UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

ON THE ANALYSIS OF DECISION PROBLEMS IN ASTROMETRY AND HYPOTHESIS TESTING

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO

SEBASTIÁN ANDRÉS ESPINOSA TRUJILLO

PROFESOR GUÍA:
JORGE SILVA SÁNCHEZ

PROFESORES CO-GUÍA:
RENÉ MÉNDEZ BUSSARD
PABLO PIANTANIDA

MIEMBROS DE LA COMISIÓN:
MARCOS ORCHARD CONCHA
FRANCISCO FÖRSTER BURÓN

SANTIAGO DE CHILE
2018

RESUMEN DE TESIS PARA OPTAR AL
TÍTULO DE MAGISTER EN CIENCIAS
DE LA INGENIERÍA, MENCIÓN ELÉCTRICA
RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE INGENIERO CIVIL ELÉCTRICO
POR: SEBASTIÁN ANDRÉS ESPINOSA TRUJILLO
FECHA: MARZO 2018
PROFESOR GUÍA: JORGE SILVA SÁNCHEZ

ON THE ANALYSIS OF DECISION PROBLEMS IN ASTROMETRY AND HYPOTHESIS TESTING

La teoría de la información surgió gracias al trabajo realizado por Claude E. Shannon: "A Mathematical Theory of Communication", donde se modela y caracteriza el desempeño óptimo de los sistemas de comunicación digitales. La idea básica es la transmisión de información mediante un canal que introduce incertidumbre en la comunicación. La señal llega a un receptor que debe decodificar la información de forma confiable en el sentido de probabilidad de error. Estableciendo una conexión con problemas de inferencia estadística, vemos que están estrechamente conectados. Ambos problemas se encargan de trabajar con observaciones y la información contenida en ellos. El objetivo final es tomar una decisión correcta basada en las observaciones. El término decisión correcta implica establecer métricas de desempeño. La teoría de la información cumple un rol muy importante al establecer límites fundamentales para problemas de decisión estadísticos, es por esto que esta tesis hace uso de las herramientas en estadística y teoría de la información para resolver dos problemas de inferencia, en el contexto de la astronomía y detección con restricción de tasa.

La primera parte de la tesis, estudia los límites fundamentales en astrometría. El foco del trabajo es estudiar la alcanzabilidad de los límites fundamentales con estimadores prácticos. El trabajo propone cotas de desempeño para estimadores clásicos (máxima verosimilitud y mínimos cuadrados) con estos resultados se verifica numéricamente la optimalidad del estimador de máxima verosimilitud en el sentido que éste alcanza la cota de Cramer-Rao en un gran espectro de regimenes observacionales.

La segunda parte de la tesis propone una cota alcanzable del error asociado al problema de detección en un contexto de hipótesis bivariado cuando una de las fuentes es transmitida con restricciones en la tasa. Este problema radica en establecer velocidades de convergencia para el error de tipo II sujeto a un error de tipo I prescrito y cuando se tiene información limitada de una de las fuentes. Para ello este trabajo establece cotas para la discrepancia que existe entre el límite fundamental asintótico y una expresión no asintótica derivada como parte de éste trabajo.

Summary

Information theory was originally iniciated by Claude E. Shannon in 1948 to find fundamental limits on data compression and digital communication, in a landmark paper entitled "A Mathematical Theory of Communication". The goal in communication is to send a message over a noisy channel, and to reconstruct it with low probability of error. Then communication theory and statistics are intimately connected. In both context the main objective is to make a minimum cost decision based on observations. In this context, information theory has played a very important role in establishing fundamental performance limits for these problems. Inspired by this theory, this thesis uses tools from statistics theory for the resolution of two punctually relevant decision making problems that deals with uncertainty and information.

On the first part of this thesis, we derive fundamental limits for the astrometry problem, where we characterize the best attainable precision limit for a family of unbiased estimators. We obtain performance bound for classical estimators (maximum likelihood and least squares) and then we validate numerically the optimality of the celebrated maximum likelihood estimator in the sense that it achieves the Cramer-Rao lower bound in a wide range of observational settings and observational regimes.

On the second part of this thesis, we derive an achievable bound of the error associated to the detection problem in a bivariate hypothesis testing problem, where one of the sources is transmitted with a rate constraint. The contribution of this work is the derivation of closed form expression for the type II error given a prescribed type I error and a limited rate. In particular, we clarified the discrepancy between practical finite length regime scenario and the fundamental asymptotic limit.

Para mi familia

Acknowledgements

A mi parecer los agradecimientos son una instancia para destacar a aquellas personas que resultaron ser fundamentales para la elaboración de este trabajo y también para el desarrollo profesional y personal de uno. A lo largo del proceso he conocido gente, varios con un aporte significativo y otros en menor cantidad pero importante de todas formas. Si bien la cantidad de personas que forman parte de mi núcleo no es la más extensa, sí resulta la más importante para mí. Dada mi personalidad, que va en relación con la estructura de la tesis, quise dividir los agradecimientos en distintos ámbitos desde lo más personal a lo más general.

En primer lugar quisiera agradecer a mi familia, por supuesto que ellos son uno de mis pilares que me inspiraron a seguir esta área del conocimiento y además son un apoyo constante frente a mis inquietudes. Nunca he dejado de aprender con ellos, por ejemplo la sabiduría de mi estimado padre que siempre tiene una respuesta para todo o me da consejos para superar mis adversidades. Mi querida madre que siempre está para el consuelo, cariño o la solución práctica para mis inquietudes, aunque me porte mal al final del día ella esta siempre para apoyarme y quererme. Finalmente mi hermano que escucho solamente noticias positivas y aliento de su parte, a quien le deseo lo mejor y las gracias por su apoyo constante pese a las diferencias que tengamos.

Por otra parte, tengo que agradecer a una persona muy importante para mi, Paula, quien ha sido otro pilar más para mi desarrollo. Su cariño y apoyo incondicional han sido fundamentales para la toma de mis decisiones y el hecho de que el camino lo hayamos podido construir juntos me trae una emoción y felicidad que son muy dificiles de describir. No seré el mejor con las palabras, por lo que creo que es dificil que este párrafo pueda demostrar todo mi cariño hacia ella, pero espero que se transmita la felicidad que me produce el escribir este agradecimiento.

En segundo lugar quiero agradecer a mis amistades más cercanas, el grupo de los amigos del colegio cuyo nombre no se puede escribir por posible censura. En particular tengo que agradecer a Ballero quien siempre ha estado disponible para que podamos pasarla bien pese a la poca frecuencia con que nos vemos. Basta una tarde para ponerse al día y las partidas de Rocket League (juegazo), Catán o Halo (del 1 al 5) son el mejor espacio de entretención que podría pedir. Ballero es un partner que siempre estará para compartir.

En tercer lugar agradezco a mis compañeros de Universidad, son muchos los que han pasado, me han ayudado a abrir mi mente, a desarrollar pensamiento crítico y a ser una persona más íntegra. En particular quiero mencionar a Marito y Franquito, al Bobby y los

Chicos Súper Coquetos (Felipe y Srta Andrea) quienes son aquellos con quien más tiempo he podido compartir y pese a la distancia les tengo un cariño muy grande.

En cuarto lugar están mis profesores, un apoyo constante desde lo académico, profesional y técnico, tengo que agradecer a mi profesor guía, Dr. Jorge Silva, quien ha sido como un segundo padre, con sus constantes charlas motivacionales que ni Steve Jobs podría dar, con su gran conocimiento y forma de enseñar que hace ver las cosas más complejas como procesos más sencillos de entender.

Finalmente quiero mencionar un ámbito que me ha motivado a seguir con mi carrera que ha sido la posibilidad de ser auxiliar, todo comenzó con un curso de introducción al álgebra del año 2012, mi experiencia fue increible y los estudiantes que he tenido hasta la fecha me han entregado un cariño y apoyo que de no haber sido docente jamás hubiese podido sentir. Esto último pasó a ser un motor para que pudiese seguir creciendo así que quiero agradecer a todos mis alumnos que he tenido hasta la fecha.

