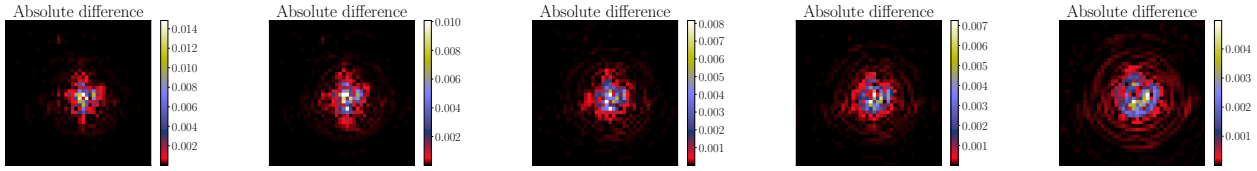
Figure 5.8: λRCA-learned dictionary atoms $\mathbf{D}$. The lowest (resp. highest) wavelength atom for each component j are shown on the left (resp. right).



(a) True monochromatic PSFs.



(b) Reconstructions from λRCA.



(c) Absolute error.

Figure 5.9: Selected monochromatic PSFs at individual wavelengths across the whole VIS band for a randomly drawn position. Row (a) show the true PSFs, row (b) their reconstruction using the λRCA model, and row (c) the absolute error between the two.

5.4.2 Test PSFs

In this section, we compare the quality of estimated galaxy PSFs. A visual example is shown for both PSF models in Figure 5.11. We further compute a set of quality criteria, for each group of galaxy SED templates, and show the results in Table 5.1. These include the pixel MSE, normalized by the l_2 norm of the true integrated test PSFs, as well as RMS errors on both ellipticity components and size. As discussed in Section 2.4.1, these are the quantities used by Cropper et al. (2013) to set requirements on PSF model quality for given WL science objectives, and was in turn used to set those for *Euclid*. Our experiments of Chapter 3 showed that the Paulin-Henriksson et al. (2008) formalism, which led to the choice of these quantities, is no longer valid for a diffraction-limited PSF. Nonetheless, it remains a useful quantification of a PSF model quality and allows for comparison with the only currently established *Euclid* requirements. Most importantly, even if our results showed these are not the *only* PSF quantities having an impact on shear estimation, they still remain among the most important contributors.

As can be seen in Figures I.2 and I.3 of the Appendices, the SED for Starburst galaxies can vary significantly from one template to the next, especially at higher redshifts. We thus split them into two groups with similar trends in high-redshift SEDs in Table 5.1. Using λRCA leads to smaller errors in pixels (as could be seen in the example of Figure 5.11c) across all types of galaxies. The error on the size of the integrated galaxy PSFs is also improved, by over a factor of 2, compared to the case where chromatic variations are not accounted for. In both cases, the amplitude of errors in the RCA model vary by over 20% from one galaxy group to another, while λRCA errors show less variations.

Looking at the RMS on ellipticity, however, we see no significant difference in performance in the second component, while λRCA actually performs slightly *worse* than RCA in the first component. A possibility is that this is due to λRCA being less efficient in capturing the spatial variations in e_1. Indeed, the graph constraint is built using the positions of stars, regardless of their SEDs. In the case of λRCA, this leads to a mixing between recovered chromatic and spatial information. If the former has more of an impact on the updates to $\mathbf{D}$

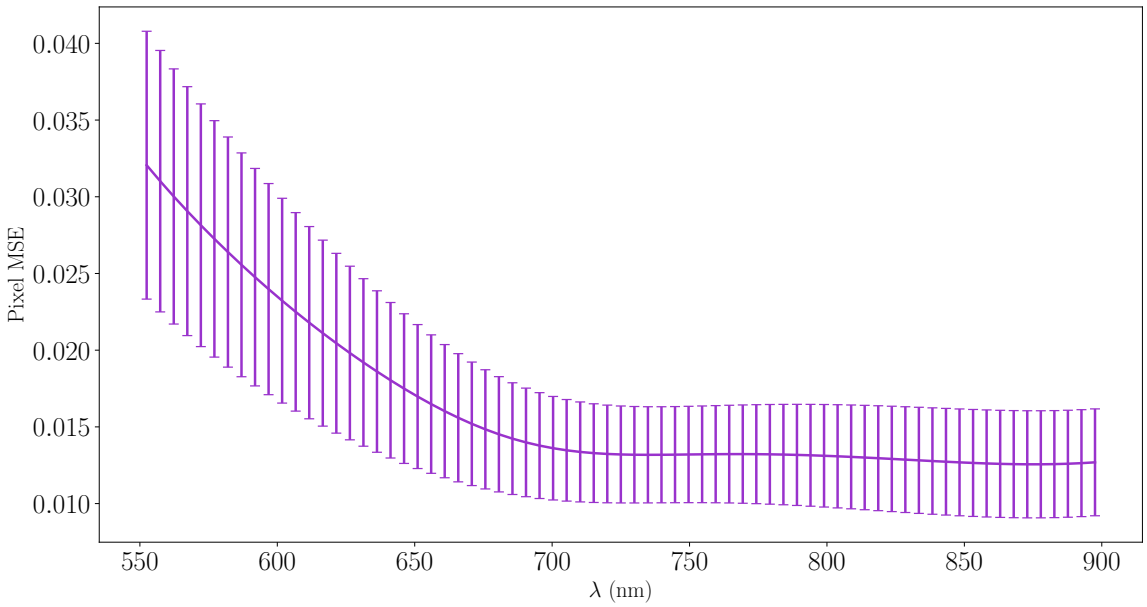


Figure 5.10: Normalized pixel MSE as a function of wavelength. Error bars show variations across all 80 star positions.

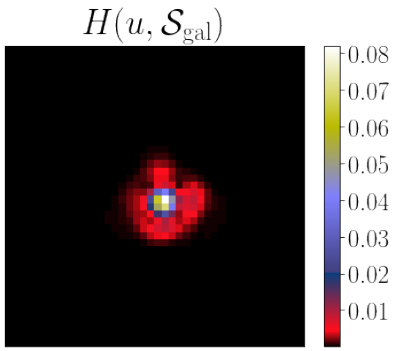
Galaxy type	Normalized pixel MSE		$\sigma(e_1)$		$\sigma(e_2)$		$\sigma(R^2)/R^2$	
	RCA	λRCA	RCA	λRCA	RCA	λRCA	RCA	λRCA
Elliptical	0.016	**0.0095**	**0.018**	0.021	0.0050	**0.0046**	0.32	**0.15**
Spiral & lenticular	0.016	**0.0095**	**0.018**	0.021	0.0051	**0.0047**	0.32	**0.15**
1st group Starburst	0.017	**0.0096**	**0.018**	0.023	0.0053	**0.0051**	0.35	**0.14**
2nd group Starburst	0.020	**0.010**	**0.020**	0.027	0.0059	0.0059	0.40	**0.13**

Table 5.1: Summary of quality criteria for both PSF models across all galaxy types. Here, σ denotes the RMS error.

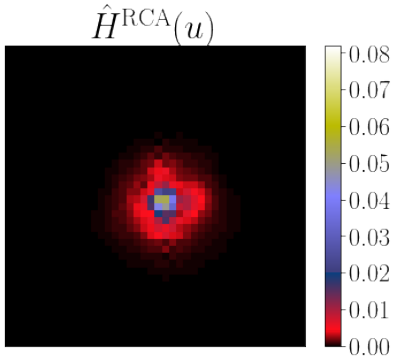
across iterations, the spatial information contained in the graph constraint is drowned out as the components "specialize" in capturing the chromatic variations. Conversely, since RCA ignores chromatic information entirely, the graph constraint has a greater impact on updates to the eigenPSFs. If this is indeed the reason for the difference in recovery of the first ellipticity component between the two models, it could be alleviated by learning both chromatic components D_j *and* achromatic eigenPSFs S_j simultaneously within λRCA. The former would still capture the chromatic variations, as is clearly the case with the components we found here (as can be seen in Figure 5.8). The latter could then leverage the graph constraint to capture remaining variations driven by the position of objects.

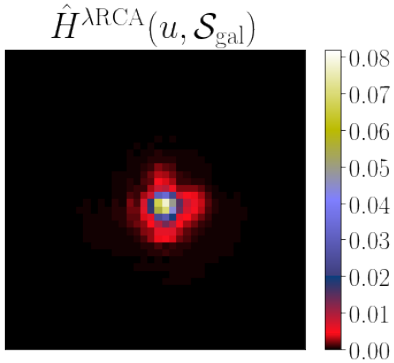
5.5 Discussion

In this chapter, we introduced λRCA, the first non-parametric PSF model capable of capturing chromatic variations. Using no information other than observed undersampled stars and estimates of their SED, we manage to capture the variations undergone by the PSF as the wavelength varies, as illustrated by the appearance of the learned components in Figure 5.8 and the reconstructed monochromatic PSFs at single wavelengths in Figure 5.9.

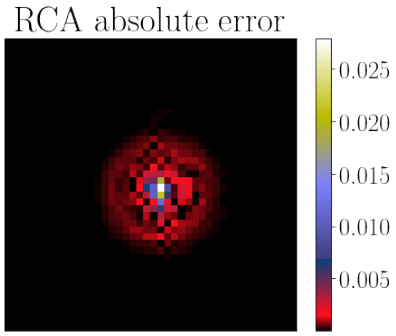


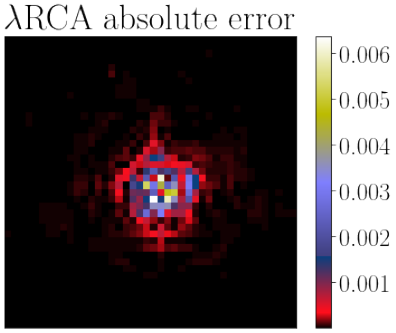
(a) True PSF integrated with a galaxy SED.





(b) Corresponding reconstructions from both PSF models.





(c) Absolute errors.

Figure 5.11: Example test PSF, reconstructions using regular RCA (i.e. neglecting chromatic variations) and λRCA, and the corresponding error maps.

Estimating the PSF that would affect galaxies with a wide range of SEDs and several different redshifts, we see a significant improvement using this proposed approach in both the recovered profile of the PSF and its size. Ignoring the chromatic variations of the PSF naturally leads to a strong correlation between the galaxy type and the PSF modelling errors. Using our chromatic model leads to more stable PSF errors across different galaxy types.

The setting of our experiments in this chapter is more realistic than that of Section 3.3, as it contains polychromatic PSFs and is estimated using the expected number of stars on a single CCD. The quality of our PSF models in terms of ellipticity is of the same order of magnitude as that observed in the previous, monochromatic setting. Interestingly, however, using our chromatic model leads to a worse RMS in the first ellipticity component.

This might be alleviated by learning both λ-dependent and achromatic components simultaneously.

Another significant difference with the results of Chapter 3 is in the RMS on the size of the PSF, which is improved by more than an order of magnitude. Using λRCA, we are now at a factor of a few 10^2 over both shape requirements, while in the earlier experiment, we found a factor of about 10^4 in size. This is not only due to the improvement to the PSF model obtained by going from achromatic RCA to λRCA, but also because integrated PSFs naturally have larger profiles, which in turn makes the superresolution easier.

Now that we have a PSF model that can capture chromatic variations, the main limitation is the number of datapoints used to fit it. As we already concluded at the end of Chapter 3, our non-parametric approach will only be able to achieve *Euclid* requirements if we can somehow fit it to more stars simultaneously. A first and fairly straightforward way to do so is to include a modelling of the discontinuities caused by the jump from one CCD to the next in the model. This will allow us to fit it to the whole field simultaneously instead of on a per-CCD basis. As shown in Appendix B, this can already lead to significant improvement of PSF model quality. An approach similar to that used by the `LensFit` PSF model, described in Equation 2.4.3, could be included into λRCA. While such an approach could not be tested with the data used in this thesis, the new simulated Euclid PSFs we mentioned in Section 5.3 do contain realistic CCD discontinuities. In the case of VIS, performing cross-CCD fits would already allow for 36 times more stars (on average) to be used. If this still proves insufficient to bridge the gap between the performance showed in this chapter and *Euclid* requirements, we will need to use stars from several different exposures.

A current caveat of λRCA is its computation time. For the experiment shown in this chapter, fitting the model to a field of 80 stars took a little over an hour on a 4-core computer. The main contribution to computation time is the repeated computations of the Wasserstein barycenter operator, $P^{(L)}$, and especially of its derivative with regards to the dictionaries. Since we need a very accurate recovery of the PSF profiles, we have to use small values for the entropic regularization parameter γ. This in turn forces us to use the log stabilized version of the backward Sinkhorn loop we presented in Section 4.4.1. Despite our efforts in last chapter to speed up this computation, and our use of a parallelized C code, this step remains the main bottleneck in the fitting of λRCA. The same was true for our WDL approach, and we suggested some alternatives to speed up the process at the end of Chapter 4, namely the Greenkhorn algorithm (Altschuler et al., 2017) and Schmitzer (2019)'s multiscale approach. Another option applicable to λRCA is offered by Courty et al. (2017): using deep neural networks, they learn the Wasserstein barycenter operator P. Once such a network is trained, computation of P becomes extremely fast. Moreover, they show these networks are particularly good at generalization; that is, despite being trained using very different data, they still manage to compute excellent approximations of the true Wasserstein barycenters. Including their approach in λRCA could lead to drastic improvements in its computational cost.

Lastly, let us mention a few effects that could have an impact our polychromatic Euclid PSF model. Uncertainty in stellar SEDs was already discussed in Section 5.1.2. Emission lines in galaxy spectra might also lead to large errors in the estimated effective PSF, even if the PSF modelling step presented in this chapter perfectly captured chromatic variations. Out-of-band leakage could also impact the chromatic PSF modelling step.

Most importantly, the central assumption on which λRCA is built is that OT tools are able to capture the true chromatic variations. While we have shown this to be true with the present dataset, very recent measurements made on *Euclid* hardware indicate the chromatic variations of the PSF may not be quite as smooth as expected. If non-smooth variations, localized in λ, occur, these would need to be taken into account separately (or on top of) the OT-based recovery of λRCA.

Chapter 6

Conclusion

This thesis presents a study of non-parametric approaches to PSF modelling for WL surveys. Several upcoming surveys, including *Euclid*, will soon start their observations. They will bring a wealth of cosmological information from WL surveys of unprecedented breadth. This will only be possible, however, if we are able to correctly handle the main source of error in galaxy shape measurement: the PSF.

Main contributions

Historically, non-parametric approaches to model the PSF have proven to be more efficient than physically-motivated ones. In the case of *Euclid*, however, no existing model can be directly applied. Indeed, the PSF of *Euclid*'s VIS instrument comes with new modelling challenges that the WL community had never been faced with before. Chief among those is the chromatic variation of the PSF over the very broad band of VIS.

In this thesis, I have improved upon existing PSF models, and devised a new approach to capture these chromatic variations. Achieving this in a non-parametric framework is an especially hard task, since the information to be recovered is highly degraded. Not only is the PSF affecting objects integrated, with their own SED, over the whole band, these broadband observations are also *undersampled*. Recovering a polychromatic PSF model from such observations amounts to solving an ill-posed *inverse problem*.

Advanced mathematical tools were deployed to solve this problem. To tackle the issues of noise, undersampling, and spatial variations of the PSF, I have made use of sparse signal processing. These methods were used to improve upon a recently proposed PSF modelling approach in Chapter 3, and extensively tested using image simulations.

In order to break down SED-integrated, undersampled star images into a set of monochromatic PSFs spanning the whole band, I have turned to Numerical OT. Following recent developments in this fairly new field, numerical tools derived from the theory of optimal transportation can now be readily computed and applied to practical problems. In particular, the *Wasserstein distance*, that describes the effort needed to "transport" a probability measure into another, has already proven to be an extremely potent tool in machine learning applications. In this vein, in Chapter 4, I have presented a new, non-linear dictionary learning approach, called WDL. This new method fully leverages the attractive properties of OT in its learned representation of "data", which can come in many forms. In particular, WDL can be applied to images – including PSFs.

In Chapter 5, I made use of several concepts and numerical tools we created as part of the development of WDL. The addition of this OT component to the PSF modelling problem led to the creation of λRCA: a non-parametric PSF model that successfully recovers chromatic variations using no information other than images of observed stars, their SEDs (or estimates thereof), and positions within the focal plane.

Beyond these methodological contributions, I have performed extensive image simulations to test the impact of PSF modelling errors on galaxy shape measurements. Through these, I showed that in the case of *Euclid*, usual assumptions about the propagation of PSF errors break down. The quality of a PSF model is traditionally assessed by looking at its size and ellipticity: quantities of second order. This assessment is also mostly done independently from the shape measurement process. Our experiments show that for *Euclid* (or any nearly diffraction-limited instrument):

- higher-order effects can affect measured galaxy shapes, and thus shear estimation;
- the propagation of PSF error varies from one shape measurement method to another.

Outlook

Experiments show that our proposed PSF models already outperform the current state-of-the art on simulations. As discussed in Appendix B, however, some technical work remains before these models can reach their full potential within actual data processing pipelines. Similarly, the computational cost of our proposed λRCA approach currently makes its use in processing *Euclid* data unrealistic. This caveat is mostly related to the computation of OT quantities, however, for which Courty et al. (2017) recently proposed a workaround. Incorporating their approach into λRCA should lead to a dramatic decrease in computational overhead.

Beyond these technical improvements, a lot of prospects for the non-parametric PSF models proposed in this thesis could lead to vast improvements. We are presently at a factor of a few hundreds beyond *Euclid*'s stringent requirements. While this may seem like a severe gap, it is important to note that this has been achieved in a fairly pessimistic scenario. Indeed, it was recently pointed out that the simulated PSFs used throughout this thesis correspond to a "toleranced" realization of the telescope model. This means the effects included due to the manufacturing of the instrument are within specifications, but still correspond to a highly perturbed final instrument. The amplitude of the PSF variations occurring within our dataset thus correspond to a highly pessimistic case, and those of the true *Euclid* flight instrument are likely to be much milder. Moreover, now that we have a model that is capable of capturing chromatic variations, the main limiting factor on its performance is the number of available observations. As discussed at the end of both Chapter 3 and Chapter 5, an immediate mean of improving the quality of a PSF model is to enable it to be fitted to more stars simultaneously.

Our last experiment used only the average number of stars expected, from a single CCD, in a science exposure. As a first mean to increase the number of available stars, we could include a way of handling all CCDs at once. We discussed this prospect in Section 5.5. This straightforward improvement would allow us to use all the stars from any single exposure. Because *Euclid*'s PSF is expected to be, by design, extremely stable, another way of increasing the number of available stars is to use those from several different exposures. The crudest way to do so would be to determine a certain amount of exposures that can be considered to have been affected by the same PSF, and use stars from all of these simultaneously. These could be selected either because they were acquired in short succession, or within similar observing conditions (e.g. with the same solar aspect angle, the main driver for the thermal state of the telescope).

A more involved approach would be to fully account for temporal variations *within* the PSF modelling. We have already gone from a PSF model that was a function of position, to one that was a function of position *and* wavelength. We could as well add the time of observation (or some proxy thereof, like measured temperature of the spacecraft) as an extra dimension. Our PSF model could even be updated in an *online* way, that is, iteratively improved as new exposures are taken. Such an approach would allow for far more datapoints to be used during the fitting. It may even allow us to use the planned calibration fields, some of which will contain a far greater star density than the science exposures of the Wide survey.

This addition of a time dimension to our non-parametric PSF model is the first of three areas where future work in the continuation of this thesis can bring significant improvements. The second is the eventual combination of the approach presented in this thesis with physically-motivated models like that of Duncan et al. (in prep). It is already highly beneficial to have two independent methods to solve a problem that is critical to the success of the *Euclid* mission. If both end up successful, their *combination* is likely to lead to an even better PSF model.

A third far-reaching area worth exploring is trying to apply the tools developed in this thesis in wavefront space. Indeed, in the present work, we have only carried out our work in direct space, that is, on *images* of the PSF itself. Another option is to try and model the wavefront error (WFE) instead of the PSF itself, which could still be done in a non-parametric way. The WFE and resulting direct space PSF are related through the modulus of the Fourier transform. Trying to recover the WFE from direct-space star images thus falls into the field of *phase retrieval* problems (Shechtman et al., 2015).

By successfully transferring our method to wavefront space, we would gain access to a wealth of new information. In particular, characterization campaigns led, on the ground, on actual *Euclid* hardware can provide the impact on the final WFE caused by various components. At the very end of Chapter 5, we briefly mentioned

newly found effects causing severe “glitches” (i.e. very strong variations occurring over a very narrow, constant ranges of λ values) in the chromatic variations of the PSF. These glitches could, perhaps, still be captured in direct space using knowledge of their location (in wavelength), for instance by considering λRCA to provide a baseline model, on top of which we would try to capture these extra variations. This task would, however, be made significantly easier if the full WFE maps could be used instead of just the glitches’ location. This specific WFE map alone is already a major motivation for further exploration of this avenue. It would also enable us to use every other available source of information about the instrument’s WFE, and potentially even the wide corpus of tools adapted for wavefront analysis, like Zernike polynomials (Zernike, 1934). This avenue for research is thus closely related to that of combining physical and empirical PSF models.

Each of those three ideas are associated with exciting new problems to solve and potential developments to explore. Conversely, there are several other extensions or improvements to the work carried out in this thesis that are more minor and/or less exploratory, some of which we already mentioned:

- technical improvements to both RCA (in order to handle real-data corner cases) and λRCA (to reduce its computational cost);
- handling multiple CCDs simultaneously;
- simultaneously fitting chromatic and achromatic components (see Sections 5.4.2 and 5.5);
- replacing the Starlet transform Φ used in the sparsity constraints by one more adapted to the PSFs we aim to model;
- treating variations of the core and the wings of the PSF separately.

Modelling the Euclid PSF in a non-parametric way is a hard task. A lot of exciting work remains to be carried out, as illustrated by the ideas listed above. The work carried out during my thesis already made it possible to take steps towards a PSF model that is useable for *Euclid*. Several other useful contributions emanated from this methodological work, both on knowledge about the propagation of PSF errors and in the form of WDL, a generic unsupervised representation learning method. It also allowed me to delve into highly interesting and useful topics, from WL image simulations and galaxy shape measurement to OT. In addition to all the aforementioned PSF modelling-related topics I hope to explore in future work, I believe numerical OT could have several other exciting and beneficial applications in cosmology. As a generic example, Wasserstein barycenters are an excellent tool to interpolate between distributions that vary smoothly with some parameters, but that we can only sample so many times. This type of issue is, of course, ubiquitous in our field.

Appendices

Appendix A

List of publications.

This appendix contains a list of publications and presentations given as part of this thesis.