Contents

1 Introduction **1**
1.1 Astrometry . . . 1
1.2 Hypothesis testing . . . 3
1.3 Contribution . . . 4
1.4 Main Hypothesis . . . 5
1.5 Structure of the thesis . . . 5

2 Estimation in astrometry **6**
2.1 Preliminaries and background . . . 6
2.1.1 Astrometry as a parameter estimation problem . . . 6
2.1.2 The Cramér-Rao bound . . . 7
2.1.3 Achievability and performance of the LS estimator . . . 8
2.2 Bounding the performance of an implicit estimator . . . 9
2.3 Application to astrometry . . . 11
2.3.1 Bounding the variance of the WLS estimator . . . 12
2.3.2 Bounding the variance of the ML estimator . . . 13
2.4 Numerical analysis . . . 14
2.4.1 Experimental setting . . . 15
2.4.2 Bias analysis . . . 16
2.4.3 Performance analysis of the WLS estimator . . . 16
2.4.4 Performance analysis of the ML estimator . . . 20
2.4.5 Comments on an adaptive WLS estimator for astrometry . . . 23
2.5 Conclusions and Final Remarks . . . 26
2.6 Appendix . . . 26
2.6.1 Proof of Theorem 2.2 . . . 26
2.6.2 Proof of Theorem 2.3 . . . 28
2.6.3 Proof of Theorem 2.4 . . . 31

3 Detection in Hypothesis Testing **35**
3.1 Preliminaries . . . 35
3.2 Inference Problem . . . 36
3.3 Non asymptotic analysis for the Error Exponent: no-rate constraint case . . . 38
3.3.1 Achievability (Upper bound) and Converse (lower bound) for the discrepancy . . . 39
3.4 Discrepancy analysis: Rate constraint case . . . 39
3.5 Conclusions and Final Remarks . . . 43

3.6 Appendix 43
3.6.1 Proof of Theorem 3.9 43
3.6.2 Proof of Theorem 3.10 45
3.6.3 Berry-Esséen theorem 47
3.6.4 Bounded difference inequality 47
3.6.5 Proof of Lemma 3.4.1 48

4 Conclusion **50**
4.1 Astrometry 50
4.2 Hypothesis testing 51

Bibliography **52**

List of Tables

2.1 Indicators of the performance quality of the ML estimator relative to the Cramér-Rao bound expressed in terms of the indicator $100\times\frac{\sqrt{\sigma^2_{ML}(n)+\beta_{ML}(n)}-\sigma_{ML}(n)}{\sigma_{ML}(n)}$ from the result in Theorem 2.4. The results are presented for different positions of the object $x_c \in \{x_o^*-\sigma, x_o^*-0.8*\sigma, x_o^*-0.6*\sigma, x_o^*-0.4*\sigma, x_o^*-0.2*\sigma, x_o^*\}$ (rows) and for different $S/N \in \{12, 32, 120, 230\}$ (columns). 23

List of Figures

2.1 Relative performance of the bias (as measured by $\log\left(100 \times \frac{\epsilon_J(n)}{x_c}\right)$) stipulated by Theorem 2.2 for the WLS estimator (left side, Eq. (2.27)) and the ML estimator (right side, Eq. (2.33)). Results are reported for different values of the source flux $\tilde{F} \in \{1080, 3224, 20004, 60160\}$, all in e$^-$ (top to bottom symbols respectively), as a function of the detector pixel size. The 0% level corresponds to having achieved no bias. 16

2.2 Range of the square root of the variance performance (in miliarcsecond=mas) for the WLS method in astrometry using uniform weights (equivalent to the LS method) predicted by Theorems 2.2 and 2.3, Eq. (2.28). Results are reported for different representative values of $\tilde{F}$ and across different pixel sizes (top-left to bottom-right): $\tilde{F} \in \{1080, 3224, 20004, 60160\}$ e$^-$. 18

2.3 Worse case discrepancies in Eq. (2.38) for the WLS estimator using the weights set indexed by the positions $\Theta = \{x_o^* - \sigma, x_o^* - 0.8*\sigma, x_o^* - 0.6*\sigma, x_o^* - 0.4*\sigma, x_o^* - 0.2*\sigma, x_o^*\}$. Results are reported for two S/N scenarios, namely $\tilde{F} = 20004$ e$^-$ (Left) and $\tilde{F} = 60160$ e$^-$ (Right), and across different pixel sizes. 20

2.4 Performance discrepancies (measuring the non-optimality) of the WLS estimator using the center position as a prior for the weight selection with respect to the CR bound obtained for the true object positions $\{x_o^* - \sigma, x_o^* - 0.8 * \sigma, x_o^* - 0.6 * \sigma, x_o^* - 0.4 * \sigma, x_o^* - 0.2 * \sigma, x_o^*\}$. Results are reported for two S/N scenarios, namely $\tilde{F} = 20004$ e$^-$ (Left) and $\tilde{F} = 60160$ e$^-$ (Right), and across different pixel sizes. 21

2.5 Range of the square root of the variance performance (in miliarcsecond=mas) for the ML method in astrometry as predicted by Theorems 2.2 and 2.4, Eq. (2.34). Results are reported for different values of $\tilde{F}$ and across different pixel sizes (top-left to bottom-right): $\tilde{F} \in \{1080, 3224, 20004, 60160\}$ e$^-$. 22

2.6 Indicator of the performance optimality of the ML indicator (computed as $100 \times \frac{\sqrt{\sigma_{ML}^2(n) + \beta_{ML}(n)} - \sigma_{ML}(n)}{\sigma_{ML}(n)}$) for different positions of the target object $x_c \in \Theta = \{x_o^* - \sigma, x_o^* - 0.8 * \sigma, x_o^* - 0.6 * \sigma, x_o^* - 0.4 * \sigma, x_o^* - 0.2 * \sigma, x_o^*\}$ in the array, as a function of pixel resolution. The left panel shows the case $\tilde{F} = 20004$ e$^-$, the right panel the case $\tilde{F} = 60160$ e$^-$. 24

2.7 Performance comparison between the $\sqrt{MSE}$ of the adaptive WLS estimator and $\sigma_{CR}(n)$, both in mas. Results are reported for different $\tilde{F}$ and across different pixel sizes: (top-left to bottom-right) $\tilde{F} \in \{1080, 3224, 20004, 60160\}$ e$^-$. 25

Chapter 1

Introduction

Information theory was originally iniciated by Claude E. Shannon in 1948, in a landmark paper titled "A Mathematical Theory of Communication". The basic goal of communication is to send a message over a noisy channel, and then to reconstruct it with low probability of error, in spite of the channel noise. In this context, information theory provides a mathematical formalization of this problem and, more importantly, offers fundamental performance limits for this decision task. This view inspires the work presented in this thesis where statistical tool are applied in two relevant problems in the area of astroinformation and hypothesis testing that are summarized in Section 1.1 and 1.2

1.1 Astrometry

Astrometry deals with the accurate and precise measurement of positions and motions of celestial objects. Is the oldest branch of observational astronomy, dating back at least to Hipparchus of Nicaea in 190 BC. Since, from its very beginnings, this branch of astronomy has required measurements over time to fulfill its goals, it could be considered the precursor of the nowadays fashionable "time-domain astronomy"[1], preceding it by at least 20 centuries. In recent years, astrometry has experienced a "coming of age" motivated by the rapid increase in positional precision allowed by the use of all-digital techniques and space observatories (see, e.g., [1], Fig. 1a in [2] for an overview spanning more than 2000 years of astrometry, [3] for a summary of the contributions from the HST (fine guide sensors), and, of course, the exquisite prospects from the [4], with applications ranging from fundamental astrophysics [5, 6], to cosmology [7]).

A number of techniques have been proposed to estimate the location and flux of celestial sources as recorded on digital detectors (CCD). In this context, estimators based on the use of a least-squares (LS) error principle have been widely adopted [8–10]. The use of this type of decision rule has been traditionally justified through heuristic reasons. First, LS methods are conceptually straightforward to formulate based on the observation model of

[1] As defined, e.g., by Wikipedia: https://en.wikipedia.org/wiki/Time_domain_astronomy

these problems. Second, they offer computationally efficient implementations and have shown reasonable performance [11–13]. Finally, the LS approach was the classical method used when the observations were obtained with analog devices [14,15], which are well characterized by a Gaussian noise model for the observations. In the Gaussian case the LS is equivalent to the maximum likelihood (ML) solution ([16–18]), and, consequently, the LS method was taken from the analogous to the digital observational (Poisson noise model) setting naturally.

Considering astrometry as an inference problem (of, usually, point sources), the astrometric community has been interested for a long time in understanding the fundamental performance limits (or information bounds) of this task ([19] and references therein). It is well understood by the community that the characterization of this precision limit offers the possibility of understanding the complexity of the task and how it depends on key attributes of the problem, like the quality of the observational site, the performance of the instrument (CCD), and the details of the experimental conditions [20, 21]. On the other hand, it provides meaningful benchmarks to define the optimality of practical estimators in the process of comparing their performance with the bounds [22].

Concerning the characterization and analysis of fundamental performance bounds, we can mention some works on the use of the Cramér-Rao (CR) bound by [23–26]; and [27]. The CR bound is a minimum variance bound (MVB) for the family of unbiased estimators [28, 29]. In astrometry and joint photometry and astrometry, Mendez et al. [20,21] have recently studied the structure of this bound, and have analyzed its dependency with respect to important observational parameters under realistic (ground-based) astronomical observing conditions. In this context, closed-form expressions for the Cramér-Rao bound were derived in a number of important settings (high pixel resolution and low and high signal-to-noise (S/N) regimes), and their trends were explored across different CCD pixel resolutions and the position of the object in the CCD array. As an interesting outcome of those studies, the analysis of the CR bound has allowed us to predict the optimal pixel resolution of the array, as well as providing a formal justification to some heuristic techniques commonly used to improve performance in astrometry, like *dithering* for undersampled images [20, Sect. 3.3]. Recently, an application of the CR bound to moving sources has been done by Bouquillon et al. [30], indicating excellent agreement between the theoretical predictions, simulations, and actual ground-based observations of the Gaia satellite, in the context of the GBOT program ([31]). The use of the CR bound on other applications is also of interest, e.g., in assessing the performance of star trackers to guide satellites with demanding pointing constraints [32], or to meaningfully compare positional differences from different catalogues (for an example involving the SDSS and Gaia see [33]). Finally, a formulation of the Bayesian CR bound in astrometry, using the so-called "Van-Trees inequality" [34], has been presented by our group in Echeverría et al [35]: This approach is particularly well suited for objects at the edge of detectability, and where some prior information is available, and has been proposed for the analysis of Gaia data for faint sources, or for those with a poor observational history [36,37].