A.1 List of Paper 2 authors

Paper II is a Euclid "Standard Project" paper. The complete list of its authors can be found below.
M.A. Schmitz[1], J.-L. Starck[1], F. Ngole Mboula[2], N. Auricchio[3], J. Brinchmann[4,5], R.I. Vito Capobianco[6], R. Clédassou[7], L. Conversi[8], L. Corcione[6], N. Fourmanoit[9], M. Frailis[10], B. Garilli[11], F. Hormuth[12], D. Hu[13], H. Israel[14], S. Kermiche[9], T. D. Kitching[13], B. Kubik[15], M. Kunz[16], S. Ligori[6], P.B. Lilje[17], I. Lloro[18,19], O. Mansutti[10], O. Marggraf[20], R.J. Massey[21], F. Pasian[10], V. Pettorino[1], F. Raison[22], J.D. Rhodes[23], M. Roncarelli[3,24], R.P. Saglia[14,22], P. Schneider[20], S. Serrano[18,25], A.N. Taylor[26], R. Toledo-Moreo[27], L. Valenziano[3], C. Vuerli[10], J. Zoubian[9]

And their corresponding affiliations:

[1] AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France
[2] Institut LIST, CEA, Université Paris-Saclay, F-91191 Gif-Sur-Yvette Cedex, France
[3] INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, I-40129 Bologna, Italy
[4] Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, PT4150-762 Porto, Portugal
[5] Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
[6] INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, I-10025 Pino Torinese (TO), Italy
[7] Centre National d'Etudes Spatiales, Toulouse, France
[8] ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, Spain
[9] Aix-Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France
[10] INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, I-34131 Trieste, Italy
[11] INAF-IASF Milano, Via Alfonso Corti 12, I-20133 Milano, Italy
[12] von Hoerner & Sulger GmbH, SchloßPlatz 8, D-68723 Schwetzingen, Germany
[13] Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
[14] Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 München, Germany
[15] Institut de Physique Nucléaire de Lyon, 4, rue Enrico Fermi, 69622, Villeurbanne cedex, France
[16] Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland
[17] Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, N-0315 Oslo, Norway
[18] Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain

[19] Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain
[20] Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
[21] Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK
[22] Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, D-85748 Garching, Germany
[23] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109, USA
[24] Dipartimento di Fisica e Astronomia, Universitá di Bologna, Via Gobetti 93/2, I-40129 Bologna, Italy
[25] Institute of Space Sciences (IEEC-CSIC), c/Can Magrans s/n, 08193 Cerdanyola del Vallés, Barcelona, Spain
[26] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[27] Depto. de Electrónica y Tecnología de Computadoras Universidad Politécnica de Cartagena, 30202, Cartagena, Spain

A.2 Publications

A.2.1 Refereed articles

- M.A. Schmitz, M. Heitz, N. Bonneel, F. Ngole, D. Coeurjolly, M. Cuturi, G. Peyré, J.-L. Starck. Wasserstein Dictionary Learning: Optimal Transport-Based Unsupervised Nonlinear Dictionary Learning (Paper I). In *SIAM Journal on Imaging Sciences*, 11(1), 643-678, 2018.

- M.A. Schmitz et al. *Euclid*: Non-parametric point spread function field recovery through interpolation on a Graph Laplacian (Paper II). *Submitted to Astronomy & Astrophysics*.

See Section A.1 for a compete list of Paper II authors.

A.2.2 Proceedings

- M.A. Schmitz, M. Heitz, N. Bonneel, F. Ngole, D. Coeurjolly, M. Cuturi, G. Peyré, J.-L. Starck. Optimal transport-based dictionary learning and its application to *Euclid*-like Point Spread Function representation. In *Wavelets and Sparsity XVII*. Vol. 10394. International Society for Optics and Photonics, 2017.

- F. Nammour, M.A. Schmitz, F. Ngolè, J.-L. Starck, J.N. Girard. Radio Astronomical Image Restoration with Shape Constraint. In *Wavelets and Sparsity XVIII*. Vol. 11138. International Society for Optics and Photonics, 2019.

A.3 Talks & posters

- Euclid Consortium Meeting 2019, Helsinki, Finland, June 2019 (STAR prize awardee plenary talk).

- France for Euclid VII, Nice, France, November 2018 (talk).

- CNES Young Researchers Days, Toulouse, France, October 2018 (flashtalk & poster).

- French-Chinese Days on Weak Lensing, Saclay, October 2018 (talk).

- Euclid Consortium Meeting 2018, Bonn, Germany, June 2018 (talk).

- Cosmo21, Valencia, Spain, May 2018 (talk).

- CosmoStat Day on Machine Learning in Astrophysics, Saclay, France, January 2018 (talk).

- France for Euclid VI, Paris, France, December 2017 (talk).
- Data Science Summer School, Palaiseau, France, August 2017 (poster).
- SPIE Optics+Photonics, San Diego, CA, July 2017 (talk, proceeding).
- SPACEOBS colloquium on data inversion and assimilation, Orsay, France, May 2017 (talk).
- Cavalieri Workshop on Optimal Transport and Optimisation, Paris, France, October 2016 (talk).

Appendix B

PSF estimation on real data

CFIS is a survey in the r and u bands carried out from the CFHT at Mauna Kea, Hawai'i. Together with Pan-STARRS (Panoramic Survey Telescope and Rapid Response System), it forms UNIONS (Ultraviolet Near-Infrared Optical Northern Survey). Its photometric measurements are among the most crucial for *Euclid*'s photo-z measurements. The area of the sky covered by the survey is shown, along with that of *Euclid*, in Figure B.1.

Ongoing efforts are currently being carried out to use CFIS data to perform WL measurements. Two shape measurement pipelines are being independently applied to allow for comparison and mutual cross-validations: `LensFit` (that we already mentioned in Equation 2.4.3) and `ShapePipe`. This chapter presents my contributions to the PSF modelling and diagnostics for the latter.

This work was carried out in collaboration with the whole `ShapePipe` team, and especially my fellow PhD student Axel Guinot. These results will appear, along with many other diagnostics of the rest of the pipeline, in Guinot et al. (in prep). Some of these extra results also involve the PSF, for instance the spatial correlations between PSF model and measured galaxy shapes. In this chapter, however, I will focus only on diagnostics of the PSF model itself, and not on the PSF *correction*.

B.1 PSF models

In order to have a first functional shape measurement pipeline, standard methods to perform for each step (see Section 2.3 for a brief overview) were first implemented in `ShapePipe`. In the case of the PSF, we thus first implemented the tried-and-tested `PSFEx` approach (Bertin, 2011). See Sections 2.4.2 and 3.3.1 for more details about the `PSFEx` model. It was originally designed to be ran in conjunction with companion software `SExtractor` (Bertin and Arnouts, 1996) in order to perform shape measurement. In the case of `ShapePipe`, however, we only want to make use of its PSF model, so that we can use any shape measurement method. Small interfacing functions need to be applied to extract the PSF model from the `PSFEx` output files and compute it at any arbitrary position. The same thing was done in DES (Jarvis et al., 2016), HSC (Bosch et al., 2017), and in our own experimental set-up of Chapter 3.

A freely available code[1], written by Erin Sheldon and Eli Rykoff, performs this step and was originally used for our processing of the CFIS data. However, after running into some unidentified issues and in order to obtain more flexibility when changing the upsampling parameter of `PSFEx`, we later changed to a version of the script[2] I used for the experiments of Chapter 3, slightly modified to fit the needs of (and the interfacing with) `ShapePipe`.

As with both current surveys using `PSFEx`, most of its parameters are left to their default values which are also very close to those shown in Figure 3.3.4. One difference is in the use `XWIN_IMAGE,YWIN_IMAGE` instead of `X_IMAGE,Y_IMAGE` as the position parameters. We found that setting the maximum polynomial degree to 3 led to a decrease in the PSF model quality, as measured from the diagnostics shown in Sections B.2 and B.3. We therefore set it to 2. Lastly, the results we show in the present chapter were obtained by setting the `PSF_SAMPLING` parameter to 1. This means `PSFEx` did not attempt to perform superresolution. Some early testing on a small area of the sky seemed to indicate using a lower value led to worse results, though this is something we plan on testing further in the future. We note, however, that Bosch et al. (2017) also found issues when the superresolution mode of `PSFEx` was activated on HSC data, and that the PSF model we obtained with

[1] `https://github.com/esheldon/psfex`
[2] `https://github.com/MorganSchmitz/PySFEx`

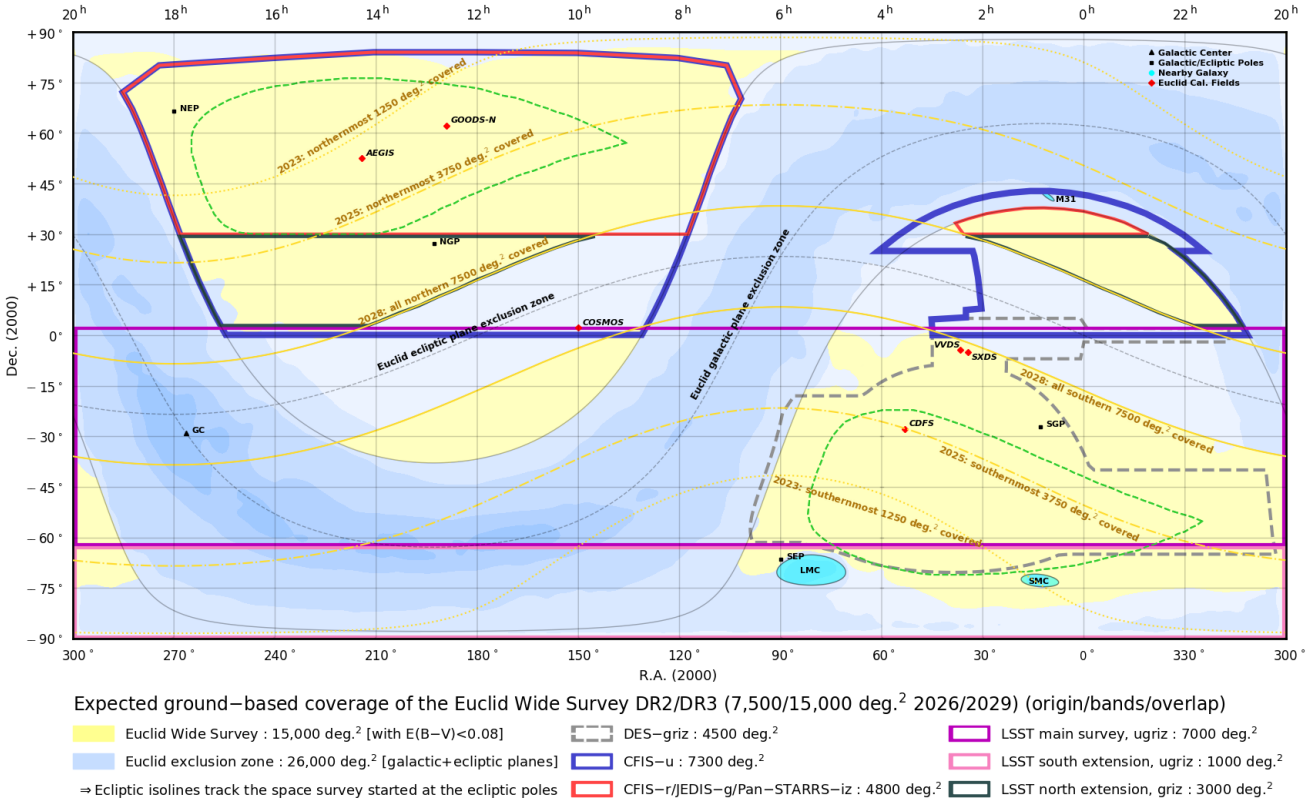


Figure B.1: Similar to Figure 2.14, the Euclid Wide Survey is shown in yellow, along with the footprint of several key ground-based surveys. In particular, CFIS-r, which produced the data used in this chapter, is shown in red. Credit: Jean-Charles Cuillandre.

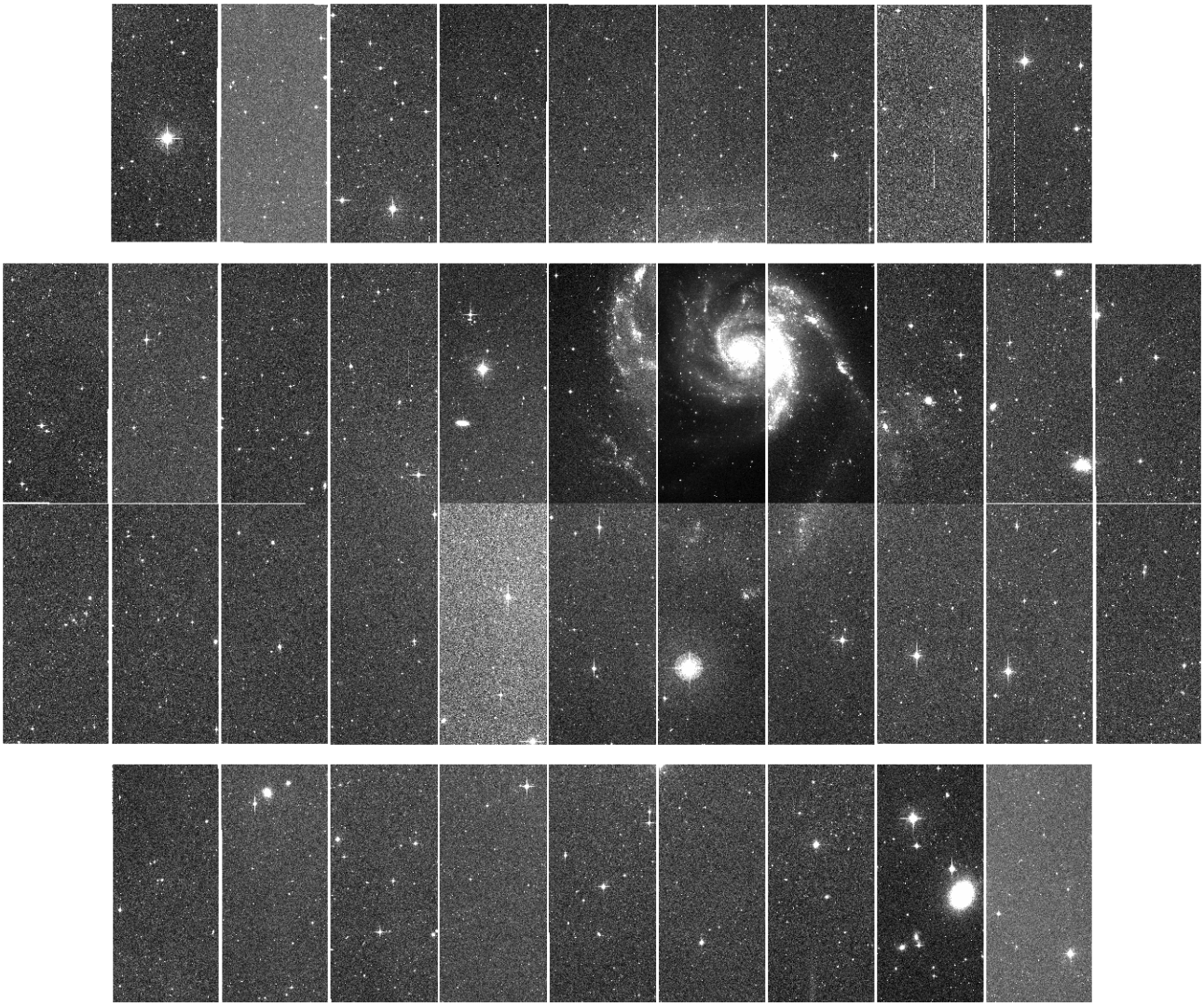


Figure B.2: A single exposure of the CFIS survey, observed with the CFHT's MegaCam. The object at the top-right of center of the mosaic is M101. Image rendered using the ds9 software (Joye and Mandel, 2003).

no superresolution already seems to be of good quality. We will show some of the diagnostics that allowed us to reach this conclusion in the rest of this chapter.

Our goal is to eventually include the PSF models described in earlier chapters of this thesis into ShapePipe. The RCA code is now implemented as a ShapePipe module and successfully runs within the pipeline. However, we found that because of several effects never encountered when applying it to simulated star fields, yet more modifications are required before it can be fully functional. As an example, the handling of SExtractor masks within the method was required. We used this opportunity to improve RCA by making it capable of handling any pixel-wise weights.

Another problematic corner case is that of a CCD where the star selection fails, for instance if the stellar locus is not found (see Figure 2.7). PSFEx's diagnostics can be used to identify and automatically remove such cases (for instance by performing a selection on the final χ^2 value). Critical failure cases could potentially be missed by RCA, however, as in its present state it assumes the catalog it receives as input contains only stars.

While not directly related to the PSF model itself, such a case of unusable star catalog occurred in the early days of our testing of ShapePipe on CFIS data. By checking instances of catastrophic failure of PSFEx individually, I came across a CCD that contained an abnormally large amount of objects identified as stars. Inspection of individual postage stamps from that catalog showed that several indeed appeared to contain a star, along with many SExtractor mask. We eventually had to return to the original exposure to identify the problem. It is reproduced in Figure B.2: the CCD was one close to object M101. While admittedly anecdotal, this example illustrates how studying the PSF can reveal issues with entirely different parts of the pipeline (in this case, local objects that had not been masked), or even the survey as a whole. Another example of an issue with ShapePipe revealed by PSF studies will be given in Section B.3.

B.2 ρ statistics

We now present diagnostics of our CFIS PSF model, computed on the original processing of an area of about $2,200$ square degrees.

In Section 2.4.1, we presented the Paulin-Henriksson et al. (2008) formalism of PSF error propagation. While we have subsequently shown its underlying assumptions no longer hold true for a diffraction-limited PSF (see Figure 3.3.6), it remains applicable to a ground-based survey like CFIS. We can thus use the real-data diagnostics we introduced in Section 2.4.2.

While there are, at this time, no requirements for CFIS WL, we can still use these to compare different PSF models applied to CFIS. As both HSC and DES use PSFEx, the ρ statistics also provide us with a mean to check whether it achieves similar performance when used in ShapeLens and on CFIS.

The ρ statistics computed across the whole processed area of CFIS are shown in Figure B.3. These were computed using the freely available Stile[3] package, developed for HSC/LSST. We note that the precise definition of the ρ statistics used by Zuntz et al. (2018) and Mandelbaum et al. (2017) differ. These correspond to the latest DES and HSC shape catalogs, respectively. In (2.113-2.117), the residual terms are uniquely defined up to a sign swap, but quantities related to the PSF shape itself can either be computed from stars (as is done in DES) or the PSF model (in HSC). Since the key element in the ρ statistics are the shape residuals, and they are computed through averaging over a large number of objects, this choice of convention is of little import. We do observe *some* variations in the resulting ρ values when computed using one convention or the other on both our own CFIS data, and the DES Year 1 star catalog.

The values we observe in Figure B.3 are of the same order as those obtained by Zuntz et al. (2018) and Mandelbaum et al. (2017). Since LensFit was also ran on CFIS data, we computed the ρ statistics using the LensFit star catalog and PSF model. Since these catalogs are not yet ready for the whole area processed by ShapePipe, this comparison was done on a smaller area, namely the W3 field of CFHTLenS (that was reobserved by CFIS). The ρ statistics computed on the same field for our own star/PSF catalogs were

[3] https://github.com/msimet/Stile

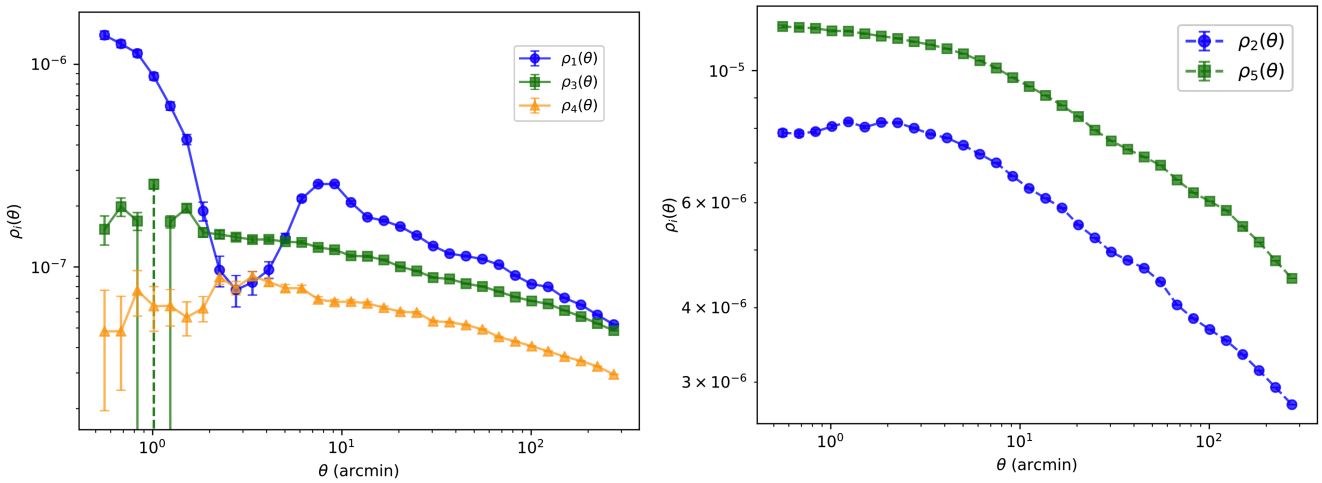


Figure B.3: ρ statistics computed from our PSFEx model of the CFIS-r dataset. Dotted lines indicate negative values, the absolute values of which are represented.

consistent with those shown in Figure B.3 (albeit, as expected, with larger error bars). We thus expect the comparison with the preliminary LensFit ρ values to be usable.

We find that LensFit leads to ρ values that are about an order of magnitude lower than ours. Since the LensFit and PSFEx models are fairly similar (see Section 2.4.3), and both were fitted to the same data, we attribute this improvement to the fact LensFit performs a cross-CCD fit. This fact highlights how a larger number of input stars can significantly increase the quality of a PSF model. Beyond the current study of the CFIS PSF, this result provides a strong motivation for the addition of multi-CCD fitting to RCA and λRCA. Moreover, the bump we observe in ρ_1 at scales of about 10arcmin in Figure B.3 is not present in the LensFIt ρ_1, and is likely related to CCD jumps.

B.3 Focal plane variations

The ρ statistics probe potential remaining spatially coherent residuals, in our PSF model, across sky positions. Another common diagnostic of PSF models on real data is to study these residuals as a function of position in the focal plane instead. This is, once again, achieved by measuring shapes of both observed stars and the PSF model estimated at their position, then averaging across a large number of objects. The two ellipticity components and their residuals across the MegaCam mosaic are shown in Figure B.4. Some patterns can be seen in the residuals, although these remain of fairly low amplitude.

Interestingly, even for a ground-based survey like CFIS, this type of diagnostic does not only provide information about potential patterns in the residual, but also about the optical PSF of the instrument. Indeed, as the atmospheric contribution to the PSF at any given position in the focal plane is expected to be largely random, they are cancelled out by averaging across a large number of exposures.