From the perspective of astrometric estimators, Lobos et al. [22] have studied in detail the performance of the widely adopted LS estimator. In particular extending the result in [38], Lobos et al. [22] derived lower and upper bounds for the mean square error (MSE) of the LS estimator. Using these bounds, the optimality of the LS estimator was analyzed, demonstrating that for high S/N there is a considerable gap between the CR bound and

the performance of the LS estimator (indicating a lack of optimality of this estimator). This work showed that for the very low S/N observational regime (weak astronomical sources), the LS estimator is near optimal, as its performance closely follows the CR bound. The limitations of the LS method in the medium to high S/N regime proved in that work opens up the question of studying alternative estimators that could achieve the CR bound on these regimes, which is the main focus of Chapter 2.

1.2 Hypothesis testing

Hypothesis testing is a classical problem of statistical decision. An hypothesis is proposed for the statistical relationship between two variables (the null hypothesis), and this is compared to an alternative hypothesis that impose no relationship between two variables. Hypothesis tests is the task of determining what outcomes of a study would lead to a rejection of the null hypothesis for a pre-specified level of significance. The process of distinguishing between the null hypothesis and the alternative hypothesis is aided by identifying two conceptual types of errors (type I and type II), and by specifying parametric limits on e.g. how much type I error will be tolerated.

Often the statistician task is to find a test with a minimal probability of an error of type II given a prescribed probability of an error of type I. It is commonly understood in statistics that the data (samples) are known to the statistician. We revisit here another dimension of this problem by assuming that the statistician does not have direct access to the data; rather, he/she has a lossy representation, more precisely, a finite rate version of it. In the problem formulated earlier, this assumption is not a significant constraint if the data are collected at a single location. In fact, the transmission of one bit enables the statistician to make an optimal decision in the sense of minimizing the probability of an error of type II for a prescribed probability of an error of type I [39]. Then, we consider the simplest problem of this kind, namely, bivariate hypothesis testing when one of the variables is measured remotely, and information about the other variable is transmitted over a noiseless channel of finite rate constraint [39].

This problem is interesting for a wide range of applications. Consider for example problems related with sensor networks and itd practical applications (self-driving cars, array of sensors for measuring, internet of things). It is clear that the ability of automated systems to make minimum risk decisions in a timely manner is crucial in the 21st century. These systems will often operate under strict constraints over their resources. In some applications, e.g. automated systems, relatively short blocklengths are common both due to delay and complexity constraints imposed in the application. It is, therefore, of critical practical interest to assess the unavoidable penalty over error exponents required to sustain the desired fidelity at a given fixed blocklength. The main goal in this thesis is to develop non asymptotic rates of convergences to the fundamental limit developed by Csiszar in [39].

Concerning the characterization and analysis of fundamental performance bounds in this task, a common approach is to determine the exponential rate of decay of the error probability of the second type with a prescribed type I error. In binary hypothesis testing, the optimal

error exponent of the Type II error is well-known and given by Stein's Lemma [40]. Extending this result, Ahlswede and Csiszar [39] characterised the asymptotic behaviour of the error exponent with communication constraint. On the other hand, Han [41] obtained fundamental bounds in the case of general systems where both sources are limited by rate constraints. Several other constributions on this topic consider extension of this kind of problem (such as asymptotic decay of the type I error or universal setting) and they can be found in [42–45].

1.3 Contribution

In the first part of this thesis, we study the ML estimator in astrometry, motivated by its well-known optimality properties in a classical parametric estimation setting with independent and identically distributed measurements (i.i.d.) [46]. We know that in the i.i.d. case this estimator is efficient with respect to the CR limit [47], but it is important to emphasize that the observational setting of astrometry deviates from the classical i.i.d. case and, consequently, the analysis of its optimality is still an open problem. In particular, we face the technical challenge of evaluating its performance, a problem that, to the best of our knowledge, has not been addressed by the astrometric community. Concerning the independent but not identically distributed case, [48] and [49] gave conditions under which ML estimators are consistent[2] and asymptotically normal[3]. Those conditions, however, are technically difficult to proof in the astrometric context.

The main challenge here is the fact that, as in the case of the LS estimator [22], the ML estimator is the solution of an optimization problem with a nonconvex cost function of the data. This implies that it is not possible to directly compute the performance of the method. To address this technical issue, we extend the approach proposed by Fessler et al. [50] to approximate the variance and the mean of an implicit estimator solution of a generic optimization problem of the data through the use of a Taylor approximation around the mean measurement (see Theorem 2.2 below). Our extension considers high order approximations of the function that allows us not only to estimate the performance of the ML estimator through an explicit nominal value, but also it provides a confidence interval around it. With this result we revisit the more general weighted least square (WLS) and ML methods providing specific upper and lower bounds for both methods (see Theorems 2.3 and 2.4). The main findings from our analysis of the bounds are two fold: first we show that the WLS exhibits a sub-optimality similar to that of the LS method for medium to high S/N regimes discovered by Lobos et al. [22] and, second, that the ML estimator achieves the CR limit for medium to high S/N and, consequently, it is optimal on those regimes. This last result is remarkable because, in conjunction with the result presented in [22], we are able to identify estimators that achieve the fundamental performance limits of astrometry in all the S/N regimes for the problem.

In the second part of this thesis, we develop an achievable bound for the error exponent

[2]this is, as the sample size increases, the sampling distribution of the estimator becomes increasingly concentrated at the true parameter value

[3]more precisely, whose distribution around the true parameter approaches a normal distribution as the sample size grows

in the bivariate hypothesis testing with communication constraints. We know that in the case with no rate restrictions, the non asymptotic performance is well characterized [42], in fact, its asymptotic limits converges to the divergence as presented in the celebrated Stein's lemma, however, little is known about the case with rate restrictions.

The main challenge faced in this work is the fact that there are many technical and mathematical difficulties when dealing with the likelihood of the ratio. Developing bounds to the induced measure given by the encoder function is not an easy task due to the freedom of such function. To address this technical issue, we extend the approach of Zhang et al. [70] to the case of noisy rate distortion theory and obtain fundamental bounds via concentration inequalitites such as the bounded difference inequality [67] and the Berry-Esséen theorem [65], which are a powerful generalization of the central limit theorem and give non asymptotic performance expression under general conditions.

1.4 Main Hypothesis

The main hypothesis of the first part of this work is that the ML estimator is efficient in a large collection of observational regimes from classical result in the i.i.d. case and empirical evidences. We conjecture that this can be proved theoretically. We also want to show that the WLS method is, in general, sub-optimal (in comparison with the minimum variance bound given by the CR result), specially at high and very high S/N.

For the problem of testing under a rate constraint, the hypothesis is that non-asymptotic expression for the error exponent can be obtained and from this being able to analize the rate of convergence of this expression to the theoretical asymptotic expression. Similar results has been established in the classical problem and we conjecture that this type of analysis can be extended from the rate constrained scenario. The main technical challenge is the fact that there exists many mathematical difficulties when dealing with the likelihood of the ratio in the rate constraint problem. Developing bounds to the induced measure given by the encoder function is not an easy task due to the freedom of such function. We want to adress this technical issue by extending the approach of Zhang et al. [70] to the case of noisy rate distortion theory and obtain fundamental bounds via concentration inequalities such as the bounded difference inequality and the Berry-Esséen theorem.

1.5 Structure of the thesis

This thesis is organized in 4 Chapters. In Chapter 2 we present the estimation problem of astrometry and characterise the fundamental limits described earlier. In Chapter 3, we present the detection problem in hypothesis testing and the main technical challenge to develop non asymptotic bounds. Finally, Chapter 4 presents the conclusions and exhibits some future work.

Chapter 2

Estimation in astrometry

2.1 Preliminaries and background

We begin by introducing the problem of astrometry. For simplicity, we focus on the 1-D scenario of a linear array detector, as it captures the key conceptual elements of the problem[1].

2.1.1 Astrometry as a parameter estimation problem

The main problem at hand is the inference of the relative position (in the array) of a point source. This source is modeled by two scalar quantities, the position of object $x_c \in \mathbb{R}$ in the array[2], and its intensity (or brightness, or flux) that we denote by $\tilde{F} \in \mathbb{R}^+$. These two parameters induce a probability distribution $\mu_{x_c,\tilde{F}}$ over an observation space that we denote by $\mathbb{X}$. Formally, given a point source represented by the pair $(x_c, \tilde{F})$, it creates a nominal intensity profile in a photon integrating device (PID), typically a CCD, which can be expressed by

$$\tilde{F}_{x_c,\tilde{F}}(x) = \tilde{F} \cdot \phi(x - x_c, \sigma), \tag{2.1}$$

where $\phi(x - x_c, \sigma)$ denotes the one dimensional normalized point spread function (PSF) and where σ is a generic parameter that determines the width (or spread) of the light distribution on the detector (typically a function of wavelength and the quality of the observing site, see Sect. 2.4) [20, 21].

The profile in Eq. (2.1) is not measured directly, but it is observed through three sources of perturbations. First, an additive background noise which captures the photon emissions of the open (diffuse) sky, and the noise of the instrument itself (the read-out noise and dark-current [51–54]), modeled by $\tilde{B}_{\mathrm{i}}$ in Eq. (2.2). Second, an intrinsic uncertainty between the aggregated intensity (the nominal object brightness plus the background) and actual

[1] The analysis can be extended to the 2-D case as presented in [20].

[2] This captures the angular position in the sky and it is measured in seconds of arc (arcsec thereafter), through the "plate-scale", which is an optical design feature of the instrument plus telescope configuration.

measurements, which is modeled by independent random variables that follow a Poisson probability law. Finally, we need to account for the spatial quantization process associated with the pixel-resolution of the PID as specified in Eqs. (2.2) and (2.3). Modeling these effects, we have a countable collection of independent and non-identically distributed random variables (observations or counts) $\{I_{\mathrm{i}} : \mathrm{i} \in \mathbb{Z}\}$, where $I_{\mathrm{i}} \sim Poisson(\lambda_{\mathrm{i}}(x_c, \tilde{F}))$, driven by the expected intensity at each pixel element i, given by

$$\lambda_{\mathrm{i}}(x_c, \tilde{F}) \equiv \mathbb{E}\{I_{\mathrm{i}}\} = \underbrace{\tilde{F} \cdot g_{\mathrm{i}}(x_c)}_{\equiv \tilde{F}_{\mathrm{i}}(x_c, \tilde{F})} + \tilde{B}_{\mathrm{i}}, \ \forall \mathrm{i} \in \mathbb{Z} \tag{2.2}$$

and

$$g_{\mathrm{i}}(x_c) \equiv \int_{x_{\mathrm{i}} - \Delta x/2}^{x_{\mathrm{i}} + \Delta x/2} \phi(x - x_c, \sigma) \, \mathrm{d}x, \ \forall \mathrm{i} \in \mathbb{Z}, \tag{2.3}$$

where $\mathbb{E}\{\}$ is the expectation value of the argument and $\{x_{\mathrm{i}} : \mathrm{i} \in \mathbb{Z}\}$ denotes the standard uniform quantization of the real line-array with resolution $\Delta x > 0$, i.e., $x_{\mathrm{i}+1} - x_{\mathrm{i}} = \Delta x$ for all $\mathrm{i} \in \mathbb{Z}$. In practice, the PID has a finite collection of measurement elements (or pixels) $I_1, .., I_n$, then a basic assumption here is that we have a good coverage of the object of interest, in the sense that for a given position x_c

$$\sum_{\mathrm{i}=1}^{n} g_{\mathrm{i}}(x_c) \approx \sum_{\mathrm{i} \in \mathbb{Z}} g_{\mathrm{i}}(x_c) = \int_{-\infty}^{\infty} \phi(x - x_c, \sigma) \, \mathrm{d}x = 1. \tag{2.4}$$

At the end, the likelihood (probability) of the joint observations $I^n = (I_1, .., I_n)$ (with values in $\mathbb{N}^n$) given the source parameters $(x_c, \tilde{F})$ is given by

$$L(I^n; x_c, \tilde{F}) = f_{\lambda_1(x_c, \tilde{F})}(I_1) \cdot f_{\lambda_2(x_c, \tilde{F})}(I_2) \cdots f_{\lambda_n(x_c, \tilde{F})}(I_n), \ \forall I^n \in \mathbb{N}^n, \tag{2.5}$$

where $f_\lambda(x) = \frac{\mathrm{e}^{-\lambda} \cdot \lambda^x}{x!}$ denotes the probability mass function (PMF) of the Poisson law [17].

Finally, if $\tilde{F}$ is assumed to be known[3], the astrometric estimation is the task of defining a decision rule $\tau_n() : \mathbb{N}^n \to \Theta$, with $\Theta = \mathbb{R}$ being the parameter space, where given an observation I^n the estimated position is given by $\hat{x}_c(I^n) = \tau_n(I^n)$.

2.1.2 The Cramér-Rao bound

In astrometry the Cramér-Rao bound has been used to bound the variance (estimation error) of any unbiased estimator [20, 21]. In general, let I^n be a collection of independent observations that follow a parametric PMF $f_{\bar{\theta}}$ defined on $\mathbb{N}$. The parameters to be estimated from I^n will be denoted in general by the vector $\bar{\theta} = (\theta_1, \theta_2, ..., \theta_m) \in \Theta = \mathbb{R}^m$. Let $\tau_n(I^n) : \mathbb{N}^n \to \Theta$ be an unbiased estimator[4] of $\bar{\theta}$, and $L(I^n; \bar{\theta}) = f_{\bar{\theta}}(I_1) \cdot f_{\bar{\theta}}(I_2) \cdots f_{\bar{\theta}}(I_n)$ be the likelihood of the observation $I^n \in \mathbb{N}^n$ given $\bar{\theta} \in \Theta$. Then, the Cramér-Rao bound [28, 29] establishes that if

$$\mathbb{E}_{I^n \sim f_{\bar{\theta}}^n} \left\{ \frac{\partial \ln L(I^n; \bar{\theta})}{\partial \theta_{\mathrm{i}}} \right\} = 0, \ \ \forall \mathrm{i} \in \{1, \ldots, m\}, \tag{2.6}$$

[3]The joint estimation of photometry and astrometry is the task of estimating both $(x_c, \tilde{F})$ from the observations, see [21].

[4]In the sense that, for all $\bar{\theta} \in \Theta$, $\mathbb{E}_{I^n \sim f_{\bar{\theta}}^n} \{\tau_n(I^n)\} = \bar{\theta}$.

then, the variance (denoted by Var), satisfies that

$$Var(\tau_n(I^n)_{\mathrm{i}}) \geq [\mathcal{I}_{\bar{\theta}}(n)^{-1}]_{\mathrm{i,i}}, \tag{2.7}$$

where $\mathcal{I}_{\bar{\theta}}(n)$ is the *Fisher information* matrix given by

$$[\mathcal{I}_{\bar{\theta}}(n)]_{\mathrm{i},j} = \mathbb{E}_{I^n \sim f^n_{\bar{\theta}}} \left\{ \frac{\partial \ln L(I^n;\bar{\theta})}{\partial \theta_{\mathrm{i}}} \cdot \frac{\partial \ln L(I^n;\bar{\theta})}{\partial \theta_j} \right\} \quad \forall \mathrm{i}, j \in \{1, \ldots, m\}. \tag{2.8}$$

In particular, for the scalar case ($m = 1$), we have that for all $\theta \in \Theta$

$$\min_{\tau_n(\cdot) \in \mathcal{T}^n} Var(\tau_n(I^n)) \geq \mathcal{I}_\theta(n)^{-1} = \mathbb{E}_{I^n \sim f^n_\theta} \left\{ \left[\left(\frac{\mathrm{d} \ln L(I^n;\theta)}{\mathrm{d}\theta} \right)^2 \right] \right\}^{-1}, \tag{2.9}$$

where $\mathcal{T}^n$ is the collection of all unbiased estimators and $I^n \sim f^n_\theta$. For astrometry, [20, 21] have characterized and analyzed the Cramér-Rao bound, leading to

Proposition 2.1 *([21, Sect. 2.4]) If $\tilde{F} \in \mathbb{R}^+$ is fixed and known, and we want to estimate x_c from $I^n \sim f_{(x_c,\tilde{F})} = L(I^n; x_c, \tilde{F})$ in Eq. (2.5), then the Fisher information is given by*

$$\mathcal{I}_{x_c}(n) = \sum_{\mathrm{i}=1}^{n} \frac{\left(\tilde{F} \frac{\mathrm{d} g_{\mathrm{i}}(x_c)}{\mathrm{d} x_c} \right)^2}{\tilde{F} g_{\mathrm{i}}(x_c) + \tilde{B}_{\mathrm{i}}}, \tag{2.10}$$

which from Eq. (2.9) induces a MVB for the astrometric estimation problem*, and where $\sigma^2_{CR}(n) \equiv \mathcal{I}_{x_c}(n)^{-1}$ denotes the (astrometric) CR bound.*

2.1.3 Achievability and performance of the LS estimator

Concerning the achievability of the CR bound with a practical estimator, [22, Proposition 2] have demonstrated that this bound cannot be attained, meaning that for any unbiased estimator $\tau_n(\cdot)$ we have that

$$Var(\tau_n(I^n)) > \sigma^2_{CR}, \tag{2.11}$$

where I^n follows the Poisson PMF $f_{(x_c,\tilde{F})}$ in Eq. (2.5).