An easily computed byproduct of these diagnostics is the distribution of identified stars as a function of position across the FOV. This is shown in Figure B.5. Some vertical lines with slightly lower star density can be seen, especially close to the CCD borders, though these are not unexpected. “Dark zones” that contain very few stars, however, are clearly present in the four corners of the CCD mosaic, and especially noticeable in the bottom right. These are not present in the LensFit star catalogs, and are thus likely indicative of an issue with Shapepipe’s star selection. As of this writing, the exact origin of this issue has not yet been identified, though a working theory is that it might be related to a poor background estimation.

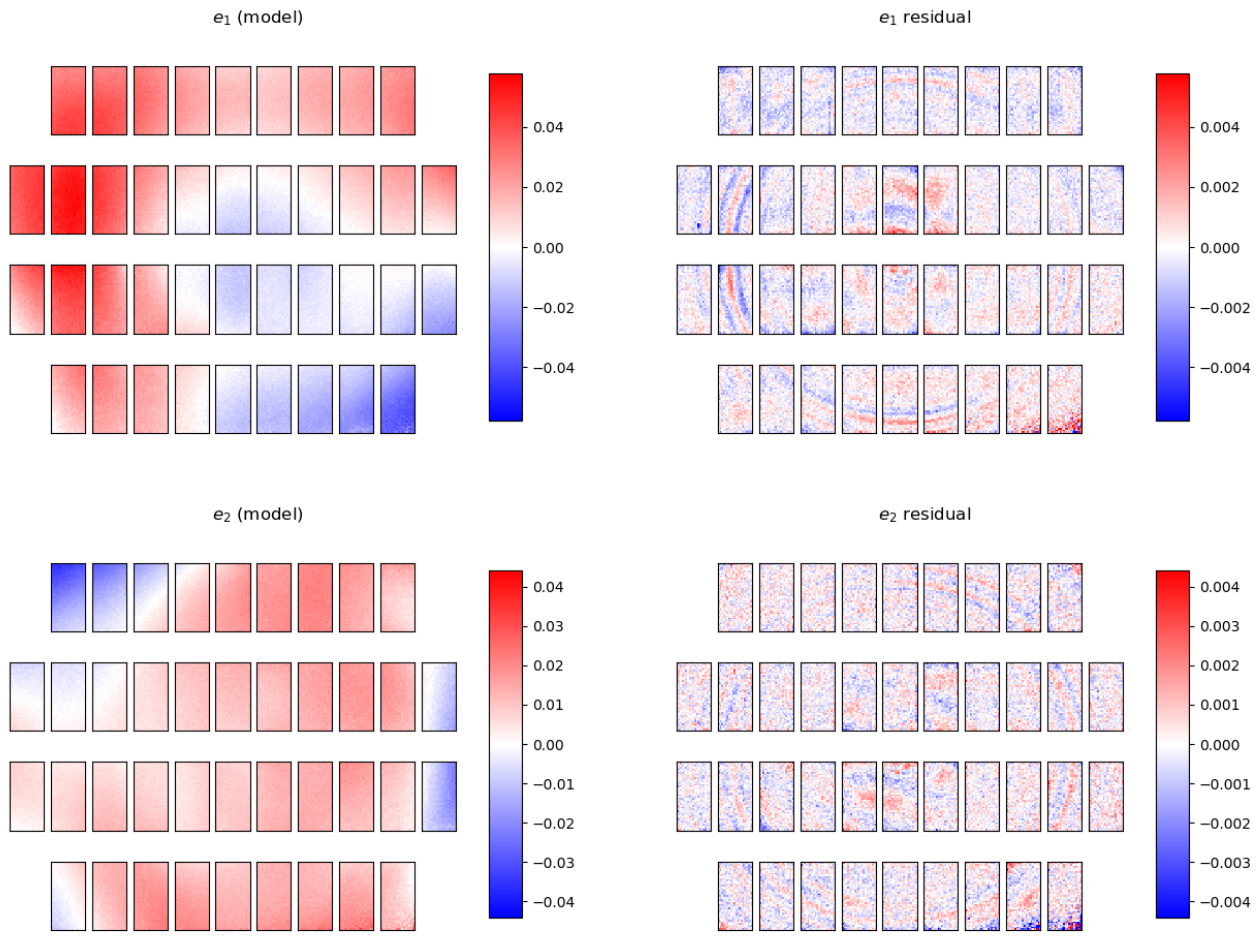


Figure B.4: Spatial distribution of the MegaCam PSF ellipticity (left), and average residuals after correction by our PSF model (right). The first (resp. second) row contains plots for the first (resp. second) ellipticity component.

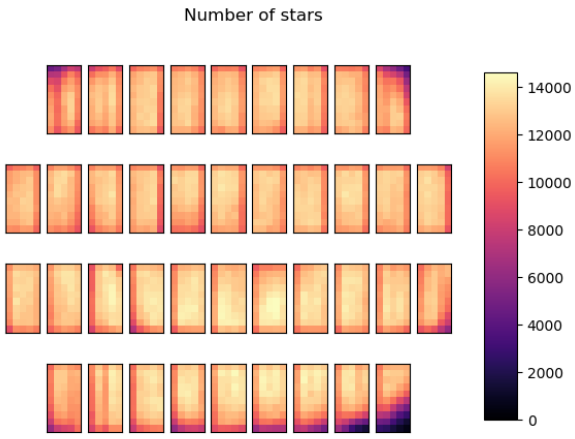


Figure B.5: Total number of stars identified, per position within the MegaCam FOV, across the whole processed CFIS data. Each CCD is separated into 5×10 bins.

B.4 Conclusion to Appendix B

This appendix presented my main contributions to the ongoing WL processing of CFIS data. We have shown several standard PSF diagnostics used in modern WL surveys. These are performed in addition to more standard checks, such as goodness-of-fit metrics and visual inspection of the residuals of star image after subtraction by the corresponding PSF model. Taken together, these indicate that our `PSFEx` model is fairly successful in modelling the PSF in the CFIS survey. In particular, Figures B.3 and B.4 seem to indicate we have achieved similar levels of PSF modelling accuracy to those observed in other recent surveys that use `PSFEx`.

An independent running of the `LensFit` pipeline on the same data, however, led to a PSF model that outperforms `PSFEx` in terms of ρ statistics. This is likely due to `LensFit`'s PSF model being fitted to the whole exposure simultaneously, rather than CCD-by-CCD.

Our proposed RCA approach, described in Chapter 3, will soon achieve the necessary code maturity to be successfully applied to these data. The diagnostics shown in this chapter can then be directly computed, and we will be able to compare its performance to `PSFEx` not just on simulations, but also on real data.

Appendix C

Shape measurement experiment: additional information

This appendix contains further details about the experiments of Chapter 3.

C.1 `im3shape` configuration

Below is the configuration file used when running `im3shape`

```
noise_sigma = 0.01
background_subtract = NO
psf_truncation_pixels = 50.0
stamp_size = 42

sersics_x0_start = 21.0
sersics_y0_start = 21.0
sersics_x0_min = 18.0
sersics_y0_min = 18.0
sersics_x0_max = 24.0
sersics_y0_max = 24.0

psf_input = psf_image_cube
perform_pixel_integration = NO
upsampling = 1
central_pixel_upsampling = NO
padding = 0
```

Note that default values are used for all parameters not specified in this config file.

C.2 KSB-HSM outlier counts

Table C.1 contains the number of objects removed from the analysis when using KSB.

	SNR10	SNR20	SNR35	SNR50
Known	26 741	26 702	26 787	26 798
PSFEx	63 541	63 356	63 236	63 111
RCA	54 345	51 664	52 100	51 739
Total	**72 951**	**71 551**	**71 128**	**70 902**

Table C.1: Number of objects where the HSM implementation of KSB fails to compute the PSF-corrected shapes per star SNR level. The total amount corresponds to the union of the 3 PSF-specific sets of such outliers.

Appendix D

Proof of (4.17-4.18)

By differentiating (4.23) with regard to the dictionary or one of the barycentric weights, we can rewrite the Jacobians in (4.16), (4.17), respectively, while separating the differentiations with regard to the dictionary D, the weights λ_i and the scaling vector b, by total differentiation and the chain rule:

$$\left[\partial_D P^{(l)}(D,\lambda)\right]^\top = \Psi_D^{(l-1)} + B_D^{(l-1)}\Psi_b^{(l-1)}, \tag{D.1}$$

$$\left[\partial_\lambda P^{(l)}(D,\lambda)\right]^\top = \Psi_\lambda^{(l-1)} + B_\lambda^{(l-1)}\Psi_b^{(l-1)}. \tag{D.2}$$

And, differentiating (4.24),

$$B_D^{(l)} = \Phi_D^{(l-1)} + B_D^{(l-1)}\Phi_b^{(l-1)}, \tag{D.3}$$

$$B_\lambda^{(l)} = \Phi_\lambda^{(l-1)} + B_\lambda^{(l-1)}\Phi_b^{(l-1)}. \tag{D.4}$$

We then have, by definitions (4.30)-(4.31) and by plugging (D.1) and (D.3) into (4.16),

$$\begin{aligned}
\nabla_D \mathcal{E}_L(D,\lambda) &= \Psi_D^{(L-1)}\left(\nabla\mathcal{L}(P^{(L)}(D,\lambda),x)\right) + B_D^{(L-1)}v^{(L-1)} \\
&= \Psi_D^{(L-1)}\left(\nabla\mathcal{L}(P^{(L)}(D,\lambda),x)\right) + \Phi_D^{(L-2)}\left(v^{(L-1)}\right) + B_D^{(L-2)}\left(v^{(L-2)}\right) \\
&= \dots \\
\nabla_D \mathcal{E}_L(D,\lambda) &= \Psi_D^{(L-1)}\left(\nabla\mathcal{L}(P^{(L)}(D,\lambda),x)\right) + \sum_{l=0}^{L-2}\Phi_D^{(l)}\left(v^{(l+1)}\right),
\end{aligned} \tag{D.5}$$

where the sum starts at 0 because $B_D^{(0)} = 0$ since we initialized $b^{(0)}$ as a constant vector. This proves (4.28). Similarly, differentiating with regard to λ yields

$$\nabla_\lambda \mathcal{E}_L(D,\lambda) = \Psi_\lambda^{(L-1)}\left(\nabla\mathcal{L}(P^{(L)}(D,\lambda),x)\right) + \sum_{l=0}^{L-2}\Phi_\lambda^{(l)}\left(v^{(l+1)}\right).$$

This proves (4.29). The detailed derivation of the differentials of φ, Φ, and Ψ with regard to all three variables is given in Appendix F.

Appendix E

Stabilized backward loop

Inputs: Data $x \in \Sigma_N$, atoms $d_1, \ldots, d_S \in \Sigma_N$, current weights $\lambda \in \Sigma_S$

comment: Sinkhorn loop

$\forall s, v_s^{(0)} := \mathbf{0}_N$

for $l = 1$ to L step 1 do

$\quad \forall s, \tilde{\varphi}_s^{(l)} := K_{LS}\left(\log(d_s) - K_{LS}(v_s^{(l-1)})\right)$

$\quad \tilde{p} := \sum_s \lambda_s \tilde{\varphi}_s^{(l)}$

$\quad \forall s, v_s^{(l)} := \tilde{p} - \tilde{\varphi}_s^{(l)}$

od

$p = \exp(\tilde{p})$

comment: Backward loop - weights

$w := \mathbf{0}_S$

$r := \mathbf{0}_{S \times N}$

$g := \nabla \mathcal{L}(p, x) \odot p$

for $l = L$ to 1 step -1 do

$\quad \forall s, w_s := w_s + \langle \tilde{\varphi}_s^{(l)}, g \rangle$

$\quad \forall s, \tilde{t}_s := K_{LS}\left(\log(\lambda_s g - r_s) - \tilde{\varphi}_s^{(l)}\right) + \log(d_s) - 2 * K_{LS}(v_s^{(l-1)})$

$\quad \forall s, r_s := \exp\left(K_{LS}(\tilde{t}_s) + v_s^{(l-1)}\right)$

$\quad g := -\sum_s r_s$

od

comment: Backward loop - dictionary

$y := \mathbf{0}_{S \times N}$

$z := \mathbf{0}_{S \times N}$

$n := \nabla \mathcal{L}(p, x)$

for $l = L$ to 1 step -1 do

$\quad \forall s, \tilde{c}_s := K_{LS}\left(\log(\lambda_s n + z_s) + v_s^{(l)}\right)$

$\quad \forall s, y_s := y_s + \exp\left(\tilde{c}_s - K_{LS}(v_s^{(l-1)})\right)$

$\quad \forall s, z_s := \exp\left(-\tilde{\varphi}_s^{(l-1)} + K_{LS}\left(\log(d_s) + \tilde{c}_s - 2 * K_{LS}(v_s^{(l-1)})\right)\right)$

$\quad n := -\sum_s z_s$

od

Outputs: $P^{(L)}(D, \lambda) := p, \nabla_D \mathcal{E}^{(L)} := y, \nabla_\lambda \mathcal{E}^{(L)} := w$

Algorithm E.1: `logSinkhornGrads`: Computation of dictionary and barycentric weights gradients in log-domain. Log-domain variables are marked with a tilde.

Appendix F

Detailed derivations

Let us first introduce the following notation:

$$\varphi: \begin{array}{c} \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N \\ b_s, d \mapsto K^\top \dfrac{d}{Kb_s} \end{array}.$$

F.1 Computation of $\partial_b \varphi$

By definition,

$$\frac{\partial \varphi}{\partial b_s}(b_s, d) = -K^\top \Delta\left(\frac{d}{(Kb_s)^2}\right) K. \tag{F.1}$$

In what follows, we will denote $\varphi_{NS}(b, D) = [\varphi(b_1, d_1)^\top, \dots, \varphi(b_S, d_S)^\top]^\top \in \mathbb{R}^{NS}$.

$$\partial_b \varphi_{NS}(b, D) = \begin{pmatrix} \frac{\partial \varphi(b_1, d_1)}{\partial b_1} & \mathbf{0}_{N\times N} & \dots & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \frac{\partial \varphi(b_2, d_2)}{\partial b_2} & \dots & \mathbf{0}_{N\times N} \\ \vdots & & \ddots & \vdots \\ \mathbf{0}_{N\times N} & \dots & \mathbf{0}_{N\times N} & \frac{\partial \varphi(b_S, d_S)}{\partial b_S} \end{pmatrix}.$$

F.2 Computation of Ψ_b

Taking the logarithm of (4.25) yields

$$\log(\Psi(b, D, \lambda)) = \sum_s \lambda_s \log(\varphi(b_s, d_s)),$$

the differentiation of which gives us

$$\begin{aligned} \Delta\left(\frac{\mathbf{1}_N}{\Psi(b, D, \lambda)}\right) \partial_b \Psi(b, D, \lambda) &= \begin{pmatrix} \lambda_1 I_N & \dots & \lambda_S I_N \end{pmatrix} \Delta\left(\frac{\mathbf{1}_{NS}}{\varphi_{NS}(b, D)}\right) \partial_b \varphi_{NS}(b, D) \\ \implies \Psi_b &= [\partial_b \varphi_{NS}(b, D)]^\top \Delta\left(\frac{\mathbf{1}_{NS}}{\varphi_{NS}(b, D)}\right) J_\lambda \Delta(\Psi(b, D, \lambda)), \end{aligned} \tag{F.2}$$

where $J_\lambda = \begin{pmatrix} \lambda_1 I_N \\ \vdots \\ \lambda_S I_N \end{pmatrix} \in \mathbb{R}^{NS\times N}$.

F.3 Computation of Ψ_D

Let $i \in \{1, \dots, S\}$.

$$\Psi(b, D, \lambda) = \prod_{s \neq i} \Delta(\varphi_c(b_s, d_s))^{\lambda_s} . \left(K^\top \frac{d_i}{K b_i} \right)^{\lambda_i},$$

and

$$\begin{aligned} \frac{\partial \left(K^\top \frac{d_i}{K b_i} \right)^{\lambda_i}}{\partial d_i} &= \lambda_i \Delta \left(K^\top \frac{d_i}{K b_i} \right)^{\lambda_i - 1} K^\top \Delta \left(\frac{\mathbf{1}_N}{K b_i} \right) \\ \implies \frac{\partial \Psi}{\partial d_i}(b, D, \lambda) &= \lambda_i \frac{\Delta(\Psi(b, D, \lambda))}{\Delta \left(K^\top \frac{d_i}{K b_i} \right)} K^\top \left(\frac{\mathbf{1}_N}{K b_s} \right). \end{aligned} \tag{F.3}$$

F.4 Computation of Φ_b

$$\begin{aligned} \partial_b \Phi(b, D, \lambda) &= \begin{pmatrix} \Delta \left(\frac{\mathbf{1}_N}{\varphi(b_1, d_1)} \right) \\ \vdots \\ \Delta \left(\frac{\mathbf{1}_N}{\varphi(b_S, d_S)} \right) \end{pmatrix} \partial_b \Psi(b, d) \\ &\quad - \begin{pmatrix} \Delta \left(\frac{\Psi(b,D,\lambda)}{\varphi(b_1,d_1)^2} \right) \frac{\partial \varphi(b_1, d_1)}{\partial b_1} & \mathbf{0}_{N \times N} & \dots & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \Delta \left(\frac{\Psi(b,D,\lambda)}{\varphi(b_2,d_2)^2} \right) \frac{\partial \varphi(b_2, d_2)}{\partial b_2} & \dots & \mathbf{0}_{N \times N} \\ \vdots & & \ddots & \vdots \\ \mathbf{0}_{N \times N} & \dots & \mathbf{0}_{N \times N} & \Delta \left(\frac{\Psi(b,D,\lambda)}{\varphi(b_S,d_S)^2} \right) \frac{\partial \varphi(b_S, d_S)}{\partial b_S} \end{pmatrix} \\ &= \Delta \left(\frac{\mathbf{1}_{NS}}{\varphi_{NS}(b, D)} \right) I_{N,S}^\top (\partial_b \Psi(b, D, \lambda)) - \Delta \left(\frac{\mathbf{1}_{NS}}{\varphi_{NS}(b, D)} \right) \Delta(\Phi(b, D, \lambda)) \partial_b \varphi_{NS}(b, D) \\ &= \Delta \left(\frac{\mathbf{1}_{NS}}{\varphi_{NS}(b, D)} \right) \left[I_{N,S}^\top (\partial_b \Psi(b, D, \lambda)) - \Delta(\Phi(b, D, \lambda)) \partial_b \varphi_{NS}(b, D) \right] \\ \implies \Phi_b &= \left[\Psi_b I_{N,S} - [\partial_b \varphi_{NS}(b, D)]^\top \Delta(\Phi(b, D, \lambda)) \right] \Delta \left(\frac{\mathbf{1}_{NS}}{\varphi_{NS}(b, D)} \right) \\ &\overset{(F.2)}{=} [[\partial_b \varphi_{NS}(b, D)]^\top \Delta \left(\frac{\mathbf{1}_{NS}}{\varphi(b, D)} \right) J_\lambda \Delta(\Psi(b, D, \lambda)) I_{N,S} \\ &\qquad - [\partial_b \varphi_{NS}(b, D)]^\top \Delta(\Phi(b, D, \lambda))] \Delta \left(\frac{\mathbf{1}_{NS}}{\varphi_{NS}(b, D)} \right) \\ &= [\partial_b \varphi_{NS}(b, D)]^\top \left[\Delta \left(\frac{\mathbf{1}_{NS}}{\varphi(b, D)} \right) J_\lambda \Delta(\Psi(b, D, \lambda)) I_{N,S} - \Delta(\Phi(b, D, \lambda)) \right] \Delta \left(\frac{\mathbf{1}_N}{\varphi_{NS}(b, D)} \right), \end{aligned} \tag{F.4}$$

where $I_{N,S} = [I_N, \dots, I_N] \in \mathbb{R}^{N \times NS}$. Moreover, we have

$$\Delta\left(\frac{\mathbf{1}_{NS}}{\varphi(b,D)}\right)J_\lambda\Delta(\Psi(b,D,\lambda)) = \begin{pmatrix}\Delta(1/\varphi(b_1,d_1)) & & \\ & \ddots & \\ & & \Delta(1/\varphi(b_S,d_S))\end{pmatrix}\begin{pmatrix}\lambda_1\Delta(\Psi(b,D,\lambda)) \\ \vdots \\ \lambda_S\Delta(\Psi(b,D,\lambda))\end{pmatrix}$$
$$= \begin{pmatrix}\lambda_1\Delta\left(\frac{\Psi(b,D,\lambda)}{\varphi(b_1,d_1)}\right) & & \\ & \ddots & \\ & & \lambda_S\Delta\left(\frac{\Psi(b,D,\lambda)}{\varphi(b_S,d_S)}\right)\end{pmatrix}$$
$$= \Delta(\Phi(b,D,\lambda))\begin{pmatrix}\lambda_1 I_N \\ \vdots \\ \lambda_S I_N\end{pmatrix}$$
$$\Delta\left(\frac{\mathbf{1}_{NS}}{\varphi(b,D)}\right)J_\lambda\Delta(\Psi(b,D,\lambda)) = \Delta(\Phi(b,D,\lambda))J_\lambda.$$

Hence, in (F.4),

$$\Phi_b = [\partial_b\varphi_{NS}(b,D)]^\top\Delta(\Phi(b,D,\lambda))[J_\lambda I_{N,S} - I_{NS}]\Delta\left(\frac{\mathbf{1}_N}{\varphi_{NS}(b,D)}\right).$$

F.5 Computation of Φ_D

Let $i \in \{1,\dots\}$. $\forall s \neq i$, the only dependency in d_i of $\Phi^s(b,D,\lambda)$ resides in Ψ (see (4.26)), hence

$$\forall s \neq i, \frac{\partial\Phi^s}{\partial d_i} = \Delta\left(\frac{\mathbf{1}_N}{\varphi(b_s,d_s)}\right)\partial_{d_i}\Psi$$
$$\overset{(F.3)}{=} \lambda_i\frac{\Delta(\Psi(B,D,\lambda))}{\Delta(\varphi(b_s,d_s))\Delta(\varphi(b_i,d_i))}K^\top\Delta\left(\frac{\mathbf{1}_N}{Kb_i}\right)$$
$$\overset{(4.26)}{=} \lambda_i\frac{\Delta(\Phi^i(B,D,\lambda))}{\Delta(\varphi(b_s,d_s))}K^\top\Delta\left(\frac{\mathbf{1}_N}{Kb_i}\right).$$

As for $s = i$, we have

$$\Phi^i(b,D,\lambda) = \frac{\Psi(b,D,\lambda)}{K^\top\frac{d_i}{Kb_i}}$$
$$\implies \frac{\partial\Phi^i}{\partial d_i}(b,D,\lambda) = \Delta\left(\frac{\mathbf{1}_N}{\varphi(b_1,d_1)}\right)\partial_D\Psi(b,D,\lambda) - \frac{\Delta(\Psi(b,D,\lambda))}{\Delta(\varphi_i(b_i,d_i)^2)}\partial_{d_i}\varphi(b_i,d_i)$$
$$= \Delta\left(\frac{\mathbf{1}_N}{\varphi(b_1,d_1)}\right)\partial_D\Psi(b,D,\lambda) - \frac{\Delta(\Phi^i(b,D,\lambda))}{\Delta(\varphi(b_i,d_i))}K^\top\left(\frac{\mathbf{1}_N}{Kb_i}\right)$$
$$= (\lambda_i - 1)\frac{\Delta(\Phi^i(b,D,\lambda))}{\Delta(\varphi(b_i,d_i))}K^\top\Delta\left(\frac{\mathbf{1}_N}{Kb_i}\right).$$

Appendix G

Generalized barycenters

Inputs: Data $x \in \Sigma_N$, atoms $d_1, \dots, d_S \in \Sigma_N$, weights $\lambda \in \Sigma_S$,
extrapolation parameter $\tau \leq 0$

$\forall s, b_s^{(0)} := \mathbf{1}_N$

for $l = 1$ to L step 1 do

$\quad \forall s, \tilde{a}_s^{(l)} := \frac{d_s}{K b_s^{(l-1)}}$

$\quad \forall s, a_s^{(l)} := \left(a_s^{(l-1)}\right)^{\tau} \left(\tilde{a}_s^{(l)}\right)^{1-\tau}$

$\quad p := \prod_s \left(K^\top a_s^{(l)}\right)^{\lambda_s}$

$\quad \forall s, \tilde{b}_s^{(l)} := \frac{p}{K^\top a_s^{(l)}}$

$\quad \forall s, b_s^{(l)} := \left(b_s^{(l-1)}\right)^{\tau} \left(\tilde{b}_s^{(l)}\right)^{1-\tau}$

od

Outputs: $P^{(L)}(D, \lambda) := p$

Algorithm G.1: `HeavyballSinkhorn`: Computation of approximate Wasserstein barycenters with acceleration.