This finding should be interpreted with caution, considering its pure theoretical meaning. This is because Eq. (2.11) does not exclude the possibility that the CR bound could be approximated arbitrarily close by a practical estimation scheme. Motivated by this refined conjecture, [22] proposed to study the performance of the widely adopted LS estimator[5] with the goal of deriving operational upper and lower performance bounds of its performance that could be used to determine how far could this scheme depart from the CR limit. Then, from this result, it was possible to evaluate the goodness of the LS estimator for concrete observational regimes. For bounding the performance of the LS estimator, the challenge was

[5]This is the solution of $\tau_{LS}(I^n) = \arg\min_{\alpha \in \mathbb{R}} \sum_{\mathrm{i}=1}^{n} (I_{\mathrm{i}} - \lambda_{\mathrm{i}}(\alpha))^2$, with $\lambda_{\mathrm{i}}(\alpha) = \tilde{F} g_{\mathrm{i}}(\alpha) + \tilde{B}_{\mathrm{i}}$, α being a generic variable representing the astrometric position, $g_{\mathrm{i}}(\cdot)$ is given by Eq. (2.3), and where arg min represents the argument that minimizes the expression. More details are presented in [22].

that $\tau_{LS}(I^n)$ is an implicit function of the data (where no close-form expression is available) and, consequently, [22] derived a result to bound the estimation error and the variance of $\tau_{LS}(I^n)$. We can briefly summarize the main result presented in [22, Theorem 1] saying that under certain mild sufficient conditions (that were shown to be realistic for astrometry), there is a constant $\delta > 0$ (that depends on the observational regime, in particular the S/N) and a nominal variance σ^2_{LS}, which is determined in closed-form in the result, from which it is possible to bound $Var(\tau_{LS}(I^n))$ by the simple expression

$$Var(\tau_{LS}(I^n)) \in \left(\frac{\sigma^2_{LS}(n)}{(1+\delta)^2}, \frac{\sigma^2_{LS}(n)}{(1-\delta)^2} \right), \tag{2.12}$$

where

$$\sigma^2_{LS}(n) = \frac{\sum_{\mathrm{i}=1}^{n} (\tilde{F} g_{\mathrm{i}}(x_c) + \tilde{B}_{\mathrm{i}}) \cdot (g'_{\mathrm{i}}(x_c))^2}{\left(\tilde{F} \sum_{\mathrm{i}=1}^{n} (g'_{\mathrm{i}}(x_c))^2 \right)^2}. \tag{2.13}$$

Note that when δ is small, $\sigma^2_{LS}(n)$ tightly determines the performance of the LS estimator, and its comparison with σ^2_{CR} can be used to evaluate the goodness of the LS estimator for astrometry. Based on a careful comparison, it was shown in [22, Sect. 4] that in general $\sigma^2_{LS}(n)$ is close to $\sigma^2_{CR}(n)$ for the small S/N regime of the problem. However for moderate to high S/N regimes, the gap between $\sigma^2_{LS}(n)$ and $\sigma^2_{CR}(n)$ becomes quite significant[6].

These unfavorable findings for the LS method have motivated us to study alternatives schemes that could potentially approach better the Cramér-Rao bound for the rich observational context of medium to high S/N regimes. This will be the focus of the following sections, where in particular we explore the performance of the ML and WLS estimators, thus extending and generalizing the analysis done for the LS estimator by our group presented in [22].

2.2 Bounding the performance of an implicit estimator

Before we go to the case of the WLS and the ML estimators, we present a general result that bounds the performance of *any* estimator that is the solution of a generic optimization problem. Let us consider a vector of observations $I^n = (I_1, \ldots, I_n) \in \mathbb{R}^n$ and a general so-called cost function $J(\alpha, I^n)$. Then the estimation of x_c from the data is the solution of the following optimization problem:

$$\tau_J(I^n) \equiv \arg\min_{\alpha \in \mathbb{R}} J(\alpha, I^n), \tag{2.14}$$

where α represents the position of the object in the context of astrometry. As in our previous work [22], the challenge here is that this estimator is implicit because no closed-form expression of the data which solves Eq. (2.14) is assumed. In particular, this implies that both the variance and the estimation error of $\tau_J(I^n)$ can not be determined directly. To

[6] In particular, for the very high S/N regime and assuming $\Delta x/\sigma \ll 1$, [22, Proposition 3] shows that this gap reaches the condition $\frac{\sigma^2_{LS}(n)}{\sigma^2_{CR}(n)} \approx \frac{8}{3\sqrt{3}}$.

address this technical issue, we extend the approach proposed by Fessler et al. [50] to approximate the variance and the mean of an implicit estimator solution of a problem described by Eq. (2.14) through the use of a Taylor approximation around the mean measurement, i.e., $\bar{I}^n = \mathbb{E}_{I^n \sim f_{(x_c, \bar{F})}}(I^n)$.

More precisely, we assume that $J(\alpha, I^n)$ has a unique global minimum at $\tau_J(I^n)$, and that it has a regular behavior, so its partial derivatives are zero, i.e.,

$$0 = \frac{\partial}{\partial \alpha} J(\alpha, I^n)\Big|_{\alpha=\tau_J(I^n)} \equiv \frac{\partial}{\partial \alpha} J(\tau_J(I^n), I^n). \tag{2.15}$$

Then we can obtain $\tau_J(I^n)$ by a first order Taylor expansion around the mean $\bar{I}^n$ by

$$\tau_J(I^n) = \tau_J(\bar{I}^n) + \sum_{\mathrm{i}=1}^{n} \frac{\partial}{\partial I_{\mathrm{i}}} \tau_J(\bar{I}^n)(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}}) + \underbrace{\frac{1}{2} \sum_{\mathrm{i}=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial I_{\mathrm{i}} \partial I_j} \tau_J(\bar{I}^n - t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j)}_{\equiv \mathrm{e}(\bar{I}, I - \bar{I})}. \tag{2.16}$$

with $t \in [0, 1]$ is fixed but unknown[7]. For simplicity, Eq. (2.16) can be written in matrix form as

$$\tau_J(I^n) = \tau_J(\bar{I}^n) + \nabla \tau_J(\bar{I}^n) \cdot (I^n - \bar{I}^n) + \mathrm{e}(\bar{I}^n, I^n - \bar{I}^n), \tag{2.17}$$

where $\nabla = [\frac{\partial}{\partial I_1} \dots \frac{\partial}{\partial I_n}]$ denotes the row gradient operator and $\mathrm{e}(\bar{I}^n, I - \bar{I}^n)$ is the residual error of the Taylor expansion. From Eq. (2.17) we can readily obtain the following expression for its variance

$$Var\{\tau_J(I^n)\} = \underbrace{\nabla \tau_J(\bar{I}^n) Cov\{I^n\} \nabla \tau_J(\bar{I}^n)^T}_{\equiv \sigma_J^2(n)} \tag{2.18}$$

$$+ \underbrace{Var\{\mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\} + 2Cov\{\nabla \tau_J(\bar{I}^n)(I^n - \bar{I}^n), \mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\}}_{\equiv \gamma_J(n)}. \tag{2.19}$$

In Eq. (2.18) we recognize two terms: $\sigma_J^2(n)$ that captures the linear behaviour of $\tau_J(\cdot)$ around $\bar{I}^n$ and $\gamma_J(n)$ which reflects the deviation from this linear trend. It should be noted that the above expression does not depend on $\tau(I^n)$ itself, but on its partial derivatives evaluated at the mean vector of observations. Then in the adoption of this approach to estimate $Var\{\tau_J(I^n)\}$, a key task is to determine $\nabla \tau_J(\bar{I}^n)$.

Remark 1 It is meaningful to note that [50] only considered the linear term in his approximate analysis, obviating the residual term $\gamma_J(n)$ in Eq. (2.18). This first order reduction is not realistic for our problem because the solution of a problem like the one posed by Eq. (2.14) in astrometry has important non-linear components that need to be considered in the analysis of Eq. (2.18).

In an effort to analyze both the linear and non-linear aspects of a general intrinsic estimator solution to Eq. (2.14), the following result offers sufficient conditions to determine $\sigma_J^2(n)$ in closed-form, and to bound the magnitude of the residual term $\gamma_J(n)$ in Eq. (2.18).

[7] It follows that $\lim_{I \to \bar{I}} \frac{\mathrm{e}(\bar{I}, I - \bar{I})}{||I - \bar{I}||_2} = 0$.

Theorem 2.2 *Let us consider a fixed and unknown parameter $x_c \in \mathbb{R}$, the observations $I^n = (I_1, ..., I_n)^T$ where $I_{\rm i} \sim f_{x_c}$, and $\tau_J(I^n)$ the estimator solution of Eq. (2.14). If we satisfy the following two rather general conditions:*

a) the cost function $J(\alpha, I^n)$ is twice differentiable with respect to I^n and x_c, and the gradient of $\tau_J(\cdot)$ evaluated in the mean data $\bar{I}^n$ offers the following decomposition

$$\nabla \tau_J(\bar{I}^n) \cdot (I^n - \bar{I}^n) = a \sum_{\rm i=1}^{N} b_{\rm i}(I_{\rm i} - \bar{I}_{\rm i}) \tag{2.20}$$

with a and $\{b_{\rm i} : {\rm i} \in \{1, ..., N\}\}$ constants, and,

b) the estimator evaluated in the mean data equals the true parameter $x_c \in \mathbb{R}$, this is,

$$\tau_J(\bar{I}^n) = x_c, \tag{2.21}$$

then we can define two new quantities $\epsilon_J(n)$ and $\beta_J(n)$ (both > 0) and $\sigma_J^2(n)$ in Eq. (2.18) with analytical expressions (details presented in Appendix 2.6.1) such that

$$|\underbrace{\mathbb{E}_{I^n \sim f_{x_c}}\{\tau_J(I^n)\} - x_c}_{bias}| \leq \epsilon_J(n) \tag{2.22}$$

and

$$Var_{I^n \sim f_{x_c}}\{\tau_J(I^n)\} \in \left(\sigma_J^2(n) - \beta_J(n), \sigma_J^2(n) + \beta_J(n)\right). \tag{2.23}$$

The proof of this result and the expression for $(\epsilon_J(n), \sigma_J^2(n), \beta_J(n))$ in Eqs. (2.22) and (2.23) are presented in detail in Appendix 2.6.1.

Revisiting the equality in Eq. (2.18), Theorem 2.2 provides general sufficient conditions to bound the residual term $\gamma_J(n)$ and by doing that, a way of bounding the variance of $\tau_J(I^n)$ which is the solution of Eq. (2.14). In particular, it is worth noting that if the ratio $\frac{\beta_J(n)}{\sigma_J^2(n)} \ll 1$, then Eq. (2.23) offers a tight bound for $Var_{I^n \sim f_{x_c}}\{\tau_J(I^n)\}$. In this last context, $\sigma_J^2(n)$ (called the nominal value of the result) provides a very good approximation for $Var_{I^n \sim f_{x_c}}\{\tau_J(I^n)\}$.

On the application of this result to the WLS and ML estimators, we will see that the main assumption in Eq. (2.20) is satisfied in both cases (see Eqs. (2.64) and (2.81) in Appendix 2.6.2 and 2.6.3, respectively), and from that $\sigma_J^2(n)$ is playing an important role to approximate the performance of ML and WLS in a wide range of observational regimes. In addition, the analysis of the bias in Eq. (2.22) shows that these estimators are unbiased for any practical purpose and, consequently, contrasting their performance (estimation error $\sim \sqrt{\text{variance}}$) with the CR bound is a meaningful way to evaluate optimality.

2.3 Application to astrometry

In this section we apply Theorem 2.2 to bound the variances of the ML and WLS estimators in the context of astrometry. Following the model presented in Sect. 2.1.1, $I^n = (I_1, \ldots, I_n)^T$

denotes the measurements acquired by each pixel of the array, and where each of them follows a Poisson distribution given by

$$I_{\mathrm{i}} \sim Poisson(\lambda_{\mathrm{i}}(x_c)), \quad \mathrm{i} = 1, \ldots, n, \tag{2.24}$$

as expressed by Eqs. (2.2) and (2.3).

2.3.1 Bounding the variance of the WLS estimator

The WLS estimator, denoted by $\tau_{WLS}(I^n)$ in Eq. (2.26), is implicitly defined through a cost function given by

$$J_{WLS}(\alpha, I^n) = \sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}}(I_{\mathrm{i}} - \lambda_{\mathrm{i}}(\alpha))^2, \tag{2.25}$$

where $(w_1, \ldots, w_n)^T \in \mathbb{R}_+^n$ is a weight vector, and α is a general source position parameter. Specifically we have that

$$\tau_{WLS}(I^n) = \arg\min_{\alpha \in \mathbb{R}} J_{WLS}(\alpha, I^n). \tag{2.26}$$

Applying Theorem 2.2 we obtain the following result:

Theorem 2.3 *Let us consider the WLS estimator solution of Eq. (2.26), then we have that*

$$|\underbrace{\mathbb{E}_{I^n \sim f_{x_c}}\{\tau_{WLS}(I^n)\} - x_c}_{bias}| \leq \epsilon_{WLS}(n) \tag{2.27}$$

and

$$Var_{I^n \sim f_{x_c}}\{\tau_{WLS}(I^n)\} \in \left(\sigma^2_{WLS}(n) - \beta_{WLS}(n), \sigma^2_{WLS}(n) + \beta_{WLS}(n)\right), \tag{2.28}$$

where $\sigma^2_{WLS}(n)$ is given by

$$\sigma^2_{WLS}(n) = \frac{\sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}}^2 \lambda_{\mathrm{i}}(x_c) \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2 \Big|_{\alpha = x_c}}{\left(\sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2 \Big|_{\alpha = x_c}\right)^2} \tag{2.29}$$

and $\beta_{WLS}(n)$ and $\epsilon_{WLS}(n)$ are well defined analytical expression of the problem (presented in Appendix 2.6.2).

The proof of this result and the expressions for $\epsilon_{WLS}(n)$ and $\beta_{WLS}(n)$in Eqs. (2.27) and (2.28), respectively, are elaborated in Appendix 2.6.2.

This result offers concrete expressions to bound the bias as well as the variance of the WLS estimator. For the bias bound in Eq. (2.27), it will be shown that $\epsilon_{WLS}(n)$ is very small (order of magnitudes smaller than x_c) for all the observational regimes explored in this work and, consequently, the WLS can be considered an unbiased estimator in astrometry, as it would be expected. Concerning the bounds for the variance in Eq. (2.28), we will show that for high and moderate S/N regimes the ratio $\beta_{WLS}(n)/\sigma^2_{WLS}(n) \ll 1$ and consequently

in this context $\sigma^2_{WLS}(n)$ is a precise estimator of $Var_{I^n\sim f_{x_c}}\{\tau_{WLS}(I^n)\}$. For the very small S/N the results offers an admissible interval $\sigma^2_{WLS}(n) \pm \beta_{WLS}(n)$ around the nominal value $\sigma^2_{WLS}(n)$. Therefore in any context $\sigma^2_{WLS}(n)$ shows to be a meaningful approximation for the performance of the WLS.

Remark 2 If we focus on the analysis on the closed form expression $\sigma^2_{WLS}(n)$ as an approximation of $Var_{I^n\sim f_{x_c}}\{\tau_{WLS}(I^n)\}$ and we compare it with the CR bound $\sigma^2_{CR}(n)$ in Eqs. (2.10) and (2.11), we note that they are very similar in their structure. In particular, it follows that $\sigma^2_{WLS}(n) = \sigma^2_{CR}(n)$, if an only if, the weights of the WLS estimator are selected in the following way

$$w_{\mathrm{i}} = K \cdot \frac{1}{\lambda_{\mathrm{i}}(x_c)}, \quad \forall \mathrm{i} \in \{1, \ldots, n\}. \tag{2.30}$$

where K is an arbitrary constant ($K > 0$). In other words, the only way in which the performance of the WLS approximates the CR limit is if we select the weights as in Eq. (2.30). However, this selection needs the information of the true position x_c, which is unfeasible as it contradicts the very essence of the inference task (indeed, x_c is unknown, and we are trying to estimate it from the data). Another interpretation is that no matter how we choose the weights of the WLS estimator, it is not possible that the WLS is close to the CR bound for every position x_c, telling us that the WLS is intrinsically not optimal from the perspective of being close to the CR limit in all the possible astrometric scenarios. In particular, this impossibility result is very strong in the high S/N regimes where $\sigma^2_{WLS}(n) \approx Var_{I^n\sim f_{x_c}}\{\tau_{WLS}(I^n)\}$. This implication is consistent with the analysis presented by [22, Fig. 4], where it was shown that the variance of the LS estimator is significantly higher than then CR bound in the high S/N regime. This justifies the study of the ML estimator.

2.3.2 Bounding the variance of the ML estimator

The ML estimator, denoted by $\tau_{ML}(I^n)$ in Eq. (2.32), is implicitly defined through a cost function

$$J(\alpha, I^n) = \sum_{\mathrm{i}=1}^{n} I_{\mathrm{i}} \ln(\lambda_{\mathrm{i}}(\alpha)) - \lambda_{\mathrm{i}}(\alpha), \tag{2.31}$$

where α is a general source position parameter. Specifically, given an observation I^n we have that

$$\begin{aligned} \tau_{ML}(I^n) &= \arg\max_{\alpha\in\mathbb{R}} J(\alpha, I^n), \\ &= \arg\min_{\alpha\in\mathbb{R}} \sum_{\mathrm{i}=1}^{n} -I_{\mathrm{i}} \ln(\lambda_{\mathrm{i}}(\alpha)) + \lambda_{\mathrm{i}}(\alpha). \end{aligned} \tag{2.32}$$

Applying Theorem 2.2 we obtain the following result:

Theorem 2.4 *Let us consider the ML estimator solution of Eq. (2.32), then we have that*

$$|\underbrace{\mathbb{E}_{I^n\sim f_{x_c}}\{\tau_{ML}(I^n)\} - x_c}_{bias}| \leq \epsilon_{ML}(n) \tag{2.33}$$

and

$$Var_{I^n \sim f_{x_c}}\{\tau_{ML}(I^n)\} \in \left(\sigma^2_{ML}(n) - \beta_{ML}(n), \sigma^2_{ML}(n) + \beta_{ML}(n)\right), \tag{2.34}$$

where

$$\sigma^2_{ML}(n) = \sigma^2_{CR}(n) = \left(\sum_{\mathrm{i}=1}^{n} \frac{\left(\tilde{F}\frac{\mathrm{d}g_{\mathrm{i}}(x_c)}{\mathrm{d}x_c}\right)^2}{\tilde{F}g_{\mathrm{i}}(x_c) + \tilde{B}_{\mathrm{i}}} \right)^{-1}, \tag{2.35}$$

and $\beta_{ML}(n)$ and $\epsilon_{ML}(n)$ are well defined analytical expression of the problem (presented in Appendix 2.6.3).

The proof of this result and the expressions for $\epsilon_{ML}(n)$ and $\beta_{ML}(n)$ in Eqs. (2.33) and (2.34), respectively, are elaborated in Appendix 2.6.3.

Remark 3 It is important to mention that the magnitude of $\epsilon_{ML}(n)$ is orders of magnitude smaller than x_c in all the observational regimes studied in this work (see this analysis in Sect. 2.4) and, consequently, for any practical purpose the ML is an unbiased estimator. This implies that the comparison with the CR bound is a meaningful indicator when evaluating the optimality of the ML estimator.