Inputs: Data $x \in \Sigma_N$, atoms $d_1, \dots, d_S \in \Sigma_N$, weights $\lambda \in \Sigma_S$,
extrapolation parameter $\tau \leq 0$, KL parameter $\rho > 0$

$\forall s, b_s^{(0)} := \mathbf{1}_N$

for $l = 1$ to L step 1 do

$\quad \forall s, \tilde{a}_s^{(l)} := \left(\frac{d_s}{K b_s^{(l-1)}}\right)^{\frac{\rho}{\rho+\gamma}}$

$\quad \forall s, a_s^{(l)} := \left(a_s^{(l-1)}\right)^{\tau} \left(\tilde{a}_s^{(l)}\right)^{1-\tau}$

$\quad p := \left(\sum_{s=1}^S \lambda_s \left(K^\top a_s^{(l)}\right)^{\frac{\gamma}{\rho+\gamma}}\right)^{\frac{\rho+\gamma}{\gamma}}$

$\quad \forall s, \tilde{b}_s^{(l)} := \left(\frac{p}{K^\top a_s^{(l)}}\right)^{\frac{\rho}{\rho+\gamma}}$

$\quad \forall s, b_s^{(l)} := \left(b_s^{(l-1)}\right)^{\tau} \left(\tilde{b}_s^{(l)}\right)^{1-\tau}$

od

Outputs: $P^{(L)}(D, \lambda) := p$

Algorithm G.2: `GeneralizedSinkhorn`: Computation of unbalanced barycenters with acceleration.

Appendix H

Additional results to Chapter 4

H.1 MNIST and Wasserstein geodesics

This section contains the additional Figures H.1, H.2, H.3, and H.4, for the application of Section 4.5.1.

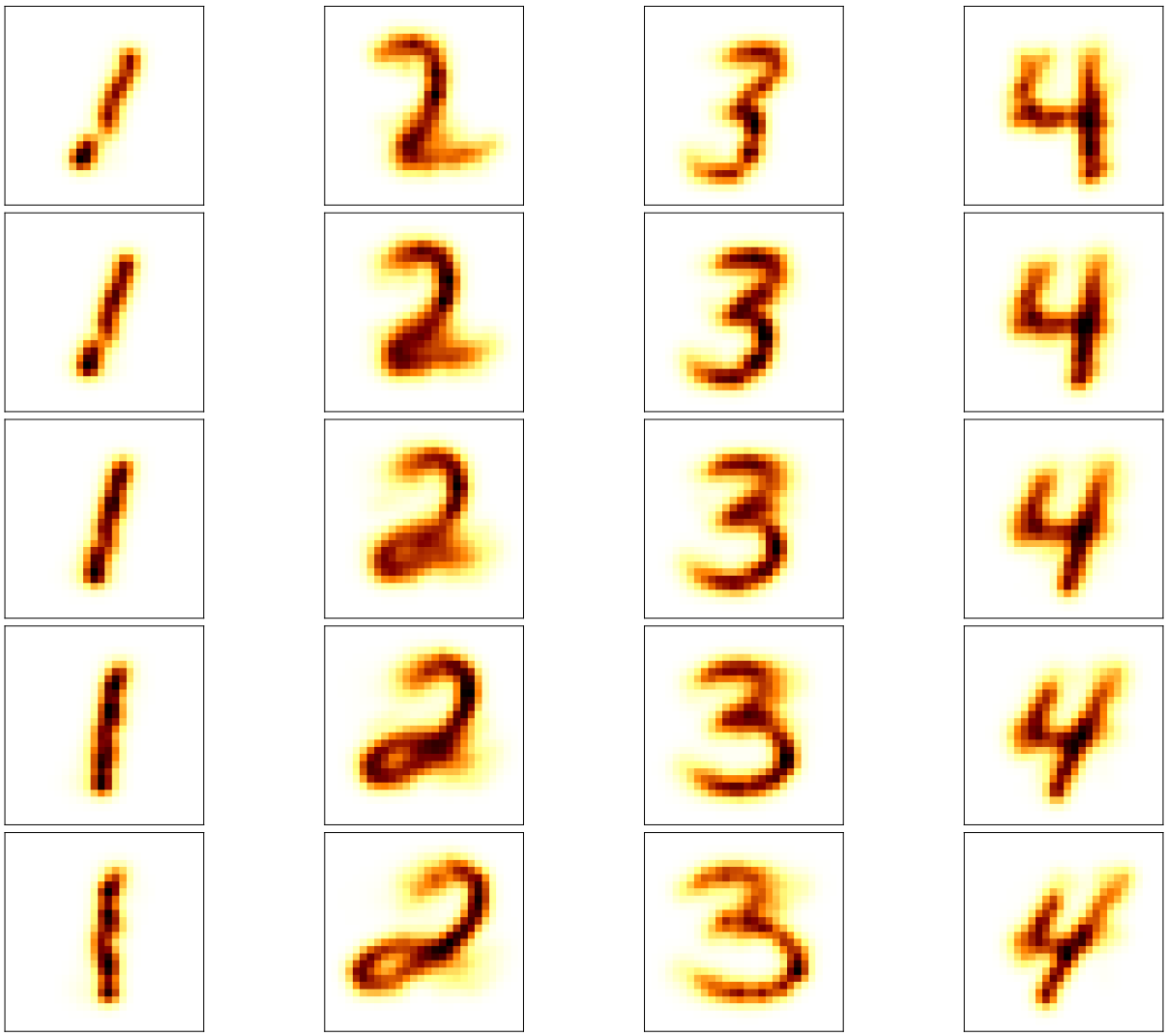


Figure H.1: Span of our two-atom dictionary for weights $(1-t, t), t \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ when trained on images of digits $1, 2, 3, 4$. See the first columns of (Seguy and Cuturi, 2015, Figure 5) for comparison with first WPGs.

H.2 PSF application

This section contains the additional Figure H.5 for the application of Section 4.5.2.

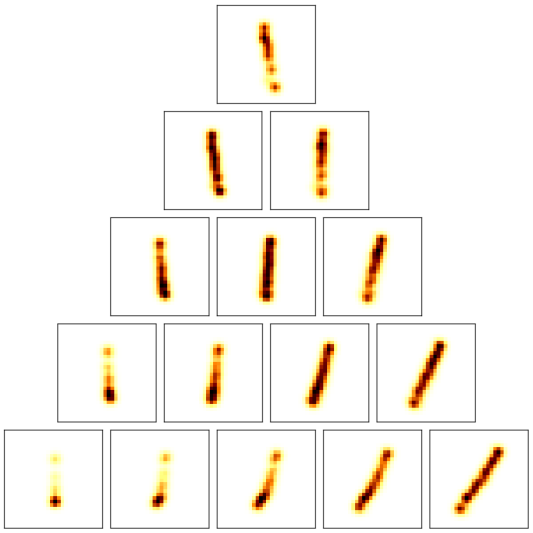


Figure H.2: Same as Figure 4.6 when training on images of the digit 1.

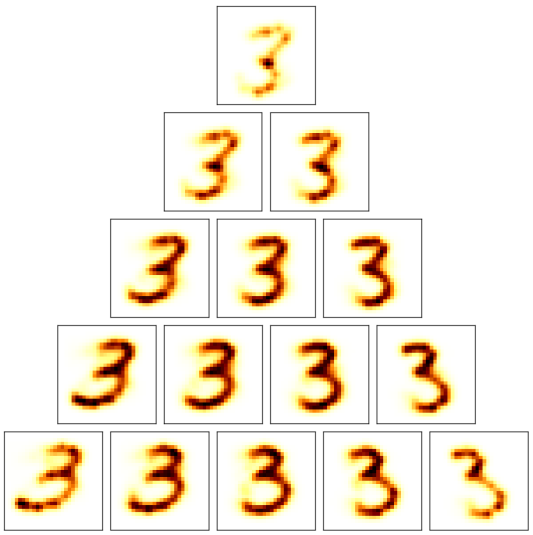


Figure H.3: Same as Figure 4.6 when training on images of the digit 3.

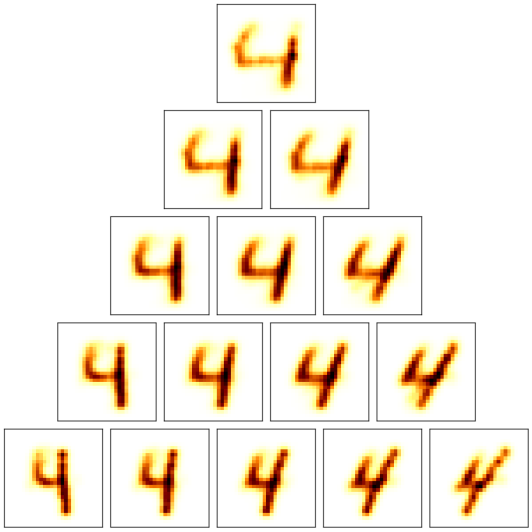


Figure H.4: Same as Figure 4.6 when training on images of the digit 4.

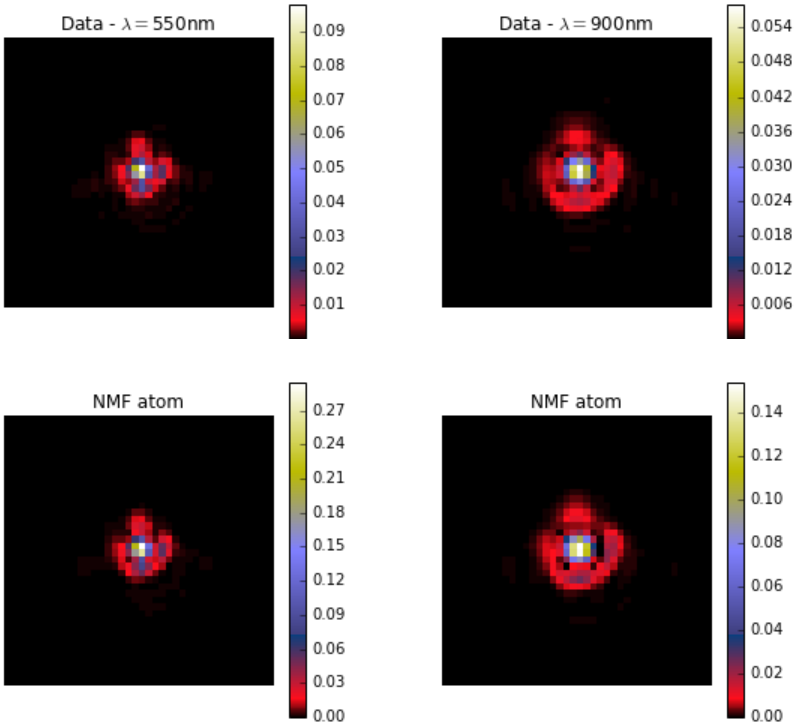


Figure H.5: Extreme wavelength PSFs in the dataset and atoms learned from NMF. See Figure 4.9 for those learned using our method.

Figure H.6: Similarly to Figure 4.14, we compare the atoms obtained using different loss functions, ranking them by mean PSNR: (a) $\overline{PSNR} = 33.81$, (b) $\overline{PSNR} = 33.72$, (c) $\overline{PSNR} = 32.95$, and (d) $\overline{PSNR} = 32.34$.

H.3 Wasserstein faces

This section contains the additional FIgures H.6 and H.7 for the application of Section 4.5.4.

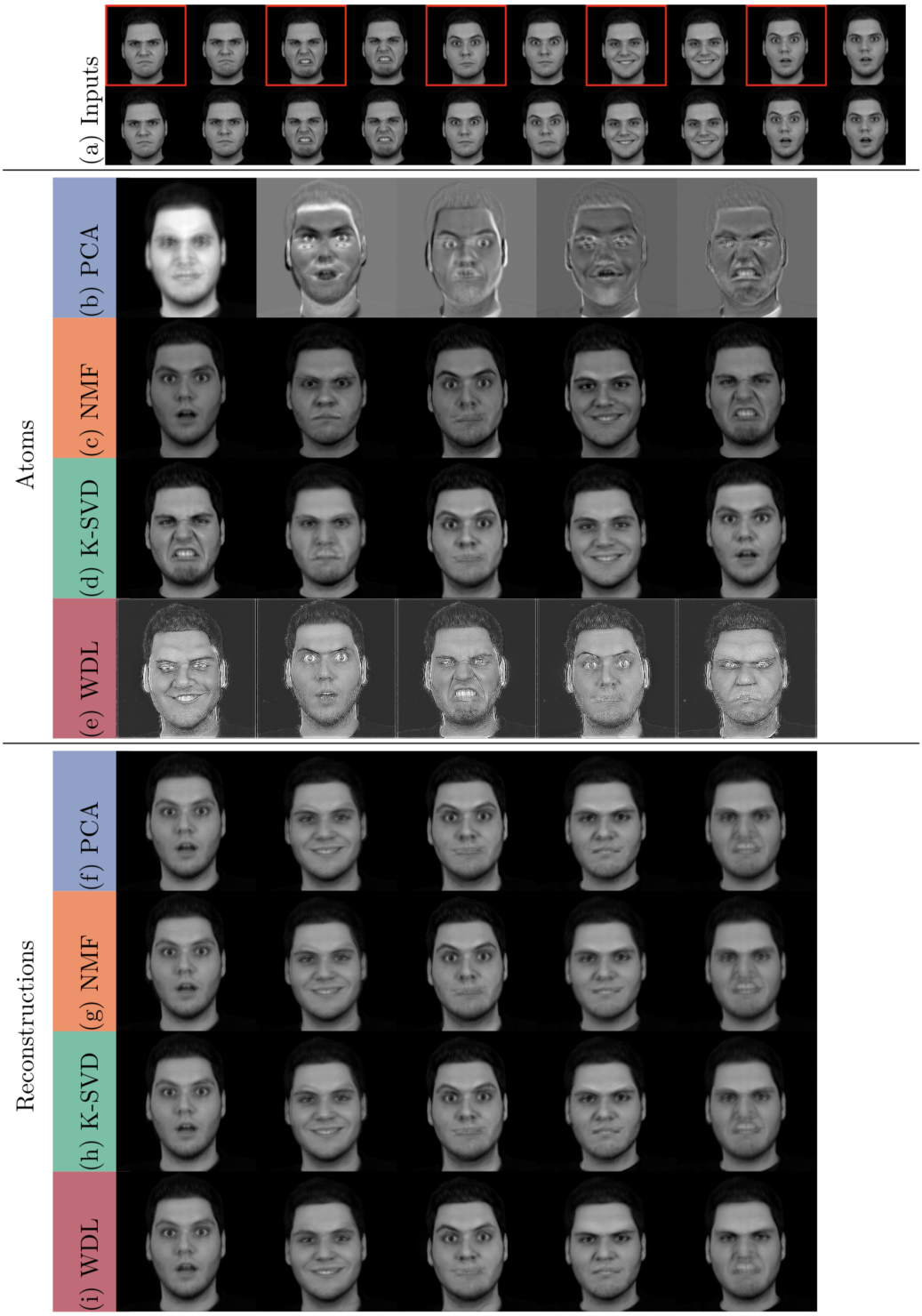


Figure H.7: Similarly to Figure 4.13, we compare our method to the Eigenfaces (Turk and Pentland, 1991) approach, NMF and K-SVD as a tool to represent faces on a low-dimensional space.

Appendix I

Galaxy SEDs

This appendix contains plots of all the template galaxy SEDs used in Chapter 5. Note unlike those shown in Figure 5.7, these have not yet been normalized or resampled with the spectral sampling step of $\Delta\lambda = 5\text{nm}$. The redshift value indicated in the titles corresponds to the lower end of each redshift bin. Figure I.1 contains the rest of the templates for elliptical galaxies; Figures I.2 and I.3 starbursts; and Figure I.4 spirals and (the) lenticular.

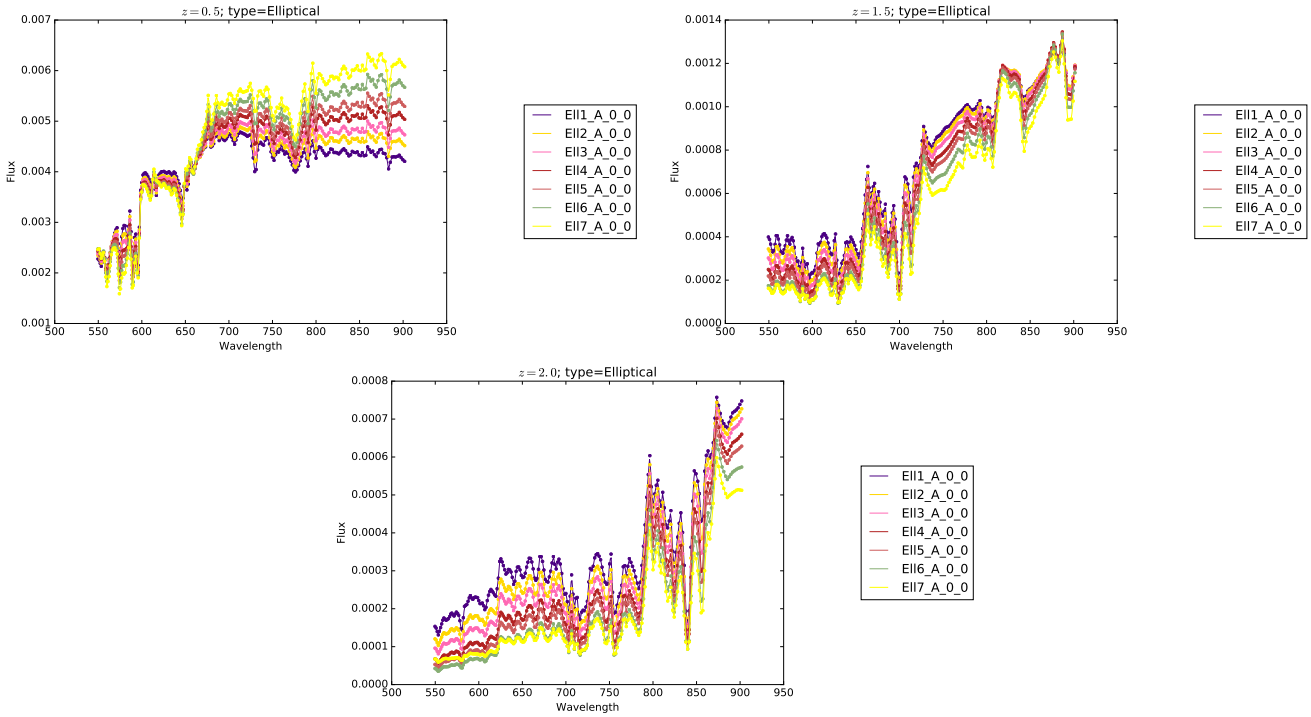


Figure I.1: Template SEDs for Elliptical galaxies.

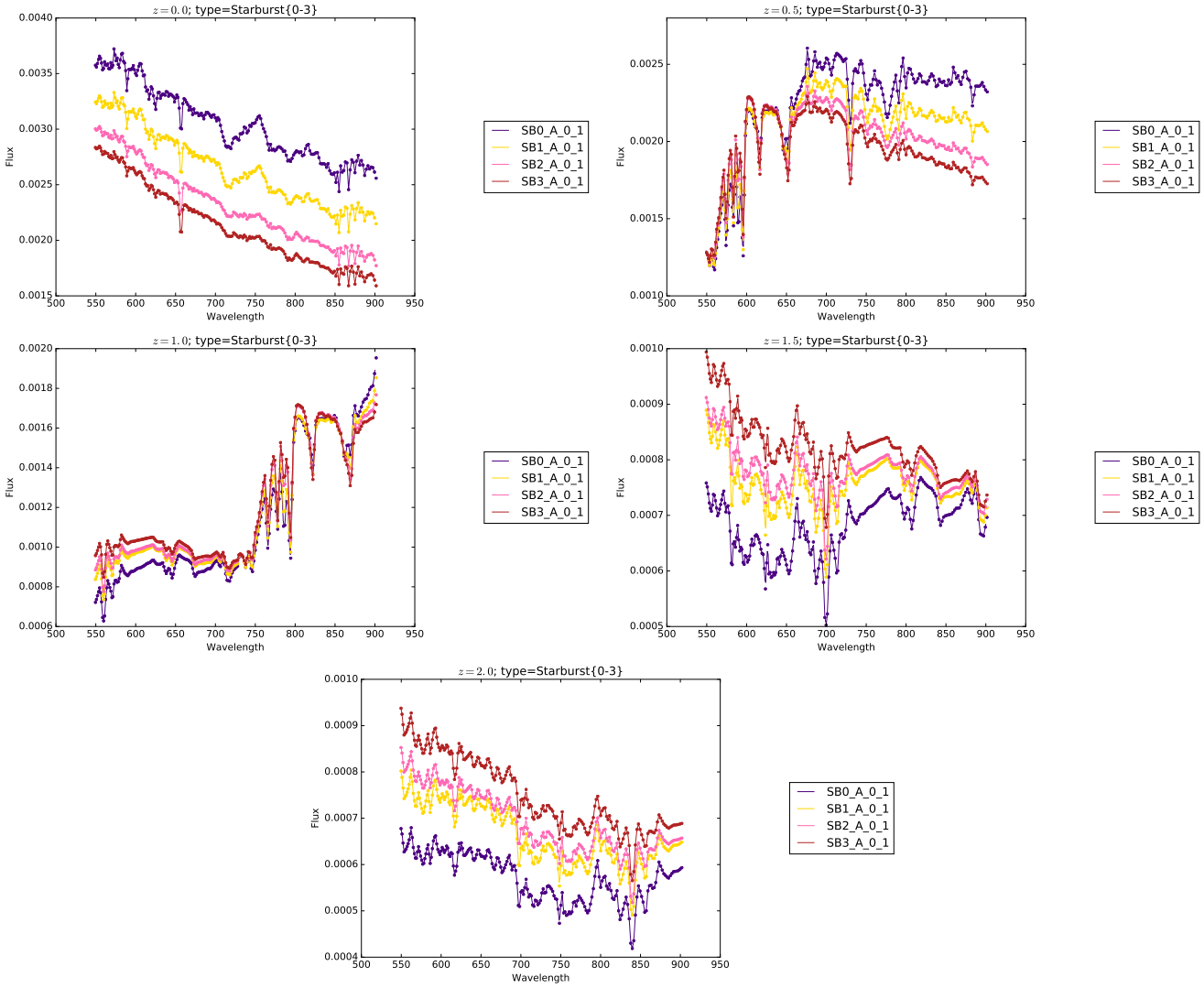


Figure I.2: Template SEDs for the first group of Starburst galaxies.