Remark 4 We observe that if the ratio $\beta_{ML}(n)/\sigma^2_{ML}(n)$ is significantly smaller than one, which is shown in Sect. 2.4 from medium to high S/N regimes, then $Var_{I^n \sim f_{x_c}}\{\tau_{ML}(I^n)\} \approx \sigma^2_{ML}(n)$. This is a very interesting result because we can approximate the performance of the ML estimator with $\sigma^2_{ML}(n)$. On this context, it is remarkable to have that the nominal value $\sigma^2_{ML}(n)$ is precisely the CR bound (see Eq. (2.35)), because this means that the ML estimator closely approximate this MVB in the interesting regime from moderate to very high S/N. Note that this medium-high S/N regime is precisely the context where the LS estimator shows significant deficiencies as presented in [22]. Therefore, ML offers optimal performances in the regime where LS type of methods are not able to match the CR bound, which satisfactorily resolves the question posted by [22] on the study of schemes that could very closely approach the CR bound in the high S/N regime.

2.4 Numerical analysis

In this section we evaluate numerically the performance bounds obtained in Sect. 2.3 for the WLS and ML estimators, and compare them with the astrometric CR bound in Proposition 2.1. The idea is to consider some realistic astrometric conditions to evaluate the expressions developed in Theorems 2.3 and 2.4 and their dependency on important observational conditions and regimes. As we shall see, key variables in this analysis are the tradeoff between the intensity of the object and the noise represented by the S/N value, and the pixel resolution of the CCD.

2.4.1 Experimental setting

We adopt some realistic design and observing variables to model the problem [20,21]. For the PSF, analytical and semi-empirical forms have been introduced, see for instance the ground-based model in [55] and the space-based models by [8] or [56]. In this work we will adopt a Gaussian PSF, i.e., $\phi(x,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x)^2}{2\sigma^2}}$ in Eq. (2.3), and where σ is the width of the PSF and is assumed to be known. This PSF has been found to be a good representation for typical astrometric-quality Ground-based data [57]. In terms of nomenclature, $FWHM \equiv 2\sqrt{2\ln 2}\,\sigma$ measured in arcsec, denotes the *Full-Width at Half-Maximum* parameter, which is an overall indicator of the image quality at the observing site [58].

The background profile, represented by $\left\{\tilde{B}_{\mathrm{i}}, \mathrm{i} = 1,..,n\right\}$ in Eq. (2.2), is a function of several variables, like the wavelength of the observations, the moon phase (which contributes significantly to the diffuse sky background), the quality of the observing site, and the specifications of the instrument itself. We will consider a uniform background across pixels underneath the PSF, i.e., $\tilde{B}_{\mathrm{i}} = \tilde{B}$ for all i. To characterize the magnitude of $\tilde{B}$, it is important to first mention that the detector does not measure photon counts [e^-] directly, but a discrete variable in "*Analog to Digital Units* (ADUs)" of the instrument, which is a linear proportion of the photon counts ([54]). This linear proportion is characterized by the gain of the instrument G in units of [e^-/ADU]. G is just a scaling value, where we can define $F \equiv \tilde{F}/G$ and $B \equiv \tilde{B}/G$ as the intensity of the object and noise, respectively, in the specific ADUs of the instrument. Then, the background (in ADUs) depends on the pixel size Δx arcsec as follows

$$B = f_s \Delta x + \frac{D + RON^2}{G}\ [ADU], \tag{2.36}$$

where f_s is the (diffuse) sky background in ADU arcsec^{-1}, while D and RON^2, both measured in e^- model the dark-current and read-out-noise of the detector on each pixel, respectively. Note that the first component in Eq. (2.36) is attributed to the site, and its effect is proportional to the pixel size. On the other hand, the second component is attributed to errors of the PID (detector), and it is pixel-size independent. This distinction is central when analyzing the performance as a function of the pixel resolution of the array (see details in [20, Sect. 4]). More important is the fact that in typical ground-based astronomical observation, long exposure times are considered, which implies that the background is dominated by diffuse light coming from the sky, and not from the detector [20, Sect. 4].

For the experimental conditions, we consider the scenario of a ground-based station located at a good site with clear atmospheric conditions and the specification of current science-grade CCDs, where $f_s = 1502.5$ ADU arcsec^{-1}, $D = 0$, $RON = 5$ e^-, $FWHM = 1$ arcsec (equivalent to $\sigma = 1/(2\sqrt{2\ln 2})$ arcsec) and $G = 2$ e^- ADU^{-1} (with these values we will have $B = 313$ ADU for $\Delta x = 0.2$ arcsec using Eq. (2.36)). In terms of scenarios of analysis, we explore different pixel resolutions for the CCD array $\Delta x \in [0.1, 0.65]$ measured in arcsec, and different signal strengths $\tilde{F} \in \{1080, 3224, 20004, 60160\}$, measured in e^-, which corresponds to $S/N \in\sim \{12, 32, 120, 230\}$. Note that increasing $\tilde{F}$ implies increasing the S/N of the problem, which can be approximately measured by the ratio $\tilde{F}/\tilde{B}$. On a given detector plus telescope setting, these different S/N scenarios can be obtained by appropriately changing the exposure time (open shutter) that generates the image.

2.4.2 Bias analysis

Considering the upper bound terms $\epsilon_{WLS}(n)$ and $\epsilon_{ML}(n)$ for the bias error obtained from Theorems 2.3 and 2.4 for the WLS and ML, respectively, Fig. 2.1 presents the relative bias error for different S/N regimes and pixel resolutions. In the case of the ML estimator, the bounds for relative bias error are very small in all the explored S/N regimes and pixel resolutions meaning that for any practical purposes this estimator is unbiased as expected from theory [17]. For the case of the WLS, we observe that from medium to high S/N the relative error bound obtained is very small but, meanwhile at low S/N, unbiasedness can not be fully guaranteed from the bound in Eq. (2.27). In general, our results show that both WLS and ML are unbiased estimators for astrometry in a wide range of relevant observational regimes (in particular from medium to high S/N) and, consequently, it is meaningful to analyze the optimality of these estimators in comparison with the CR bound in those regimes.

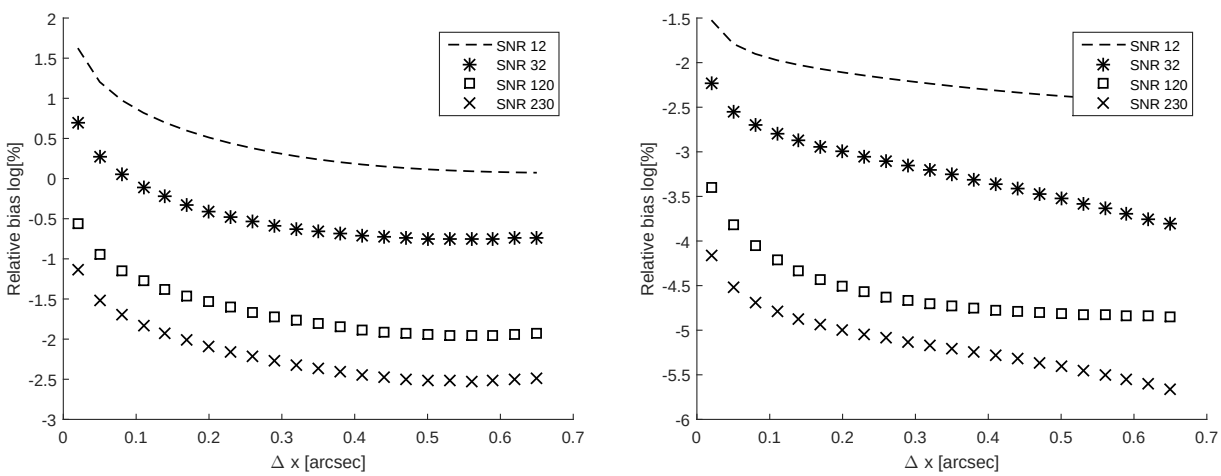

Figure 2.1: Relative performance of the bias (as measured by $\log\left(100 \times \frac{\epsilon_J(n)}{x_c}\right)$) stipulated by Theorem 2.2 for the WLS estimator (left side, Eq. (2.27)) and the ML estimator (right side, Eq. (2.33)). Results are reported for different values of the source flux $\tilde{F} \in \{1080, 3224, 20004, 60160\}$, all in e^- (top to bottom symbols respectively), as a function of the detector pixel size. The 0% level corresponds to having achieved no bias.

In the following sections, we move to the analysis of the variance of the WLS and ML with particular focus on the medium to high S/N regimes across all pixel resolutions using the performance bounds derived in Eqs. (2.28) and (2.34), respectively.

2.4.3 Performance analysis of the WLS estimator

In this section, we evaluate numerically the expression derived in Theorem 2.3 to bound the variance of the WLS estimator in Eq. (2.28). For that we characterize the admissible regime

predicted for the variance of the WLS estimator, i.e., the interval

$$\left(\sigma^2_{WLS}(n) - \beta_{WLS}(n), \sigma^2_{WLS}(n) + \beta_{WLS}(n)\right), \tag{2.37}$$

for $S/N \in \{12, 32, 120, 230\}$ and $\Delta x \in [0.01, 0.65]$ arcsec. In these bounds, we recognize its central value (or nominal value) $\sigma^2_{WLS}(n)$ in Eq. (2.29) and the length of the interval $2\beta_{WLS}(n)$ that is determined in closed form for its numerical evaluation in Eqs. (2.43) and (2.67). Note that $2\beta_{WLS}(n)$ can be considered an indicator of the precision of our result to approximate the variance of the WLS in astrometry.