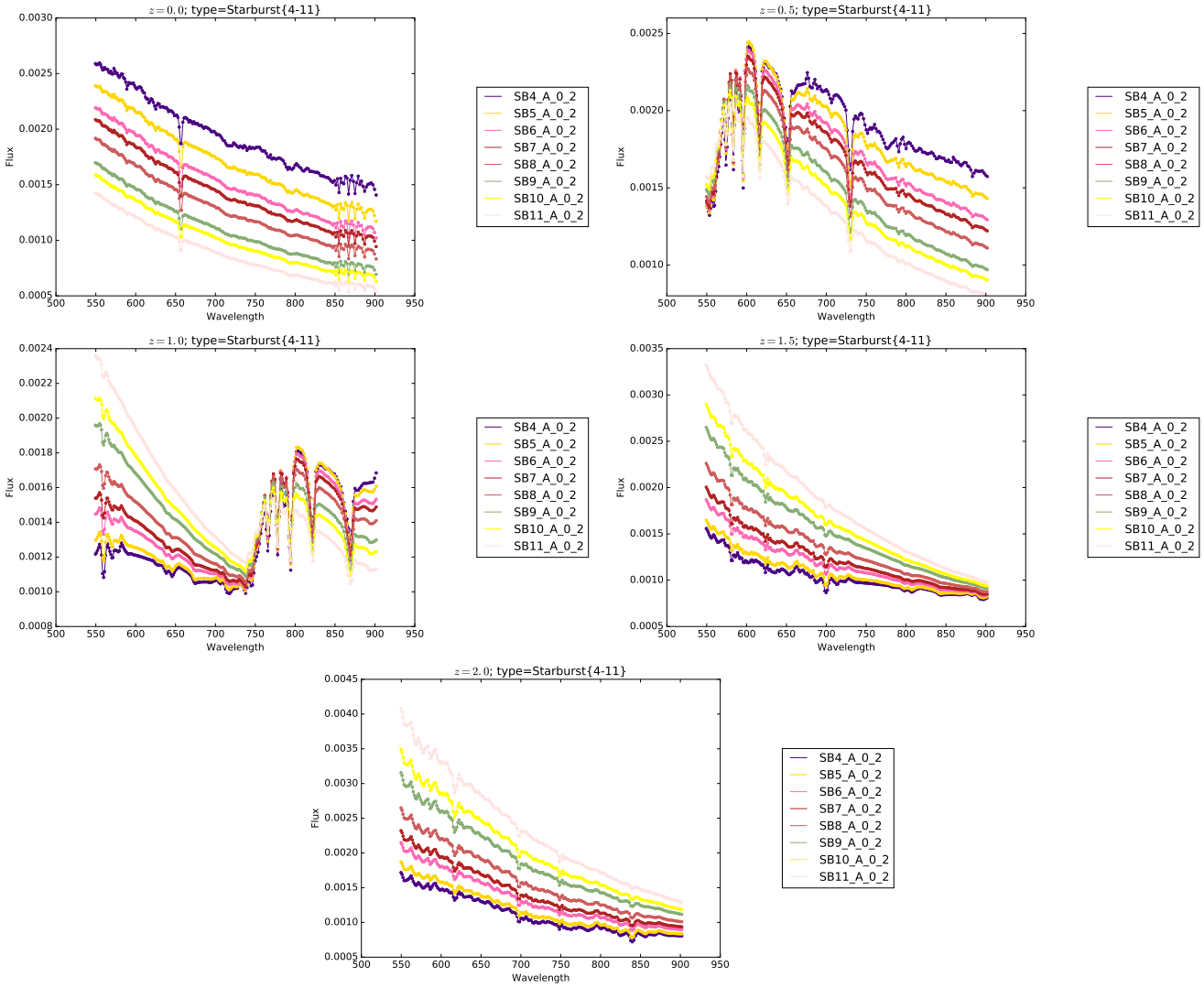


Figure I.3: Template SEDs for the second group of Starburst galaxies.

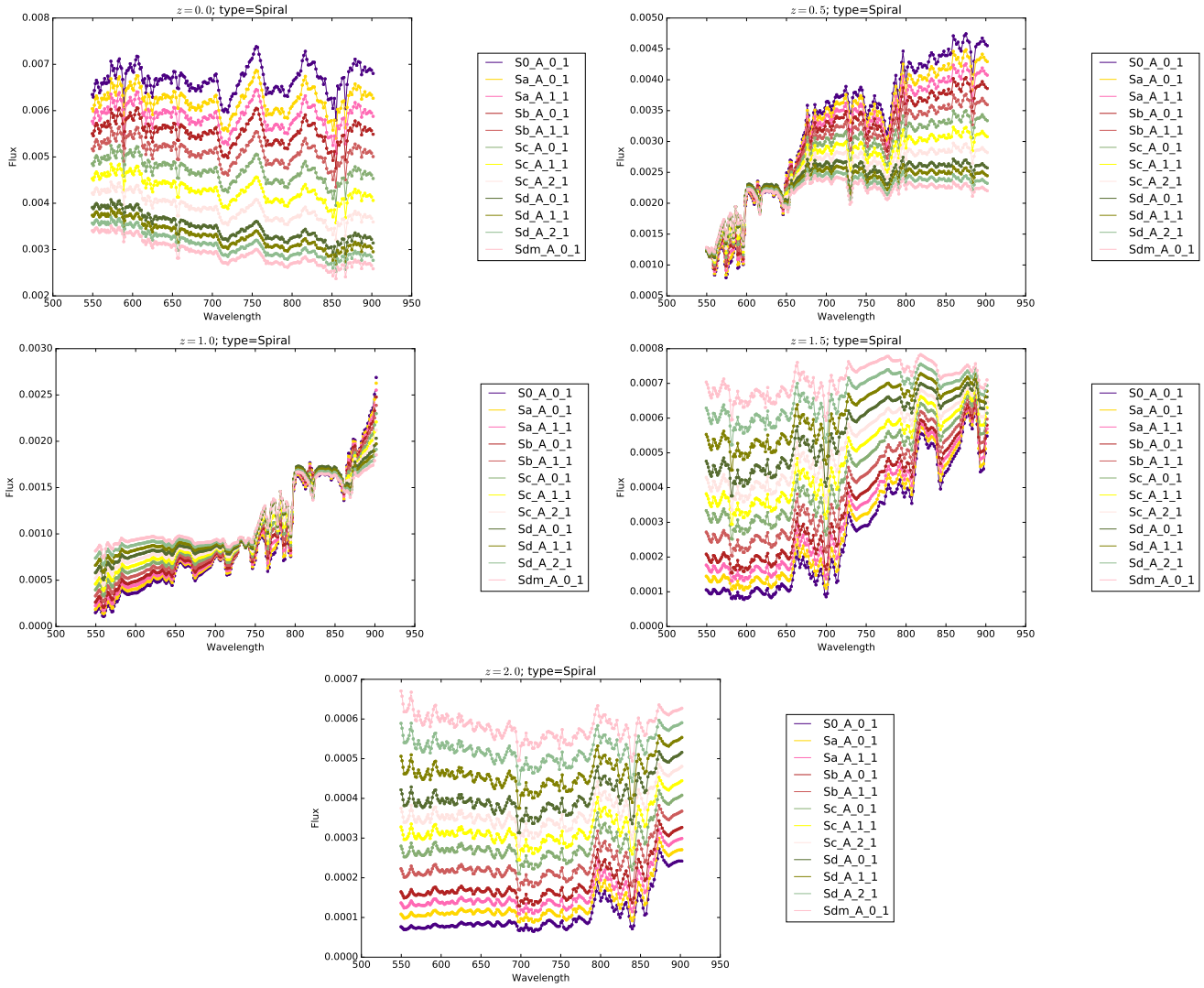


Figure I.4: Template SEDs for the spiral and lenticular galaxies.

Acronyms & Abbreviations

1D 1-dimensional.

2D 2-dimensional.

3D 3-dimensional.

2PCF two-point correlation function.

ACS Advanced Camera for Surveys.

ANN artificial neural network.

AOCS Attitude and Orbit Control System.

BAO Baryon Acoustic Oscillations.

BFD Bayesian Fourier Domain.

BFE Brighter-Fatter effect.

BFGS Broyden-Fletcher-Goldfarb-Shanno.

BTC Big Throughput Camera.

CCD Charge-Coupled Device.

CDFS Chandra Deep Field South.

CDM Cold Dark Matter.

CEA Commissariat à l'énergie atomique et aux énergies alternatives.

CFHT Canada-France-Hawaii Telescope.

CFHTLenS CFHT Lensing Survey.

CFHTLS CFHT Legacy Survey.

CFIS Canada France Imaging Survey.

CMB Cosmic Microwave Background.

COMBO Classifying Objects by Medium-Band Observations.

COSMOS COSmic evOlution Survey.

CSS-OS Chinese Space Station Optical Survey.

CTI Charge Transfer Inefficiency.

DECam Dark Energy Camera.

DES Dark Energy Survey.

DESCART Dark matter from Ellipticity Sources CARTography.

DLS Deep Lens Survey.

EC Euclid Consortium.

EDF Euclid Deep Field.

ELAIS European Large-Area Infrared Space observatory.

EOS Equation of State.

EPFL Ecole Polytechnique Fédérale de Lausanne.

ES1 ELAIS South 1.

ESA European Space Agency.

ESI Echelle Spectrograph and Imager.

FFT Fast Fourier Transform.

FLRW Friedmann-Lemaître-Robertson-Walker.

FORS FOcal Reducer and low dispersion Spectrograph.

FOV field of view.

FWHM full width at half-maximum.

GaBoDS Garching-Bonn Deep Survey.

GEMS Galaxy Evolution From Morphology And SEDs.

GR General Relativity.

GREAT GRavitational lEnsing Accuracy Testing.

HSC Hyper Suprime-Cam.

HSM Hirata-Seljak-Mandelbaum.

HST Hubble Space Telescope.

KiDS Kilo Degree Survey.

KL Kullback-Leibler.

KSB Kaiser-Squires-Broadhurst.

L-BFGS Limited memory BFGS.

LMC Large Magellanic Cloud.

LSS Large Scale Structure.

LSST Large Synoptic Survey Telescope.

MNIST Modified National Institute of Standards and Technology database.

MPG-ESO Max Planck Gesellschaft-European Southern Observatory.

MSE mean square error.

MUG Multimedia Understanding Groups database.

NISP Near Infrared Spectrometer and Photometer.

NMF Non-negative Matrix Factorization.

OT Optimal Transport.

PCA Principal Components Analysis.

PDF probabilty density function.

PFIC Prime Focus Imaging Camera.

photo-z photometric redshift.

PPS Primordial Power Spectrum.

PSF Point Spread Function.

PSNR Peak SNR.

RBF Radial Basis Function.

RCA Resolved Components Analysis.

RCS Red-sequence Cluster Survey.

RMS root mean square.

SDSS Sloan Digital Sky Survey.

SED spectral energy distribution.

SNR Signal-to-Noise Ratio.

SSP Strategic Subaru Proposal.

STIS Space Telescope Imaging Spectrograph.

SVD spectral value decomposition.

UH University of Hawai‘i.

VIRMOS Visible and InfraRed Multi-Object Spectrographs.

VIS Euclid Visible instrument.

VLT Very Large Telescope.

VOICE VST Optical Imaging of the CDFS and ES1 fields.

VST VLT Survey Telescope.

WDL Wasserstein Dictionary Learning.

WFE wavefront error.

WFI Wide Field Instrument.

WFI (ESO) Wide Field Imager.

WFIRST Wide Field Infrared Survey Telescope.

WFPC2 Wide Field Planetary Camera 2.

WHT William Herschel Telescope.

WL weak (gravitational) lensing.

WPG Wasserstein Principal Geodesic.

Bibliography

Agueh, M. and Carlier, G. (2011). Barycenters in the Wasserstein space. *SIAM Journal on Mathematical Analysis*, 43(2):904–924.

Aharon, M., Elad, M., and Bruckstein, A. (2006). K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on signal processing*, 54(11):4311–4322.

Aifanti, N., Papachristou, C., and Delopoulos, A. (2010). The MUG facial expression database. In *Image Analysis for Multimedia Interactive Services (WIAMIS), 2010 11th International Workshop on*, pages 1–4. IEEE.

Aihara, H., Arimoto, N., Armstrong, R., Arnouts, S., Bahcall, N. A., Bickerton, S., Bosch, J., Bundy, K., Capak, P. L., Chan, J. H., et al. (2017). The Hyper Suprime-Cam SSP survey: overview and survey design. *Publications of the Astronomical Society of Japan*, 70(SP1):S4.

Albrecht, A., Bernstein, G., Cahn, R., Freedman, W. L., Hewitt, J., Hu, W., Huth, J., Kamionkowski, M., Kolb, E. W., Knox, L., et al. (2006). Report of the dark energy task force. *arXiv preprint astro-ph/0609591*.

Altschuler, J., Weed, J., and Rigollet, P. (2017). Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration. In *Advances in Neural Information Processing Systems*, pages 1964–1974.

Arjovsky, M., Chintala, S., and Bottou, L. (2017). Wasserstein Generative Adversarial Networks. In *International conference on machine learning*, pages 214–223.

Armijo, L. (1966). Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of mathematics*, 16(1):1–3.

Bacon, D. J., Massey, R. J., Refregier, A. R., and Ellis, R. S. (2003). Joint cosmic shear measurements with the Keck and William Herschel telescopes. *Monthly Notices of the Royal Astronomical Society*, 344(3):673–685.

Bacon, D. J., Refregier, A. R., and Ellis, R. S. (2000). Detection of weak gravitational lensing by large-scale structure. *Monthly Notices of the Royal Astronomical Society*, 318(2):625–640.

Bardeen, J. M. (1980). Gauge-invariant cosmological perturbations. *Physical Review D*, 22(8):1882.

Bartelmann, M. and Schneider, P. (2001). Weak gravitational lensing. *Physics Reports*, 340(4-5):291–472.

Bassetti, F., Bodini, A., and Regazzini, E. (2006). On minimum Kantorovich distance estimators. *Statistics & probability letters*, 76(12):1298–1302.

Bassetti, F. and Regazzini, E. (2006). Asymptotic properties and robustness of minimum dissimilarity estimators of location-scale parameters. *Theory of Probability & Its Applications*, 50(2):171–186.

Benamou, J.-D., Carlier, G., Cuturi, M., Nenna, L., and Peyré, G. (2015). Iterative Bregman projections for regularized transportation problems. *SIAM Journal on Scientific Computing*, 37(2):A1111–A1138.

Benjamin, J., Heymans, C., Semboloni, E., Van Waerbeke, L., Hoekstra, H., Erben, T., Gladders, M. D., Hetterscheidt, M., Mellier, Y., and Yee, H. (2007). Cosmological constraints from the 100-deg^2 weak-lensing survey. *Monthly Notices of the Royal Astronomical Society*, 381(2):702–712.

Bergé, J., Price, S., Amara, A., and Rhodes, J. (2012). On point spread function modelling: towards optimal interpolation. *Monthly Notices of the Royal Astronomical Society*, 419(3):2356–2368.

Bernardeau, F. (1998). The effects of source clustering on weak lensing statistics. *Astronomy & Astrophysics*, 338:375–382.

Bernardeau, F., Colombi, S., Gaztanaga, E., and Scoccimarro, R. (2002). Large-scale structure of the universe and cosmological perturbation theory. *Physics reports*, 367(1-3):1–248.

Bernstein, G. and Jarvis, M. (2002). Shapes and shears, stars and smears: optimal measurements for weak lensing. *The Astronomical Journal*, 123(2):583.

Bernstein, G. M. and Armstrong, R. (2014). Bayesian lensing shear measurement. *Monthly Notices of the Royal Astronomical Society*, 438(2):1880–1893.

Bernstein, G. M., Armstrong, R., Krawiec, C., and March, M. C. (2016). An accurate and practical method for inference of weak gravitational lensing from galaxy images. *Monthly Notices of the Royal Astronomical Society*, 459(4):4467–4484.

Bernton, E., Jacob, P. E., Gerber, M., and Robert, C. P. (2017). On parameter estimation with the Wasserstein distance. *Information and Inference: A Journal of the IMA*.

Bertin, E. (2011). Automated Morphometry with SExtractor and PSFEx. In Evans, I. N., Accomazzi, A., Mink, D. J., and Rots, A. H., editors, *Astronomical Data Analysis Software and Systems XX*, volume 442 of *Astronomical Society of the Pacific Conference Series*, page 435.

Bertin, E. and Arnouts, S. (1996). SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series*, 117(2):393–404.

Bertsekas, D. P. (1988). The Auction algorithm: A distributed relaxation method for the assignment problem. *Annals of operations research*, 14(1):105–123.

Bigot, J., Gouet, R., Klein, T., López, A., et al. (2017). Geodesic PCA in the Wasserstein space by convex PCA. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 53, pages 1–26. Institut Henri Poincaré.

Blei, D. M. and Lafferty, J. D. (2009). Topic models. In *Text Mining*, pages 101–124. Chapman and Hall/CRC.

Boissard, E., Le Gouic, T., Loubes, J.-M., et al. (2015). Distribution template estimate with Wasserstein metrics. *Bernoulli*, 21(2):740–759.

Bolte, J., Daniilidis, A., Ley, O., and Mazet, L. (2010). Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. *Transactions of the American Mathematical Society*, 362(6):3319–3363.

Bonneel, N., Peyré, G., and Cuturi, M. (2016). Wasserstein barycentric coordinates: Histogram regression using optimal transport. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2016)*, 35(4).

Bonneel, N., Rabin, J., Peyré, G., and Pfister, H. (2015). Sliced and Radon Wasserstein Barycenters of Measures. *Journal of Mathematical Imaging and Vision*, 51(1):22–45.

Boone, K., Aldering, G., Copin, Y., Dixon, S., Domagalski, R., Gangler, E., Pecontal, E., and Perlmutter, S. (2018). A binary offset effect in CCD readout and its impact on astronomical data. *Publications of the Astronomical Society of the Pacific*, 130(988):064504.

Bosch, J., Armstrong, R., Bickerton, S., Furusawa, H., Ikeda, H., Koike, M., Lupton, R., Mineo, S., Price, P., Takata, T., et al. (2017). The Hyper Suprime-Cam software pipeline. *Publications of the Astronomical Society of Japan*, 70(SP1):S5.

Boucaud, A., Huertas-Company, M., Heneka, C., Ishida, E. E., Sedaghat, N., de Souza, R. S., Moews, B., Dole, H., Castellano, M., Merlin, E., et al. (2019). Photometry of high-redshift blended galaxies using deep learning. *arXiv preprint arXiv:1905.01324*.

Bridle, S., Balan, S. T., Bethge, M., Gentile, M., Harmeling, S., Heymans, C., Hirsch, M., Hosseini, R., Jarvis, M., Kirk, D., et al. (2010). Results of the GREAT08 challenge: An image analysis competition for cosmological lensing. *Monthly Notices of the Royal Astronomical Society*, 405(3):2044–2061.

Bridle, S., Shawe-Taylor, J., Amara, A., Applegate, D., Balan, S. T., Berge, J., Bernstein, G., Dahle, H., Erben, T., Gill, M., et al. (2009). Handbook for the GREAT08 challenge: An image analysis competition for cosmological lensing. *The Annals of Applied Statistics*, pages 6–37.

Brown, A., Vallenari, A., Prusti, T., De Bruijne, J., Babusiaux, C., Bailer-Jones, C., Biermann, M., Evans, D. W., Eyer, L., Jansen, F., et al. (2018). Gaia Data Release 2-Summary of the contents and survey properties. *Astronomy & Astrophysics*, 616:A1.

Brown, M. L., Taylor, A. N., Bacon, D. J., Gray, M. E., Dye, S., Meisenheimer, K., and Wolf, C. (2003). The shear power spectrum from the COMBO-17 survey. *Monthly Notices of the Royal Astronomical Society*, 341(1):100–118.

Buhmann, M. D. (2003). *Radial basis functions: theory and implementations*, volume 12. Cambridge University Press.

Burd, A., Cwiok, M., Czyrkowski, H., Dabrowski, R., Dominik, W., Grajda, M., Husejko, M., Jegier, M., Kalicki, A., Kasprowicz, G., et al. (2005). Pi of the Sky–all-sky, real-time search for fast optical transients. *New Astronomy*, 10(5):409–416.

Calabretta, M. R. and Greisen, E. W. (2002). Representations of celestial coordinates in FITS. *Astronomy & Astrophysics*, 395(3):1077–1122.

Carlier, G., Oberman, A., and Oudet, E. (2015). Numerical methods for matching for teams and Wasserstein barycenters. *ESAIM: Mathematical Modelling and Numerical Analysis*, 49(6):1621–1642.

Carlsten, S., Strauss, M. A., Lupton, R. H., Meyers, J. E., and Miyazaki, S. (2018). Wavelength-dependent PSFs and their impact on weak lensing measurements. *Monthly Notices of the Royal Astronomical Society*, 479(2):1491–1504.

Casertano, S., Ratnatunga, K. U., and Griffiths, R. E. (2003). Cosmic gravitational shear from the Hubble Space Telescope Medium Deep Survey. *The Astrophysical Journal Letters*, 598(2):L71.

Chang, C., Marshall, P., Jernigan, J., Peterson, J., Kahn, S., Gull, S., AlSayyad, Y., Ahmad, Z., Bankert, J., Bard, D., et al. (2012). Atmospheric point spread function interpolation for weak lensing in short exposure imaging data. *Monthly Notices of the Royal Astronomical Society*, 427(3):2572–2587.

Chizat, L., Peyré, G., Schmitzer, B., and Vialard, F.-X. (2018). Scaling algorithms for unbalanced optimal transport problems. *Mathematics of Computation*, 87(314):2563–2609.

Clowe, D., Markevitch, M., Bradač, M., Gonzalez, A. H., Chung, S. M., Massey, R., and Zaritsky, D. (2012). On dark peaks and missing mass: a weak-lensing mass reconstruction of the merging cluster system A520. *The Astrophysical Journal*, 758(2):128.

Collister, A. A. and Lahav, O. (2004). ANNz: estimating photometric redshifts using artificial neural networks. *Publications of the Astronomical Society of the Pacific*, 116(818):345.

Combettes, P. L. and Pesquet, J.-C. (2011). Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer.

Condat, L. (2013). A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *Journal of Optimization Theory and Applications*, 158(2):460–479.

Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of computation*, 19(90):297–301.

Cooray, A. and Hu, W. (2001). Weak gravitational lensing bispectrum. *The Astrophysical Journal*, 548(1):7.

Coulton, W. R., Armstrong, R., Smith, K. M., Lupton, R. H., and Spergel, D. N. (2018). Exploring the brighter-fatter effect with the Hyper Suprime-Cam. *The Astrophysical Journal*, 155(6):258.

Courty, N., Flamary, R., and Ducoffe, M. (2017). Learning Wasserstein embeddings. *arXiv preprint arXiv:1710.07457*.

Cropper, M., Hoekstra, H., Kitching, T., Massey, R., Amiaux, J., Miller, L., Mellier, Y., Rhodes, J., Rowe, B., Pires, S., et al. (2013). Defining a weak lensing experiment in space. *Monthly Notices of the Royal Astronomical Society*, 431(4):3103–3126.

Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300.

Cuturi, M. and Doucet, A. (2014). Fast computation of Wasserstein barycenters. In *Proceedings of The 31st International Conference on Machine Learning*, pages 685–693.

Cuturi, M. and Peyré, G. (2016). A smoothed dual approach for variational Wasserstein problems. *SIAM Journal on Imaging Sciences*, 9(1):320–343.

Cypriano, E., Amara, A., Voigt, L., Bridle, S., Abdalla, F., Refregier, A., Seiffert, M., and Rhodes, J. (2010). Cosmic shear requirements on the wavelength dependence of telescope point spread functions. *Monthly Notices of the Royal Astronomical Society*, 405(1):494–502.

d'Aspremont, A., El Ghaoui, L., Jordan, M. I., and Lanckriet, G. R. (2007). A direct formulation for Sparse PCA using semidefinite programming. *SIAM review*, 49(3):434–448.

Davis, C. P., Rodriguez, J., and Roodman, A. (2016). Wavefront-based PSF estimation. In *Ground-based and Airborne Telescopes VI*, volume 9906, page 990668. International Society for Optics and Photonics.

de Jong, J. T., Kleijn, G. A. V., Kuijken, K. H., Valentijn, E. A., et al. (2013). The Kilo-Degree Survey. *Experimental Astronomy*, 35(1-2):25–44.

De Karman, T. and Howarth, L. (1938). On the statistical theory of isotropic turbulence. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, 164(917):192–215.

Deming, W. E. and Stephan, F. F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, 11(4):427–444.

DES Collaboration et al. (2005). The Dark Energy Survey. *International Journal of Modern Physics A*, 20(14):3121–3123.

Desai, S., Mohr, J., Bertin, E., Kümmel, M., and Wetzstein, M. (2016). Detection and removal of artifacts in astronomical images. *Astronomy and computing*, 16:67–78.

Deshpande, A. C., Kitching, T. D., Cardone, V., Taylor, P. L., Paykari, P., and Casas, S. (in prep.). On the reduced shear approximation and magnification bias for Stage IV cosmic shear experiments (working title).

Dodd, M. S., Papineau, D., Grenne, T., Slack, J. F., Rittner, M., Pirajno, F., O'neil, J., and Little, C. T. (2017). Evidence for early life in Earth's oldest hydrothermal vent precipitates. *Nature*, 543(7643):60.

Dodelson, S. (2003). *Modern Cosmology*. Elsevier.

Donoho, D. L. and Huo, X. (2001). Uncertainty principles and ideal atomic decomposition. *IEEE transactions on information theory*, 47(7):2845–2862.

Drlica-Wagner, A., Sevilla-Noarbe, I., Rykoff, E. S., Gruendl, R., Yanny, B., Tucker, D., Hoyle, B., Rosell, A. C., Bernstein, G., Bechtol, K., et al. (2018). Dark Energy Survey Year 1 results: the photometric data set for cosmology. *The Astrophysical Journal Supplement Series*, 235(2):33.

Duncan, C., Miller, L., Schreiber, C., et al. (in prep.). Euclid: PSF modelling of the VIS instrument (official title TBA).

Dunteman, G. H. (1989). *Principal Components Analysis*. Number 69. Sage.

Dyson, F. W., Eddington, A. S., and Davidson, C. (1920). A determination of the deflection of light by the Sun's gravitational field, from observations made at the total eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 220:291–333.

Einstein, A. (1916). The foundation of the general theory of relativity. *Annalen Phys.*, 14:769–822.

Eisenstein, D. J. and Hu, W. (1998). Baryonic features in the matter transfer function. *The Astrophysical Journal*, 496(2):605.

Elad, M. and Hel-Or, Y. (2001). A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur. *IEEE Transactions on Image Processing*, 10(8):1187–1193.

Er, X., Hoekstra, H., Schrabback, T., Cardone, V. F., Scaramella, R., Maoli, R., Vicinanza, M., Gillis, B., and Rhodes, J. (2018). Calibration of colour gradient bias in shear measurement using HST/CANDELS data. *Monthly Notices of the Royal Astronomical Society*, 476(4):5645–5657.

Erben, T., Hildebrandt, H., Miller, L., van Waerbeke, L., Heymans, C., Hoekstra, H., Kitching, T., Mellier, Y., Benjamin, J., Blake, C., et al. (2013). CFHTLenS: the Canada–France–Hawaii Telescope Lensing Survey–imaging data and catalogue products. *Monthly Notices of the Royal Astronomical Society*, 433(3):2545–2563.

Eriksen, M. and Hoekstra, H. (2018). Implications of a wavelength-dependent PSF for weak lensing measurements. *Monthly Notices of the Royal Astronomical Society*, 477(3):3433–3448.

Erlander, S. and Stewart, N. F. (1990). *The gravity model in transportation analysis: theory and extensions*, volume 3. Vsp.

Euclid Collaboration, Blanchard, A., Camera, S., Carbone, C., Cardone, V. F., Casas, S., Ilić, S., Kilbinger, M., Kitching, T., Kunz, M., Lacasa, F., Linder, E., Majerotto, E., Markovič, K., Martinelli, M., Pettorino, V., Pourtsidou, A., Sakr, Z., Sánchez, A. G., Sapone, D., Tutusaus, I., Yahia-Cherif, S., Yankelevich, V., et al. (2019). Euclid preparation: VII. Forecast validation for Euclid cosmological probes. *arXiv e-prints*, page arXiv:1910.09273.

Farrens, S., Mboula, F. N., and Starck, J.-L. (2017). Space variant deconvolution of galaxy survey images. *Astronomy & Astrophysics*, 601:A66.

Fletcher, P. T., Lu, C., Pizer, S. M., and Joshi, S. (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23(8):995–1005.

Franklin, J. and Lorenz, J. (1989). On the scaling of multidimensional matrices. *Linear Algebra and its applications*, 114:717–735.

Fréchet, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. In *Annales de l'institut Henri Poincaré*, volume 10, pages 215–310. Presses universitaires de France.

Frogner, C., Zhang, C., Mobahi, H., Araya, M., and Poggio, T. A. (2015). Learning with a Wasserstein loss. In *Advances in Neural Information Processing Systems*, pages 2053–2061.

Fruchter, A. and Hook, R. (2002). Drizzle: A method for the linear reconstruction of undersampled images. *Publications of the Astronomical Society of the Pacific*, 114(792):144.

Fu, L. and Kilbinger, M. (2010). A new cosmic shear function: optimized E-/B-mode decomposition on a finite interval. *Monthly Notices of the Royal Astronomical Society*, 401(2):1264–1274.

Fu, L., Liu, D., Radovich, M., Liu, X., Pan, C., Fan, Z., Covone, G., Vaccari, M., Amaro, V., Brescia, M., et al. (2018). Weak-lensing study in VOICE survey–I. Shear measurement. *Monthly Notices of the Royal Astronomical Society*, 479(3):3858–3872.

Fu, L., Semboloni, E., Hoekstra, H., Kilbinger, M., Van Waerbeke, L., Tereno, I., Mellier, Y., Heymans, C., Coupon, J., Benabed, K., et al. (2008). Very weak lensing in the CFHTLS wide: cosmology from cosmic shear in the linear regime. *Astronomy & astrophysics*, 479(1):9–25.

Gao, W., Chen, J., Richard, C., and Huang, J. (2014). Online dictionary learning for kernel LMS. *IEEE Transactions on Signal Processing*, 62(11):2765–2777.

Genevay, A., Peyre, G., and Cuturi, M. (2018). Learning generative models with Sinkhorn divergences. In *International Conference on Artificial Intelligence and Statistics*, pages 1608–1617.

Gentile, M., Courbin, F., and Meylan, G. (2013). Interpolating point spread function anisotropy. *Astronomy & Astrophysics*, 549:A1.

Giblin, J. T. J., Mertens, J. B., Starkman, G. D., and Zentner, A. R. (2017). General relativistic corrections to the weak lensing convergence power spectrum. *Physical Review D*, 96(10):103530.

Gladders, M. D. and Yee, H. K. (2005). The Red-Sequence Cluster Survey. I. The survey and cluster catalogs for patches RCS 0926+37 and RCS 1327+29. *The Astrophysical Journal Supplement Series*, 157(1):1.

Gong, Y., Liu, X., Cao, Y., Chen, X., Fan, Z., Li, R., Li, X.-D., Li, Z., Zhang, X., and Zhan, H. (2019). Cosmology from the Chinese Space Station Optical Survey (CSS-OS). *arXiv preprint arXiv:1901.04634*.

Green, J., Schechter, P., Baltay, C., Bean, R., Bennett, D., Brown, R., Conselice, C., Donahue, M., Fan, X., Gaudi, B., et al. (2012). Wide-Field Infrared Survey Telescope (WFIRST) final report. *arXiv preprint arXiv:1208.4012*.

Greisen, E. W. and Calabretta, M. R. (2002). Representations of world coordinates in FITS. *Astronomy & Astrophysics*, 395(3):1061–1075.

Griewank, A. and Walther, A. (2008). *Evaluating derivatives: principles and techniques of algorithmic differentiation*. SIAM.

Guinot, A., Kilbinger, M., Farrens, S., Schmitz, M. A., Pujol, A., Peel, A., Starck, J.-L., Sureau, F., and Gentile, M. (in prep.). ShapePipe: a new shape measurement pipeline (working title).

Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, pages 49–52.

Haker, S., Zhu, L., Tannenbaum, A., and Angenent, S. (2004). Optimal mass transport for registration and warping. *International Journal of Computer Vision*, 60(3):225–240.

Hamana, T., Colombi, S. T., Thion, A., Devriendt, J. E., Mellier, Y., and Bernardeau, F. (2002). Source-lens clustering effects on the skewness of the lensing convergence. *Monthly Notices of the Royal Astronomical Society*, 330(2):365–377.

Hamana, T., Miyazaki, S., Shimasaku, K., Furusawa, H., Doi, M., Hamabe, M., Imi, K., Kimura, M., Komiyama, Y., Nakata, F., et al. (2003). Cosmic shear statistics in the Suprime-Cam 2.1 square degree field: constraints on Ω_m and σ_8. *The Astrophysical Journal*, 597(1):98.

Hämmerle, H., Miralles, J.-M., Schneider, P., Erben, T., Fosbury, R., Freudling, W., Pirzkal, N., Jain, B., and White, S. (2002). Cosmic shear from STIS pure parallels-II. Analysis. *Astronomy & Astrophysics*, 385(3):743–760.

Hammond, D. K., Vandergheynst, P., and Gribonval, R. (2011). Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150.

Harandi, M. and Salzmann, M. (2015). Riemannian coding and dictionary learning: Kernels to the rescue. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3926–3935.

Harandi, M., Sanderson, C., Shen, C., and Lovell, B. C. (2013). Dictionary learning and sparse coding on Grassmann manifolds: An extrinsic solution. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3120–3127.

Harandi, M. T., Sanderson, C., Hartley, R., and Lovell, B. C. (2012). Sparse coding and dictionary learning for symmetric positive definite matrices: A kernel approach. In *Computer Vision–ECCV 2012*, pages 216–229. Springer.

Harrison, E. R. (1970). Fluctuations at the threshold of classical cosmology. *Physical review D*, 1(10):2726.

Heavens, A. (2003). 3D weak lensing. *Monthly Notices of the Royal Astronomical Society*, 343(4):1327–1334.

Herbel, J., Kacprzak, T., Amara, A., Refregier, A., and Lucchi, A. (2018). Fast point spread function modeling with deep learning. *Journal of Cosmology and Astroparticle Physics*, 2018(07):054.

Hetterscheidt, M., Simon, P., Schirmer, M., Hildebrandt, H., Schrabback, T., Erben, T., and Schneider, P. (2007). GaBoDS: The Garching-Bonn Deep Survey-VII. Cosmic shear analysis. *Astronomy & Astrophysics*, 468(3):859–876.

Heymans, C., Brown, M. L., Barden, M., Caldwell, J. A., Jahnke, K., Peng, C. Y., Rix, H.-W., Taylor, A., Beckwith, S. V., Bell, E. F., et al. (2005). Cosmological weak lensing with the HST GEMS survey. *Monthly Notices of the Royal Astronomical Society*, 361(1):160–176.

Heymans, C., Rowe, B., Hoekstra, H., Miller, L., Erben, T., Kitching, T., and Van Waerbeke, L. (2012a). The impact of high spatial frequency atmospheric distortions on weak-lensing measurements. *Monthly Notices of the Royal Astronomical Society*, 421(1):381–389.

Heymans, C., Van Waerbeke, L., Bacon, D., Berge, J., Bernstein, G., Bertin, E., Bridle, S., Brown, M. L., Clowe, D., Dahle, H., et al. (2006). The Shear Testing Programme–I. Weak lensing analysis of simulated ground-based observations. *Monthly Notices of the Royal Astronomical Society*, 368(3):1323–1339.

Heymans, C., Van Waerbeke, L., Miller, L., Erben, T., Hildebrandt, H., Hoekstra, H., Kitching, T. D., Mellier, Y., Simon, P., Bonnett, C., et al. (2012b). CFHTLenS: the Canada–France–Hawaii Telescope Lensing Survey. *Monthly Notices of the Royal Astronomical Society*, 427(1):146–166.

Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507.

Hirata, C. and Seljak, U. (2003). Shear calibration biases in weak-lensing surveys. *Monthly Notices of the Royal Astronomical Society*, 343(2):459–480.

Ho, J., Xie, Y., and Vemuri, B. (2013). On a nonlinear generalization of sparse coding and dictionary learning. In *International conference on machine learning*, pages 1480–1488.

Hoekstra, H. (2004). The effect of imperfect models of point spread function anisotropy on cosmic shear measurements. *Monthly Notices of the Royal Astronomical Society*, 347(4):1337–1344.

Hoekstra, H., Franx, M., Kuijken, K., and Squires, G. (1998). Weak lensing analysis of CL 1358+62 using Hubble Space Telescope observations. *The Astrophysical Journal*, 504(2):636.

Hoekstra, H., Herbonnet, R., Muzzin, A., Babul, A., Mahdavi, A., Viola, M., and Cacciato, M. (2015). The Canadian Cluster Comparison Project: detailed study of systematics and updated weak lensing masses. *Monthly Notices of the Royal Astronomical Society*, 449(1):685–714.

Hoekstra, H., Viola, M., and Herbonnet, R. (2017). A study of the sensitivity of shape measurements to the input parameters of weak-lensing image simulations. *Monthly Notices of the Royal Astronomical Society*, 468(3):3295–3311.

Hoekstra, H., Yee, H. K., and Gladders, M. D. (2002a). Constraints on Ω_m and σ_8 from weak lensing in Red-Sequence Cluster Survey fields. *The Astrophysical Journal*, 577(2):595.

Hoekstra, H., Yee, H. K., Gladders, M. D., Barrientos, L. F., Hall, P. B., and Infante, L. (2002b). A measurement of weak lensing by large-scale structure in Red-Sequence Cluster Survey fields. *The Astrophysical Journal*, 572(1):55.

Hoffmann, S. and Anderson, J. (2017). A study of PSF models for ACS/WFC. *Instrument Science Report ACS 2017-8, 12 pages*.

Hu, W. (1999). Power spectrum tomography with weak lensing. *The Astrophysical Journal Letters*, 522(1):L21.

Hudelot, P., Goranova, Y., Mellier, Y., McCracken, H. J., Magnard, F., Monnerville, M., Smah, G., Cuillandre, J.-C., Withington, K., Regnault, N., et al. (2012). T0007: The final CFHTLS release.

Huff, E. and Mandelbaum, R. (2017). Metacalibration: Direct self-calibration of biases in shear measurement. *arXiv preprint arXiv:1702.02600*.

Huff, E. M., Hirata, C. M., Mandelbaum, R., Schlegel, D., Seljak, U., and Lupton, R. H. (2014). Seeing in the dark–I. Multi-epoch alchemy. *Monthly Notices of the Royal Astronomical Society*, 440(2):1296–1321.

Hyvärinen, A., Karhunen, J., and Oja, E. (2004). *Independent component analysis*, volume 46. John Wiley & Sons.

Ibata, R. A., McConnachie, A., Cuillandre, J.-C., Fantin, N., Haywood, M., Martin, N. F., Bergeron, P., Beckmann, V., Bernard, E., Bonifacio, P., et al. (2017). The Canada–France Imaging Survey: First results from the u-band component. *The Astrophysical Journal*, 848(2):128.

Ilbert, O., Arnouts, S., McCracken, H., Bolzonella, M., Bertin, E., Le Fèvre, O., Mellier, Y., Zamorani, G., Pellò, R., Iovino, A., et al. (2006). Accurate photometric redshifts for the CFHT Legacy Survey calibrated using the VIMOS VLT deep survey. *Astronomy & Astrophysics*, 457(3):841–856.

Ilbert, O., Capak, P., Salvato, M., Aussel, H., McCracken, H., Sanders, D., Scoville, N., Kartaltepe, J., Arnouts, S., Le Floc'H, E., et al. (2008). COSMOS photometric redshifts with 30-bands for 2-deg^2. *The Astrophysical Journal*, 690(2):1236.

Irace, Z. and Batatia, H. (2013). Motion-based interpolation to estimate spatially variant PSF in positron emission tomography. In *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, pages 1–5. IEEE.

Israel, H., Massey, R., Prod'homme, T., Cropper, M., Cordes, O., Gow, J., Kohley, R., Marggraf, O., Niemi, S., Rhodes, J., et al. (2015). How well can charge transfer inefficiency be corrected? A parameter sensitivity study for iterative correction. *Monthly Notices of the Royal Astronomical Society*, 453(1):561–580.

Israel, H., Massey, R., Prod'homme, T., Cropper, M., Cordes, O., Gow, J., Kohley, R., Marggraf, O., Niemi, S., Rhodes, J., et al. (2017). Erratum: How well can charge transfer inefficiency be corrected? A parameter sensitivity study for iterative correction. *Monthly notices of the Royal Astronomical Society.*, 467(4):4218–4219.

Jarvis, M., Bernstein, G., Fischer, P., Smith, D., Jain, B., Tyson, J., and Wittman, D. (2003). Weak-lensing results from the 75 square degree Cerro Tololo Inter-American Observatory survey. *The Astronomical Journal*, 125(3):1014.

Jarvis, M. and Jain, B. (2004). Principal component analysis of PSF variation in weak lensing surveys. *arXiv preprint astro-ph/0412234*.

Jarvis, M., Jain, B., Bernstein, G., and Dolney, D. (2006). Dark energy constraints from the CTIO lensing survey. *The Astrophysical Journal*, 644(1):71.

Jarvis, M., Schechter, P., and Jain, B. (2008). Telescope optics and weak lensing: PSF patterns due to low order aberrations. *arXiv preprint arXiv:0810.0027*.

Jarvis, M., Sheldon, E., Zuntz, J., Kacprzak, T., Bridle, S., Amara, A., Armstrong, R., Becker, M., Bernstein, G., Bonnett, C., et al. (2016). The DES Science Verification weak lensing shear catalogues. *Monthly Notices of the Royal Astronomical Society*, 460(2):2245–2281.

Jee, M., Blakeslee, J., Sirianni, M., Martel, A., White, R., and Ford, H. (2007). Principal component analysis of the time-and position-dependent point-spread function of the Advanced Camera for Surveys. *Publications of the Astronomical Society of the Pacific*, 119(862):1403.

Jee, M., Mahdavi, A., Hoekstra, H., Babul, A., Dalcanton, J., Carroll, P., and Capak, P. (2012). A study of the dark core in A520 with the Hubble Space Telescope: The mystery deepens. *The Astrophysical Journal*, 747(2):96.

Jee, M. J., Hoekstra, H., Mahdavi, A., and Babul, A. (2014). Hubble Space Telescope/Advanced Camera for Surveys confirmation of the dark substructure in A520. *The Astrophysical Journal*, 783(2):78.

Jee, M. J., Tyson, J. A., Schneider, M. D., Wittman, D., Schmidt, S., and Hilbert, S. (2013). Cosmic shear results from the Deep Lens Survey. I. Joint constraints on Ω_m and σ_8 with a two-dimensional analysis. *The Astrophysical Journal*, 765(1):74.

Joachimi, B., Cacciato, M., Kitching, T. D., Leonard, A., Mandelbaum, R., Schäfer, B. M., Sifón, C., Hoekstra, H., Kiessling, A., Kirk, D., et al. (2015). Galaxy alignments: An overview. *Space Science Reviews*, 193(1-4):1–65.

Jones, E., Oliphant, T., Peterson, P., et al. (2001). SciPy: Open source scientific tools for Python.

Joseph, R., Courbin, F., and Starck, J.-L. (2016). Multi-band morpho-spectral component analysis deblending tool (MuSCADeT): Deblending colourful objects. *Astronomy & Astrophysics*, 589:A2.

Joye, W. and Mandel, E. (2003). New features of SAOImage DS9. In *Astronomical data analysis software and systems XII*, volume 295, page 489.

Kaiser, N. (1992). Weak gravitational lensing of distant galaxies. *The Astrophysical Journal*, 388:272–286.

Kaiser, N. (2000). A new shear estimator for weak-lensing observations. *The Astrophysical Journal*, 537(2):555.

Kaiser, N. and Squires, G. (1993). Mapping the dark matter with weak gravitational lensing. *The Astrophysical Journal*, 404:441–450.

Kaiser, N., Squires, G., and Broadhurst, T. (1995). A method for weak lensing observations. *The Astrophysical Journal*, 449:460.

Kaiser, N., Wilson, G., and Luppino, G. A. (2000). Large-scale cosmic shear measurements. *arXiv preprint astro-ph/0003338*.

Kilbinger, M. (2015). Cosmology with cosmic shear observations: a review. *Reports on Progress in Physics*, 78(8):086901.

Kilbinger, M., Heymans, C., Asgari, M., Joudaki, S., Schneider, P., Simon, P., Van Waerbeke, L., Harnois-Déraps, J., Hildebrandt, H., Köhlinger, F., et al. (2017). Precision calculations of the cosmic shear power spectrum projection. *Monthly Notices of the Royal Astronomical Society*, 472(2):2126–2141.

Kitching, T., Amara, A., Gill, M., Harmeling, S., Heymans, C., Massey, R., Rowe, B., Schrabback, T., Voigt, L., Balan, S., et al. (2011). Gravitational lensing accuracy testing 2010 (GREAT10) challenge handbook. *The Annals of Applied Statistics*, pages 2231–2263.

Kitching, T., Balan, S., Bridle, S., Cantale, N., Courbin, F., Eifler, T., Gentile, M., Gill, M., Harmeling, S., Heymans, C., et al. (2012). Image analysis for cosmology: results from the GREAT10 galaxy challenge. *Monthly Notices of the Royal Astronomical Society*, 423(4):3163–3208.

Kitching, T., Miller, L., Heymans, C., Van Waerbeke, L., and Heavens, A. (2008). Bayesian galaxy shape measurement for weak lensing surveys–II. Application to simulations. *Monthly Notices of the Royal Astronomical Society*, 390(1):149–167.

Kitching, T., Rowe, B., Gill, M., Heymans, C., Massey, R., Witherick, D., Courbin, F., Georgatzis, K., Gentile, M., Gruen, D., et al. (2013). Image analysis for cosmology: results from the GREAT10 star challenge. *The Astrophysical Journal Supplement Series*, 205(2):12.

Kratochvil, J. M., Lim, E. A., Wang, S., Haiman, Z., May, M., and Huffenberger, K. (2012). Probing cosmology with weak lensing Minkowski functionals. *Physical Review D*, 85(10):103513.

Krause, E. and Hirata, C. M. (2010). Weak lensing power spectra for precision cosmology-multiple-deflection, reduced shear, and lensing bias corrections. *Astronomy & Astrophysics*, 523:A28.

Kreisch, C. D., Cyr-Racine, F.-Y., and Doré, O. (2019). The neutrino puzzle: Anomalies, interactions, and cosmological tensions. *arXiv preprint arXiv:1902.00534*.

Krist, J. (1995). Simulation of HST PSFs using Tiny Tim. In Shaw, R. A., Payne, H. E., and Hayes, J. J. E., editors, *Astronomical Data Analysis Software and Systems IV*, volume 77 of *Astronomical Society of the Pacific Conference Series*, page 349.

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97.

Kuijken, K. (1999). Weak weak lensing: correcting weak shear measurements accurately for PSF anisotropy. *arXiv preprint astro-ph/9904418*.

Kuijken, K., Heymans, C., Dvornik, A., Hildebrandt, H., de Jong, J. T. A., Wright, A. H., Erben, T., Bilicki, M., Giblin, B., Shan, H.-Y., Getman, F., Grado, A., Hoekstra, H., Miller, L., Napolitano, N., Paolilo, M., Radovich, M., Schneider, P., Sutherland, W., Tewes, M., Tortora, C., Valentijn, E. A., and Verdoes Kleijn, G. A. (2019). The fourth data release of the Kilo-Degree Survey: ugri imaging and nine-band optical-IR photometry over 1000 square degrees. *Astronomy & Astrophysics*, 625:A2.

Kuijken, K., Heymans, C., Hildebrandt, H., Nakajima, R., Erben, T., de Jong, J. T., Viola, M., Choi, A., Hoekstra, H., Miller, L., et al. (2015). Gravitational lensing analysis of the Kilo-Degree Survey. *Monthly Notices of the Royal Astronomical Society*, 454(4):3500–3532.

Kuntzer, T. and Courbin, F. (2017). Detecting unresolved binary stars in Euclid VIS images. *Astronomy & Astrophysics*, 606:A119.

Kuntzer, T., Tewes, M., and Courbin, F. (2016). Stellar classification from single-band imaging using machine learning. *Astronomy & Astrophysics*, 591:A54.

Kuntzer, T. A. (2018). *Machine-learning methods for weak lensing analysis of the ESA Euclid sky survey*. PhD thesis, EPFL.

Lauer, T. R. (1999). Combining undersampled dithered images. *Publications of the Astronomical Society of the Pacific*, 111(756):227.

Laureijs, R., Amiaux, J., Arduini, S., Augueres, J.-L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., et al. (2011). Euclid definition study report. *arXiv preprint arXiv:1110.3193*.

Leauthaud, A., Massey, R., Kneib, J.-P., Rhodes, J., Johnston, D. E., Capak, P., Heymans, C., Ellis, R. S., Koekemoer, A. M., Le Fevre, O., et al. (2007). Weak gravitational lensing with COSMOS: galaxy selection and shape measurements. *The Astrophysical Journal Supplement Series*, 172(1):219.

Lee, D. D. and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791.

Lee, H., Battle, A., Raina, R., and Ng, A. Y. (2007). Efficient sparse coding algorithms. In *Advances in neural information processing systems*, pages 801–808.

Leistedt, B., McEwen, J. D., Büttner, M., and Peiris, H. V. (2016). Wavelet reconstruction of E and B modes for CMB polarization and cosmic shear analyses. *Monthly Notices of the Royal Astronomical Society*, 466(3):3728–3740.

Lemaître, G. (1927). Un univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. In *Annales de la Société scientifique de Bruxelles*, volume 47, pages 49–59.

Léonard, C. (2014). A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems - Series A (DCDS-A)*, 34(4):1533–1574.

Lesgourgues, J. (2011). The cosmic linear anisotropy solving system (CLASS) I: Overview. *arXiv preprint arXiv:1104.2932*.

Lesgourgues, J. and Pastor, S. (2006). Massive neutrinos and cosmology. *Physics Reports*, 429(6):307–379.

Li, P., Wang, Q., Zuo, W., and Zhang, L. (2013). Log-Euclidean kernels for sparse representation and dictionary learning. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1601–1608.

Limber, D. N. (1953). The analysis of counts of the extragalactic nebulae in terms of a fluctuating density field. *The Astrophysical Journal*, 117:134.

Lin, C.-A. (2016). *Cosmology with weak-lensing peak counts*. PhD thesis, CEA Saclay.

Lin, H., Dodelson, S., Seo, H.-J., Soares-Santos, M., Annis, J., Hao, J., Johnston, D., Kubo, J. M., Reis, R. R., and Simet, M. (2012). The SDSS co-add: cosmic shear measurement. *The Astrophysical Journal*, 761(1):15.

Linder, E. V. (2003). Exploring the expansion history of the Universe. *Physical Review Letters*, 90(9):091301.

Liu, H., Qin, J., Cheng, H., and Sun, F. (2015). Robust kernel dictionary learning using a whole sequence convergent algorithm. In *IJCAI*, volume 1, page 5.

LoVerde, M. and Afshordi, N. (2008). Extended Limber approximation. *Physical Review D*, 78(12):123506.

LSST Science Collaboration, Abell, P. A., Allison, J., Anderson, S. F., Andrew, J. R., Angel, J. R. P., Armus, L., Arnett, D., Asztalos, S. J., Axelrod, T. S., et al. (2009). LSST Science Book, Version 2.0. *arXiv preprint arXiv:0912.0201*.

Lu, T., Zhang, J., Dong, F., Li, Y., Liu, D., Fu, L., Li, G., and Fan, Z. (2017). Testing PSF interpolation in weak lensing with real data. *The Astronomical Journal*, 153(4):197.

Lu, T., Zhang, J., Dong, F., Li, Y., Liu, D., Fu, L., Li, G., and Fan, Z. (2018). Removing the impact of correlated PSF uncertainties in weak lensing. *The Astrophysical Journal*, 858(2):122.

Luppino, G. and Kaiser, N. (1997). Detection of weak lensing by a cluster of galaxies at z=0.83. *The Astrophysical Journal*, 475(1):20.

Lupton, R., Gunn, J. E., Ivezic, Z., Knapp, G. R., Kent, S., and Yasuda, N. (2001). The SDSS imaging pipelines. *arXiv preprint astro-ph/0101420*.

Ma, Z., Bernstein, G., Weinstein, A., and Sholl, M. (2008). Diagnosing space telescope misalignment and jitter using stellar images. *Publications of the Astronomical Society of the Pacific*, 120(874):1307.

Magnier, E. and Cuillandre, J.-C. (2004). The Elixir system: Data characterization and calibration at the Canada-France-Hawaii Telescope. *Publications of the Astronomical Society of the Pacific*, 116(819):449.

Mahdavi, A., Hoekstra, H., Babul, A., Balam, D. D., and Capak, P. L. (2007). A dark core in Abell 520. *The Astrophysical Journal*, 668(2):806.

Mairal, J., Bach, F., Ponce, J., and Sapiro, G. (2010). Online learning for matrix factorization and sparse coding. *Journal of Machine Learning Research*, 11(Jan):19–60.

Mallat, S. (1999). *A wavelet tour of signal processing*. Elsevier.

Mandelbaum, R. (2018). Weak lensing for precision cosmology. *Annual Review of Astronomy and Astrophysics*, 56:393–433.

Mandelbaum, R., Hirata, C. M., Seljak, U., Guzik, J., Padmanabhan, N., Blake, C., Blanton, M. R., Lupton, R., and Brinkmann, J. (2005). Systematic errors in weak lensing: application to SDSS galaxy-galaxy weak lensing. *Monthly Notices of the Royal Astronomical Society*, 361(4):1287–1322.

Mandelbaum, R., Lanusse, F., Leauthaud, A., Armstrong, R., Simet, M., Miyatake, H., Meyers, J. E., Bosch, J., Murata, R., Miyazaki, S., et al. (2018). Weak lensing shear calibration with simulations of the HSC survey. *Monthly Notices of the Royal Astronomical Society*, 481(3):3170–3195.

Mandelbaum, R., Miyatake, H., Hamana, T., Oguri, M., Simet, M., Armstrong, R., Bosch, J., Murata, R., Lanusse, F., Leauthaud, A., et al. (2017). The first-year shear catalog of the Subaru Hyper Suprime-Cam Subaru Strategic Program Survey. *Publications of the Astronomical Society of Japan*, 70(SP1):S25.

Mandelbaum, R., Rowe, B., Armstrong, R., Bard, D., Bertin, E., Bosch, J., Boutigny, D., Courbin, F., Dawson, W. A., Donnarumma, A., et al. (2015). GREAT3 results–I. Systematic errors in shear estimation and the impact of real galaxy morphology. *Monthly Notices of the Royal Astronomical Society*, 450(3):2963–3007.

Mandelbaum, R., Rowe, B., Bosch, J., Chang, C., Courbin, F., Gill, M., Jarvis, M., Kannawadi, A., Kacprzak, T., Lackner, C., et al. (2014). The third gravitational lensing accuracy testing (GREAT3) challenge handbook. *The Astrophysical Journal Supplement Series*, 212(1):5.

Maoli, R., Van Waerbeke, L., Mellier, Y., Schneider, P., Jain, B., Bernardeau, F., Erben, T., and Fort, B. (2001). Cosmic shear analysis in 50 uncorrelated VLT fields. implications for Ω_0, σ_8. *Astronomy & Astrophysics*, 368(3):766–775.

Martinet, N., Schneider, P., Hildebrandt, H., Shan, H., Asgari, M., Dietrich, J. P., Harnois-Déraps, J., Erben, T., Grado, A., Heymans, C., et al. (2017). KiDS-450: cosmological constraints from weak-lensing peak statistics–II: Inference from shear peaks using N-body simulations. *Monthly Notices of the Royal Astronomical Society*, 474(1):712–730.

Martinet, N., Schrabback, T., Hoekstra, H., Tewes, M., Herbonnet, R., Schneider, P., Hernandez-Martin, B., Taylor, A., Brinchmann, J., et al. (2019). Euclid preparation: IV. impact of undetected galaxies on weak-lensing shear measurements. *Astronomy & Astrophysics*.

Massey, R., Heymans, C., Bergé, J., Bernstein, G., Bridle, S., Clowe, D., Dahle, H., Ellis, R., Erben, T., Hetterscheidt, M., et al. (2007). The Shear Testing Programme 2: Factors affecting high-precision weak-lensing analyses. *Monthly Notices of the Royal Astronomical Society*, 376(1):13–38.

Massey, R., Hoekstra, H., Kitching, T., Rhodes, J., Cropper, M., Amiaux, J., Harvey, D., Mellier, Y., Meneghetti, M., Miller, L., et al. (2012). Origins of weak lensing systematics, and requirements on future instrumentation (or knowledge of instrumentation). *Monthly Notices of the Royal Astronomical Society*, 429(1):661–678.

Massey, R. and Refregier, A. (2005). Polar shapelets. *Monthly Notices of the Royal Astronomical Society*, 363(1):197–210.

Massey, R., Schrabback, T., Cordes, O., Marggraf, O., Israel, H., Miller, L., Hall, D., Cropper, M., Prod'homme, T., and Matias Niemi, S. (2014). An improved model of charge transfer inefficiency and correction algorithm for the Hubble Space Telescope. *Monthly Notices of the Royal Astronomical Society*, 439(1):887–907.

McCann, R. J. (1997). A convexity principle for interacting gases. *Advances in mathematics*, 128(1):153–179.

Melchior, P., Moolekamp, F., Jerdee, M., Armstrong, R., Sun, A.-L., Bosch, J., and Lupton, R. (2018). Scarlet: Source separation in multi-band images by constrained matrix factorization. *Astronomy and Computing*, 24:129–142.

Melchior, P., Viola, M., Schäfer, B. M., and Bartelmann, M. (2011). Weak gravitational lensing with DEIMOS. *Monthly Notices of the Royal Astronomical Society*, 412(3):1552–1558.

Mérigot, Q. (2011). A multiscale approach to optimal transport. In *Computer Graphics Forum*, volume 30, pages 1583–1592. Wiley Online Library.

Merten, J., Giocoli, C., Baldi, M., Meneghetti, M., Peel, A., Lalande, F., Starck, J.-L., and Pettorino, V. (2019). On the dissection of degenerate cosmologies with machine learning. *Monthly Notices of the Royal Astronomical Society*, 487(1):104–122.

Meyers, J. E. and Burchat, P. R. (2015). Impact of atmospheric chromatic effects on weak lensing measurements. *The Astrophysical Journal*, 807(2):182.

Miller, L., Heymans, C., Kitching, T., Van Waerbeke, L., Erben, T., Hildebrandt, H., Hoekstra, H., Mellier, Y., Rowe, B., Coupon, J., et al. (2013). Bayesian galaxy shape measurement for weak lensing surveys–III. Application to the Canada–France–Hawaii Telescope Lensing Survey. *Monthly Notices of the Royal Astronomical Society*, 429(4):2858–2880.

Miller, L., Kitching, T., Heymans, C., Heavens, A., and Van Waerbeke, L. (2007). Bayesian galaxy shape measurement for weak lensing surveys–I. Methodology and a fast-fitting algorithm. *Monthly Notices of the Royal Astronomical Society*, 382(1):315–324.

Moffat, A. (1969). A theoretical investigation of focal stellar images in the photographic emulsion and application to photographic photometry. *Astronomy and Astrophysics*, 3:455.

Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie Royale des Sciences de Paris*.

Montavon, G., Müller, K.-R., and Cuturi, M. (2016). Wasserstein training of restricted Boltzmann machines. In *Advances in Neural Information Processing Systems*, pages 3711–3719.

Morales, J. L. and Nocedal, J. (2011). Remark on "algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization". *ACM Transactions on Mathematical Software (TOMS)*, 38(1):7.

Moreau, J. J. (1962). Fonctions convexes duales et points proximaux dans un espace Hilbertien. In Elsevier, editor, *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, volume 255, pages 2897–2899.

Morgan, W. W. and Keenan, P. (1973). Spectral classification. *Annual Review of Astronomy and Astrophysics*, 11(1):29–50.

Morrison, C. B., Hildebrandt, H., Schmidt, S. J., Baldry, I. K., Bilicki, M., Choi, A., Erben, T., and Schneider, P. (2017). THE-WIZZ: clustering redshift estimation for everyone. *Monthly Notices of the Royal Astronomical Society*, 467(3):3576–3589.

Nesterov, Y. (2018). *Lectures on convex optimization*, volume 137. Springer.

Ngolè, F. and Starck, J.-L. (2017). Point spread function field learning based on optimal transport distances. *SIAM Journal on Imaging Sciences*, 10(3):1549–1578.

Ngolè, F., Starck, J.-L., Okumura, K., Amiaux, J., and Hudelot, P. (2016). Constraint matrix factorization for space variant PSFs field restoration. *Inverse Problems*, 32(12):124001.

Ngolè, F., Starck, J.-L., Ronayette, S., Okumura, K., and Amiaux, J. (2015). Super-resolution method using sparse regularization for point-spread function recovery. *Astronomy & Astrophysics*, 575:A86.

Nightingale, J., Massey, R., et al. (in prep.). Euclid: Charge-transfer inefficiency modelling and correction (official title TBA).

Okabe, N. and Umetsu, K. (2008). Subaru weak lensing study of seven merging clusters: distributions of mass and baryons. *Publications of the Astronomical Society of Japan*, 60(2):345–375.

Papadakis, N. (2015). *Optimal Transport for Image Processing*. Habilitation à diriger des recherches, Université de Bordeaux.

Parikh, N. and Boyd, S. (2014). Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):127–239.

Paulin-Henriksson, S., Amara, A., Voigt, L., Refregier, A., and Bridle, S. (2008). Point spread function calibration requirements for dark energy from cosmic shear. *Astronomy & Astrophysics*, 484(1):67–77.

Paulin-Henriksson, S., Refregier, A., and Amara, A. (2009). Optimal point spread function modeling for weak lensing: complexity and sparsity. *Astronomy & Astrophysics*, 500(2):647–655.

Paykari, P., Kitching, T. D., Hoekstra, H., Azzollini, R., Cardone, V., Cropper, M., Duncan, C., Kannawadi, A., Miller, L., Wallis, C., Aussel, H., Van Uitert, E., Fenech Conti, I., et al. (2019). Euclid preparation: VI. Verifying the performance of cosmic shear experiments. *under review by ECEB*.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572.

Peebles, P. J. E. (1980). *The large-scale structure of the Universe*. Princeton university press.

Peel, A., Lalande, F., Starck, J.-L., Pettorino, V., Merten, J., Giocoli, C., Meneghetti, M., and Baldi, M. (2019). Distinguishing standard and modified gravity cosmologies with machine learning. *Physical Review D*, 100(2):023508.

Peel, A., Lanusse, F., and Starck, J.-L. (2017). Sparse reconstruction of the merging A520 cluster system. *The Astrophysical Journal*, 847(1):23.

Pennington, J., Socher, R., and Manning, C. D. (2014). Glove: Global Vectors for Word Representation. In *EMNLP*, volume 14, pages 1532–1543.

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R., Nugent, P., Castro, P., Deustua, S., Fabbro, S., Goobar, A., Groom, D., et al. (1999). Measurements of Ω and Λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2):565.

Perryman, M., de Boer, K. S., Gilmore, G., Høg, E., Lattanzi, M., Lindegren, L., Luri, X., Mignard, F., Pace, O., and de Zeeuw, P. (2001). Gaia: Composition, formation and evolution of the galaxy. *Astronomy & Astrophysics*, 369(1):339–363.

Peyré, G., Chizat, L., Vialard, F.-X., and Solomon, J. (2016). Quantum optimal transport for tensor field processing. *arXiv preprint arXiv:1612.08731*.

Peyré, G. and Cuturi, M. (2019). Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5-6):355–607.

Pickles, A. (1998). A stellar spectral flux library: 1150–25000Å. *Publications of the Astronomical Society of the Pacific*, 110(749):863.

Piotrowski, L., Batsch, T., Czyrkowski, H., Cwiok, M., Dabrowski, R., Kasprowicz, G., Majcher, A., Majczyna, A., Malek, K., Mankiewicz, L., et al. (2013). PSF modelling for very wide-field CCD astronomy. *Astronomy & Astrophysics*, 551:A119.

Pitié, F., Kokaram, A. C., and Dahyot, R. (2005). N-dimensional probablility density function transfer and its application to colour transfer. In *Proceedings of the Tenth IEEE International Conference on Computer Vision - Volume 2*, ICCV '05, pages 1434–1439, Washington, DC, USA. IEEE Computer Society.

Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A., Barreiro, R., Bartolo, N., Basak, S., et al. (2018). Planck 2018 results. VI. Cosmological parameters. *arXiv preprint arXiv:1807.06209*.

Polyak, B. T. (1964). Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17.

Price, M. A., McEwen, J. D., Cai, X., Kitching, T. D., and Wallis, C. G. (2018). Sparse bayesian mass-mapping with uncertainties: hypothesis testing of structure. *arXiv preprint arXiv:1812.04014*.

Pujol, A., Kilbinger, M., Sureau, F., and Bobin, J. (2019). A highly precise shear bias estimator independent of the measured shape noise. *Astronomy & Astrophysics*, 621:A2.

Pujol, A., Sureau, F., Bobin, J., Courbin, F., Gentile, M., and Kilbinger, M. (2017). Shear measurement bias: dependencies on methods, simulation parameters and measured parameters. *arXiv preprint arXiv:1707.01285*.

Quan, Y., Bao, C., and Ji, H. (2016). Equiangular kernel dictionary learning with applications to dynamic texture analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 308–316.

Rabin, J., Peyré, G., Delon, J., and Bernot, M. (2011). Wasserstein barycenter and its application to texture mixing. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 435–446. Springer.

Rachev, S. and Rüschendorf, L. (1998). *Mass Transportation Problems: Theory*, volume 1. Springer Verlag.

Raguet, H., Fadili, J., and Peyré, G. (2013). A generalized forward-backward splitting. *SIAM Journal on Imaging Sciences*, 6(3):1199–1226.

Ravanbakhsh, S., Lanusse, F., Mandelbaum, R., Schneider, J., and Poczos, B. (2017). Enabling dark energy science with deep generative models of galaxy images. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Refregier, A. (2003). Shapelets—I. A method for image analysis. *Monthly Notices of the Royal Astronomical Society*, 338(1):35–47.

Refregier, A. and Bacon, D. (2003). Shapelets—II. A method for weak lensing measurements. *Monthly Notices of the Royal Astronomical Society*, 338(1):48–56.

Rhodes, J., Nichol, R. C., Aubourg, É., Bean, R., Boutigny, D., Bremer, M. N., Capak, P., Cardone, V., Carry, B., Conselice, C. J., et al. (2017). Scientific synergy between LSST and Euclid. *Astrophysical journal supplement series.*, 233(2):21.

Rhodes, J., Refregier, A., Collins, N. R., Gardner, J. P., Groth, E. J., and Hill, R. S. (2004). Measurement of cosmic shear with the Space Telescope Imaging Spectrograph. *The Astrophysical Journal*, 605(1):29.

Rhodes, J., Refregier, A., and Groth, E. J. (2000). Weak lensing measurements: A revisited method and application to Hubble Space Telescope images. *The Astrophysical Journal*, 536(1):79.

Rhodes, J. D., Massey, R. J., Albert, J., Collins, N., Ellis, R. S., Heymans, C., Gardner, J. P., Kneib, J.-P., Koekemoer, A., Leauthaud, A., et al. (2007). The stability of the point-spread function of the Advanced Camera for Surveys on the Hubble Space Telescope and implications for weak gravitational lensing. *The Astrophysical Journal Supplement Series*, 172(1):203.

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., et al. (1998). Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009.

Rix, H.-W., Barden, M., Beckwith, S. V., Bell, E. F., Borch, A., Caldwell, J. A., Häussler, B., Jahnke, K., Jogee, S., McIntosh, D. H., et al. (2004). GEMS: Galaxy evolution from morphologies and SEDs. *The Astrophysical Journal Supplement Series*, 152(2):163.

Rolet, A., Cuturi, M., and Peyré, G. (2016). Fast dictionary learning with a smoothed Wasserstein loss. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pages 630–638.

Romano, A., Fu, L., Giordano, F., Maoli, R., Martini, P., Radovich, M., Scaramella, R., Antonuccio-Delogu, V., Donnarumma, A., Ettori, S., et al. (2010). Abell 611-I. Weak lensing analysis with LBC. *Astronomy & Astrophysics*, 514:A88.

Rowe, B. (2010). Improving PSF modelling for weak gravitational lensing using new methods in model selection. *Monthly Notices of the Royal Astronomical Society*, 404(1):350–366.

Rowe, B., Hirata, C., and Rhodes, J. (2011). Optimal linear image combination. *The Astrophysical Journal*, 741(1):46.

Rowe, B., Jarvis, M., Mandelbaum, R., Bernstein, G. M., Bosch, J., Simet, M., Meyers, J. E., Kacprzak, T., Nakajima, R., Zuntz, J., et al. (2015). GalSim: The modular galaxy image simulation toolkit. *Astronomy and Computing*, 10:121–150.

Rubinstein, R., Zibulevsky, M., and Elad, M. (2008). Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit. Technical report, Computer Science Department, Technion.

Rubner, Y., Tomasi, C., and Guibas, L. J. (2000). The Earth Mover's Distance as a metric for image retrieval. *Int. J. Comput. Vision*, 40(2):99–121.

Salton, G. and McGill, M. J. (1983). *Introduction to modern information retrieval*. McGraw-Hill, Inc.

Salvato, M., Ilbert, O., and Hoyle, B. (2019). The many flavours of photometric redshifts. *Nature Astronomy*, 3(3):212.

Sandler, R. and Lindenbaum, M. (2009). Non-negative matrix factorization with Earth Mover's Distance metric. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1873–1880. IEEE.

Schirmer, M., Erben, T., Schneider, P., Pietrzynski, G., Gieren, W., Carpano, S., Micol, A., and Pierfederici, F. (2003). GaBoDS: The Garching-Bonn Deep Survey-I. Anatomy of galaxy clusters in the background of NGC 300. *Astronomy & Astrophysics*, 407(3):869–888.

Schmitz, M. A., Heitz, M., Bonneel, N., Ngolè, F., Coeurjolly, D., Cuturi, M., Peyré, G., and Starck, J.-L. (2017). Optimal transport-based dictionary learning and its application to Euclid-like point spread function representation. In *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics.

Schmitz, M. A., Heitz, M., Bonneel, N., Ngole, F., Coeurjolly, D., Cuturi, M., Peyré, G., and Starck, J.-L. (2018). Wasserstein dictionary learning: Optimal transport-based unsupervised nonlinear dictionary learning. *SIAM Journal on Imaging Sciences*, 11(1):643–678.

Schmitz, M. A., Starck, J.-L., Mboula, F. N., Auricchio, N., Brinchmann, J., Capobianco, R., Clédassou, R., Conversi, L., Corcione, L., Fourmanoit, N., et al. (2019). Euclid: Non-parametric point spread function field recovery through interpolation on a graph Laplacian. *arXiv preprint arXiv:1906.07676v1*.

Schmitzer, B. (2019). Stabilized sparse scaling algorithms for entropy regularized transport problems. *SIAM Journal on Scientific Computing*, 41(3):A1443–A1481.

Schneider, P., Eifler, T., and Krause, E. (2010). COSEBIs: Extracting the full E-/B-mode information from cosmic shear correlation functions. *Astronomy & Astrophysics*, 520:A116.

Schneider, P., Van Waerbeke, L., Jain, B., and Kruse, G. (1998). A new measure for cosmic shear. *Monthly Notices of the Royal Astronomical Society*, 296(4):873–892.

Schneider, P., van Waerbeke, L., Kilbinger, M., and Mellier, Y. (2002a). Analysis of two-point statistics of cosmic shear-I. Estimators and covariances. *Astronomy & Astrophysics*, 396(1):1–19.

Schneider, P., Van Waerbeke, L., and Mellier, Y. (2002b). B-modes in cosmic shear from source redshift clustering. *Astronomy & Astrophysics*, 389(3):729–741.

Schölkopf, B., Smola, A., and Müller, K.-R. (1997). Kernel principal component analysis. *Artificial Neural Networks — ICANN'97*, pages 583–588.

Schrabback, T., Erben, T., Simon, P., Miralles, J.-M., Schneider, P., Heymans, C., Eifler, T., Fosbury, R., Freudling, W., Hetterscheidt, M., et al. (2007). Cosmic shear analysis of archival HST/ACS data-I. Comparison of early ACS pure parallel data to the HST/GEMS survey. *Astronomy & Astrophysics*, 468(3):823–847.

Schrabback, T., Hartlap, J., Joachimi, B., Kilbinger, M., Simon, P., Benabed, K., Bradač, M., Eifler, T., Erben, T., Fassnacht, C. D., et al. (2010). Evidence of the accelerated expansion of the universe from weak lensing tomography with COSMOS. *Astronomy & Astrophysics*, 516:A63.

Schrödinger, E. (1931). *Über die umkehrung der naturgesetze*. Verlag Akademie der wissenschaften in kommission bei Walter de Gruyter u. Company.

Schuhmann, R. L., Heymans, C., and Zuntz, J. (2019). Galaxy shape measurement synergies between LSST and Euclid. *arXiv preprint arXiv:1901.08586*.

Scoville, N., Aussel, H., Brusa, M., Capak, P., Carollo, C. M., Elvis, M., Giavalisco, M., Guzzo, L., Hasinger, G., Impey, C., et al. (2007). The cosmic evolution survey (COSMOS): overview. *The Astrophysical Journal Supplement Series*, 172(1):1.

Seguy, V. and Cuturi, M. (2015). Principal geodesic analysis for probability measures under the optimal transport metric. In *Advances in Neural Information Processing Systems*, pages 3312–3320.

Seitz, C. and Schneider, P. (1997). Steps towards nonlinear cluster inversion through gravitational distortions. III. Including a redshift distribution of the sources. *Astronomy and Astrophysics*, 318:687–699.

Seitz, S., Schneider, P., and Ehlers, J. (1994). Light propagation in arbitrary spacetimes and the gravitational lens approximation. *Classical and Quantum Gravity*, 11(9):2345.

Semboloni, E., Hoekstra, H., Huang, Z., Cardone, V., Cropper, M., Joachimi, B., Kitching, T., Kuijken, K., Lombardi, M., Maoli, R., et al. (2013). On the shear estimation bias induced by the spatial variation of colour across galaxy profiles. *Monthly Notices of the Royal Astronomical Society*, 432(3):2385–2401.

Sérsic, J. (1963). Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy. *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, 6:41.

Shapiro, C. and Cooray, A. (2006). The Born and lens–lens corrections to weak gravitational lensing angular power spectra. *Journal of Cosmology and Astroparticle Physics*, 2006(03):007.

Shechtman, Y., Eldar, Y. C., Cohen, O., Chapman, H. N., Miao, J., and Segev, M. (2015). Phase retrieval with application to optical imaging: a contemporary overview. *IEEE signal processing magazine*, 32(3):87–109.

Sheldon, E. (2015). NGMIX: Gaussian mixture models for 2D images. *Astrophysics Source Code Library*.

Sheldon, E. S. and Huff, E. M. (2017). Practical weak-lensing shear measurement with metacalibration. *The Astrophysical Journal*, 841(1):24.

Shirdhonkar, S. and Jacobs, D. W. (2008). Approximate Earth Mover's Distance in linear time. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE.

Simon, P. and Schneider, P. (2017). Weak-lensing shear estimates with general adaptive moments, and studies of bias by pixellation, PSF distortions, and noise. *Astronomy & Astrophysics*, 604:A109.

Sinkhorn, R. (1967). Diagonal equivalence to matrices with prescribed row and column sums. *The American Mathematical Monthly*, 74(4):402–405.

Smith, R. E., Peacock, J. A., Jenkins, A., White, S., Frenk, C., Pearce, F., Thomas, P. A., Efstathiou, G., and Couchman, H. (2003). Stable clustering, the halo model and non-linear cosmological power spectra. *Monthly Notices of the Royal Astronomical Society*, 341(4):1311–1332.

Solomon, J., De Goes, F., Peyré, G., Cuturi, M., Butscher, A., Nguyen, A., Du, T., and Guibas, L. (2015). Convolutional Wasserstein distances: Efficient optimal transportation on geometric domains. *ACM Transactions on Graphics (TOG)*, 34(4):66.

Solomon, J., Rustamov, R., Guibas, L., and Butscher, A. (2014). Wasserstein propagation for semi-supervised learning. In *Proceedings of The 31st International Conference on Machine Learning*, pages 306–314.

Springer, O. M., Ofek, E. O., Weiss, Y., and Merten, J. (2018). Weak lensing shear estimation beyond the shape-noise limit: a machine learning approach. *arXiv preprint arXiv:1808.07491*.

Spurio Mancini, A., Reischke, R., Pettorino, V., Schäfer, B., and Zumalacárregui, M. (2018). Testing (modified) gravity with 3D and tomographic cosmic shear. *Monthly Notices of the Royal Astronomical Society*, 480(3):3725–3738.

Starck, J.-L. and Murtagh, F. (2007). *Astronomical image and data analysis*. Springer Science & Business Media.

Starck, J.-L., Murtagh, F., and Bertero, M. (2011). Starlet transform in astronomical data processing. In *Handbook of Mathematical Methods in Imaging*, pages 1489–1531. Springer.

Starck, J.-L., Murtagh, F., and Fadili, J. (2015). *Sparse image and signal processing: Wavelets and related geometric multiscale analysis*. Cambridge University Press.

Suksmono, A. B. (2013). Reconstruction of complex-valued fractional Brownian motion fields based on compressive sampling and its application to PSF interpolation in weak lensing survey. *arXiv preprint arXiv:1311.0124*.

Talagrand, M. (1996). Transportation cost for Gaussian and other product measures. *Geometric and Functional Analysis*, 6(3):587–600.

Tanaka, M., Coupon, J., Hsieh, B.-C., Mineo, S., Nishizawa, A. J., Speagle, J., Furusawa, H., Miyazaki, S., and Murayama, H. (2017). Photometric redshifts for Hyper Suprime-Cam Subaru Strategic Program Data Release 1. *Publications of the Astronomical Society of Japan*, 70(SP1):S9.

Tewes, M., Kuntzer, T., Nakajima, R., Courbin, F., Hildebrandt, H., and Schrabback, T. (2019). Weak-lensing shear measurement with machine learning—Teaching artificial neural networks about feature noise. *Astronomy & Astrophysics*, 621:A36.

Theano Development Team (2016). Theano: A Python framework for fast computation of mathematical expressions. *arXiv e-prints*, abs/1605.02688.

Treu, T. (2010). Strong lensing by galaxies. *Annual Review of Astronomy and Astrophysics*, 48:87–125.

Troxel, M. and Ishak, M. (2015). The intrinsic alignment of galaxies and its impact on weak gravitational lensing in an era of precision cosmology. *Physics Reports*, 558:1–59.

Turk, M. and Pentland, A. (1991). Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86.

Tyson, J., Wittman, D., Dell'Antonio, I., Becker, A., Margoniner, V., Team, D., et al. (2001). The Deep Lens Survey: Overview. In *Bulletin of the American Astronomical Society*, volume 33, page 1464.

Vaccari, M., Covone, G., Radovich, M., Grado, A., Limatola, L., Botticella, M., Cappellaro, E., Paolillo, M., Pignata, G., De Cicco, D., et al. (2017). The VOICE survey: VST Optical imaging of the CDFS and ES1 fields. *arXiv preprint arXiv:1704.01495*.

Van Nguyen, H., Patel, V. M., Nasrabadi, N. M., and Chellappa, R. (2013). Design of non-linear kernel dictionaries for object recognition. *IEEE Transactions on Image Processing*, 22(12):5123–5135.

Van Waerbeke, L., Mellier, Y., Erben, T., Cuillandre, J., Bernardeau, F., Maoli, R., Bertin, E., Mc Cracken, H., Fevre, O. L., Fort, B., et al. (2000). Detection of correlated galaxy ellipticities on CFHT data: first evidence for gravitational lensing by large-scale structures. *Arxiv preprint astro-ph/0002500*.

Van Waerbeke, L., Mellier, Y., and Hoekstra, H. (2005). Dealing with systematics in cosmic shear studies: New results from the VIRMOS-DESCART survey. *Astronomy & Astrophysics*, 429(1):75–84.

Van Waerbeke, L., Mellier, Y., Radovich, M., Bertin, E., Dantel-Fort, M., McCracken, H., Le Fevre, O., Foucaud, S., Cuillandre, J.-C., Erben, T., et al. (2001). Cosmic shear statistics and cosmology. *Astronomy & Astrophysics*, 374(3):757–769.

Villani, C. (2003). *Topics in optimal transportation*. Number 58. American Mathematical Soc.

Villani, C. (2008). *Optimal transport: old and new*, volume 338. Springer Science & Business Media.

Viola, M., Melchior, P., and Bartelmann, M. (2011). Biases in, and corrections to, KSB shear measurements. *Monthly Notices of the Royal Astronomical Society*, 410(4):2156–2166.

Voigt, L. and Bridle, S. (2010). Limitations of model-fitting methods for lensing shear estimation. *Monthly Notices of the Royal Astronomical Society*, 404(1):458–467.

Wang, W., Slepcev, D., Basu, S., Ozolek, J. A., and Rohde, G. K. (2013). A linear optimal transportation framework for quantifying and visualizing variations in sets of images. *International Journal of Computer Vision*, 101(2):254–269.

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., and Bernstein, G. (2000). Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales. *Nature*, 405(6783):143.

Wolf, C., Dye, S., Kleinheinrich, M., Meisenheimer, K., Rix, H.-W., and Wisotzki, L. (2001). Deep BVR photometry of the Chandra Deep Field South from the COMBO-17 survey. *Astronomy & Astrophysics*, 377(2):442–449.

Wolz, L., Kilbinger, M., Weller, J., and Giannantonio, T. (2012). On the validity of cosmological Fisher matrix forecasts. *Journal of Cosmology and Astroparticle Physics*, 2012(09):009.

Xin, B., Ivezić, Ž., Lupton, R. H., Peterson, J. R., Yoachim, P., Jones, R. L., Claver, C. F., and Angeli, G. (2018). A study of the point-spread function in SDSS images. *The Astronomical Journal*, 156(5):222.

Xu, H., Wang, W., Liu, W., and Carin, L. (2018). Distilled Wasserstein learning for word embedding and topic modeling. In *Advances in Neural Information Processing Systems*, pages 1716–1725.

Ye, J., Wu, P., Wang, J. Z., and Li, J. (2017). Fast discrete distribution clustering using Wasserstein barycenter with sparse support. *IEEE Transactions on Signal Processing*, 65(9):2317–2332.

York, D. G., Adelman, J., Anderson Jr, J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J., Barkhouser, R., Bastian, S., Berman, E., et al. (2000). The Sloan Digital Sky Survey: Technical summary. *The Astronomical Journal*, 120(3):1579.

Zavriev, S. and Kostyuk, F. (1993). Heavy-ball method in nonconvex optimization problems. *Computational Mathematics and Modeling*, 4(4):336–341.

Zel'dovich, Y. B. (1972). A hypothesis, unifying the structure and the entropy of the universe. *Monthly Notices of the Royal Astronomical Society*, 160(1):1P–3P.

Zernike, F. (1934). Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode. *Physica*, 1(7-12):689–704.

Zhang, J., Dong, F., Li, H., Li, X., Li, Y., Liu, D., Luo, W., Fu, L., Li, G., and Fan, Z. (2019). Testing shear recovery with field distortion. *The Astrophysical Journal*, 875(1):48.

Zhang, J., Luo, W., and Foucaud, S. (2015). Accurate shear measurement with faint sources. *Journal of Cosmology and Astroparticle Physics*, 2015(01):024.

Zuntz, J., Kacprzak, T., Voigt, L., Hirsch, M., Rowe, B., and Bridle, S. (2013). im3shape: a maximum likelihood galaxy shear measurement code for cosmic gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, 434(2):1604–1618.

Zuntz, J., Sheldon, E., Samuroff, S., Troxel, M. A., Jarvis, M., MacCrann, N., Gruen, D., Prat, J., Sánchez, C., Choi, A., et al. (2018). Dark Energy Survey Year 1 results: weak lensing shape catalogues. *Monthly Notices of the Royal Astronomical Society*, 481(1):1149–1182.

ÉCOLE DOCTORALE
Astronomie et Astrophysique d'île-de-France

Titre: Estimation du champ de PSF pour l'effet de lentille gravitationnelle faible avec *Euclid*

Mots clés: Lentille gravitationnelle faible, Cosmologie, Euclid, Fonction d'étalement du point

Résumé: Le chemin parcouru par la lumière, lors de sa propagation dans l'Univers, est altéré par la présence d'objets massifs. Cela entraine une déformation des images de galaxies lointaines. La mesure de cet effet, dit de lentille gravitationnelle faible, nous permet de sonder la structure, aux grandes échelles, de notre Univers. En particulier, nous pouvons ainsi étudier la distribution de la matière noire et les propriétés de l'Energie Sombre, proposée comme origine de l'accélération de l'expansion de l'Univers. L'étude de l'effet de lentille gravitationnelle faible constitue l'un des objectifs scientifiques principaux d'*Euclid*, un télescope spatial de l'Agence Spatiale Européenne dont le lancement est prévu en 2022.
En pratique, ce signal est obtenu en mesurant la forme des galaxies. Toute image produite par un instrument optique est altérée par sa fonction d'étalement du point (PSF). Celle-ci a diverses origines : diffraction, imperfections dans les composantes optiques de l'instrument, effets atmosphériques (pour les télescopes au sol)... Puisque la PSF affecte aussi les formes des galaxies, il est crucial de la prendre en compte lorsque l'on étudie l'effet de lentille gravitationnelle faible, ce qui nécessite de connaître la PSF avec une très grande précision. Celle-ci varie en fonction de la position dans le plan focal. Une mesure de la PSF, à certaines positions, est donnée par l'observation d'étoiles non-résolues dans le champ, à partir desquelles on peut construire un modèle de PSF. Dans le cas d'*Euclid*, ces images d'étoiles seront sous-échantillonnée et il est nécessaire d'intégrer une étape de super-résolution dans la construction du modèle. En raison de la très large bande d'intégration de l'imageur visible d'*Euclid*, il sera également nécessaire de capturer les variations en longueur d'onde de la PSF.
La contribution principale de cette thèse est le développement de méthodes novatrices d'estimation de la PSF, reposant sur plusieurs outils : la notion de représentation parcimonieuse, et le transport optimal numérique. Ce dernier nous permet de proposer la première méthode capable de fournir un modèle polychromatique de la PSF, construit uniquement à partir d'images sous-échantillonnées d'étoiles et leur spectre. Une étude de la propagation des erreurs de PSF sur la mesure de forme de galaxies est également proposée.

Title: *Euclid* weak lensing: PSF field estimation

Keywords: Weak Lensing, Cosmology, Euclid, Point Spread Function

Abstract: As light propagates through the Universe, its path is altered by the presence of massive objects. This causes a distortion of the images of distant galaxies. Measuring this effect, called weak gravitational lensing, allows us to probe the large scale structure of the Universe. This makes it a powerful source of cosmological insight, and can in particular be used to study the distribution of dark matter and the nature of Dark Energy. The European Space Agency's upcoming *Euclid* mission is a spaceborne telescope with weak lensing as one of its primary science objectives. It is expected to launch in 2022.
In practice, the weak lensing signal is recovered from the measurement of the shapes of galaxies. The images obtained by any optical instrument are altered by its Point Spread Function (PSF), caused by various effects: diffraction, imperfect optics, atmospheric turbulence (for ground-based telescopes)... Since the PSF also alters galaxy shapes, it is crucial to correct for it when performing weak lensing measurements. This, in turn, requires precise knowledge of the PSF itself.
The PSF varies depending on the position of objects within the instrument's focal plane. Unresolved stars in the field provide a measurement of the PSF at given positions, from which a PSF model can be built. In the case of *Euclid*, star images will suffer from undersampling. The PSF model will thus need to perform a super-resolution step. In addition, because of the very wide band of its visible instrument, variations of the PSF with the wavelength of incoming light will also need to be accounted for. The main contribution of this thesis is the building of novel PSF modelling approaches. These rely on sparsity and numerical optimal transport. The latter enables us to propose the first method capable of building a polychromatic PSF model, using no information other than undersampled star images, their position and spectra. We also study the propagation of errors in the PSF to the measurement of galaxy shapes.

Université Paris-Saclay
CEA Paris-Saclay, Département d'Astrophysique
Orme des Merisiers, F-91191 Gif-sur-Yvette, France