Revisiting the uniform weight case

To begin the analysis, we consider the setting of uniform weights across pixels, i.e., the case of the LS estimator and, without loss of generality, we locate the object in the center of the field of view[8], which can be considered a reasonable scenario to represent the complexity of the astrometry task. At this point, it is important to remind the reader that from the analysis of $\sigma^2_{WLS}(n)$ in Remark 2, the nominal value $\sigma^2_{WLS}(n)$ is equivalent to the CR bound when the w_{i} are selected as a function of the true position in Eq. (2.30). In view of this observation, the selection of non-uniform weights can be interpreted as biasing the estimation to a particular area of the angular space, which goes in contradiction with the essence of the inference problem that estimates the position with no prior information, and only relies on the measured counts. From this interpretation, revisiting the LS estimator is an important first step in the analysis of the WLS framework.

On the specifics, the boundaries of the interval in Eq. (2.37) and its nominal values are illustrated in Fig. 2.2 for the different observational regimes. In addition, Fig. 2.2 shows the CR bound as a reference to evaluate the optimality of the LS scheme across settings. We observe that for the low $S/N \sim 12$ regime, the nominal values precisely match the CR bound, however our result is not conclusive as the interval around the nominal performance is significantly large. This is the regime where our result is not conclusive regarding the performance of the LS estimator. In the regime $S/N \in (30, 50)$ (top right panel on Fig. 2.2), we notice an important reduction on the range of admissible performance, and our result becomes more informative and meaningful. In this context, the nominal values is very close to the CR bound, and we could assert that the LS estimator offers sufficiently good performance in the sense that is very competitive with the MVB. Importantly, when we move to the regime of relatively high S/N and very hight S/N (from 100 to 300), our results is very accurate to predict the performance of the LS method, and we find that the gap between the CR and the nominal value is very significant (the deviation from the MVB is 16% and 30% for S/N 120 and 230 at $\Delta x = 0.2$ arcsec, respectively). This last result confirms one of the main findings presented in [22], who showed that for medium to hight S/N the LS estimator is suboptimal with respect to the MVB. In Fig. 2.2 we also show the square root of the empirical variance ($Var(\tau_{\hat{WLS}})$) with respect to the empirically-determined mean position $\hat{x}_c$ (using the WLS estimator), all as derived from the simulations, showing good consistency with our predictions.

[8]It is important to remind that the CR bound is a function of the value of the parameter to be estimated, in this case the position x_c, see the Fisher information in Eq. (2.10).

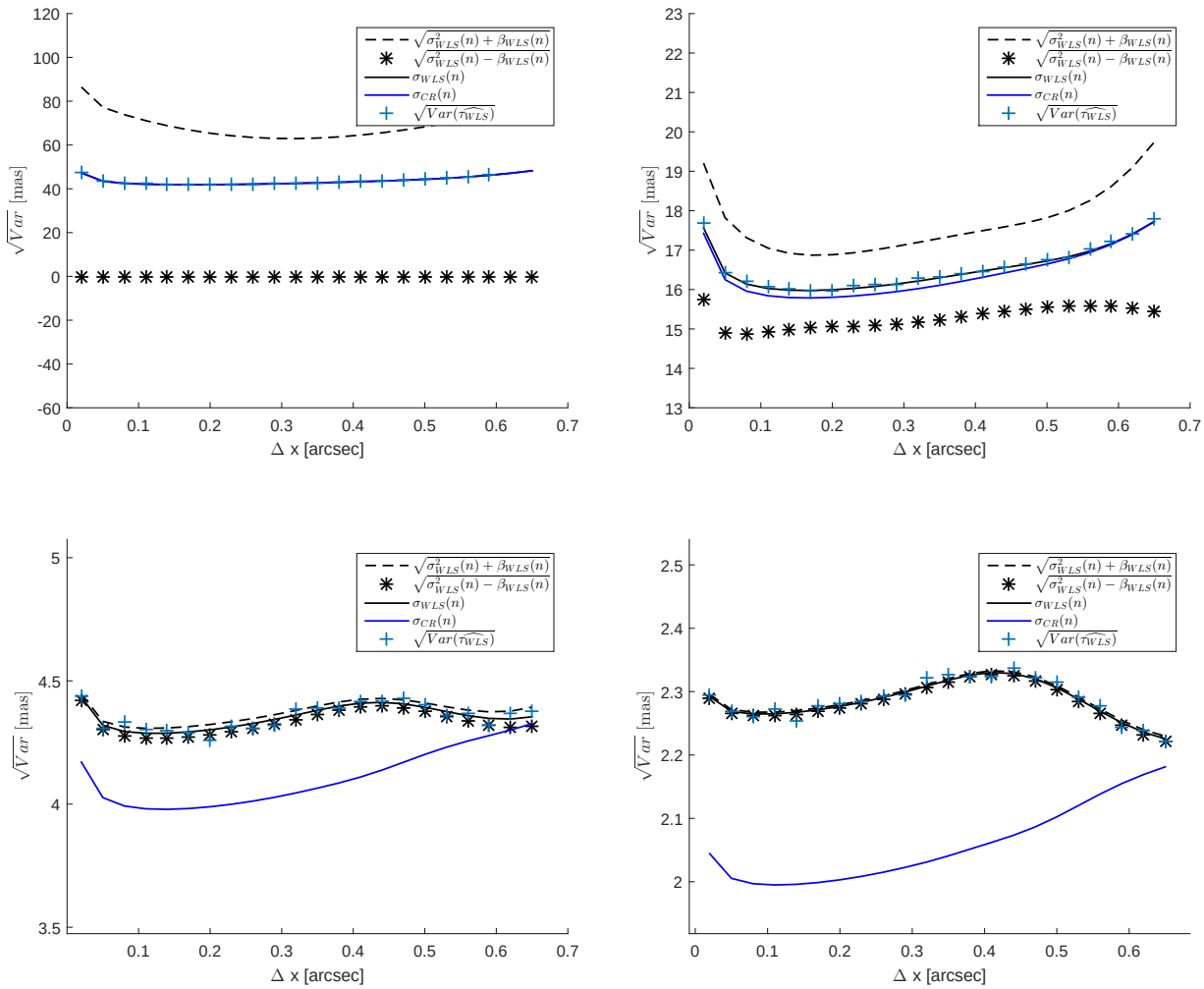

Figure 2.2: Range of the square root of the variance performance (in miliarcsecond=mas) for the WLS method in astrometry using uniform weights (equivalent to the LS method) predicted by Theorems 2.2 and 2.3, Eq. (2.28). Results are reported for different representative values of $\tilde{F}$ and across different pixel sizes (top-left to bottom-right): $\tilde{F} \in \{1080, 3224, 20004, 60160\}$ e^-.

Non-uniform weight case

The sub-optimality of the WLS scheme from moderate to high S/N seen in Fig. 2.2 can be extended for any arbitrary selection of a fixed set of weights, as it would be expected. Given that the space of weights selection is literally unlimited, we use the insight obtained from Remark 2 that states that a selection of weights can be interpreted as an specific prior on the position of the object where the optimum, but unfortunately unknown, selection (achieving the CR bound) of weights in Eq. (2.30) is an explicit function of the unknown position of the object. Then, we consider a finite set of positions $\{x_{c,1}, .., x_{c,M}\}$ that uniformly partition the field of view, and their corresponding weights sets $\{w_{\mathrm{i}}(x_{c,1}) : 1 = 1, .., n\}$, $\{w_{\mathrm{i}}(x_{c,2}) : 1 = 1, .., n\} \ldots$ and $\{w_{\mathrm{i}}(x_{c,M}) : 1 = 1, .., n\}$ using Eq. (2.30) to cover a representative collections of weights for the problem of astrometry.

Then, for a specific selection of weights in our admissible collection (attributed to a prior believe of the position of the object in the field of view), we evaluate the worse discrepancy between $\sigma^2_{WLS}(n) - \beta_{WLS}(n)$ (which is the most favourable expression for the variance predicted from Theorem 2.3), and the CR bound in Proposition 2.1, across a collection of presumed object positions in the following range of positions

$$\Theta = \{x_o^* - \sigma, x_o^* - 0.8 * \sigma, x_o^* - 0.6 * \sigma, x_o^* - 0.4 * \sigma, x_o^* - 0.2 * \sigma, x_o^*\}$$

where x_o^* denotes the center of the array (which, as indicated at the beginning of Sect. 2.4.3, is equal to the true object position x_c) and $\sigma = FWHM/2\sqrt{2\ln 2}$ is the dispersion parameter of the PSF. The idea of using this worse case difference is justified from the fact that in this parameter estimation problem we do not know the position of the object, and consequently, the optimality of any WLS estimator should be evaluated in the worse case situation, as the scheme should be able to estimate the position of the object in any scenario (position). More precisely, for a given weight selection $\{w_{\mathrm{i}}, \mathrm{i} = 1, .., n\}$, we use the following worse case discrepancy

$$\sup_{x_c \in \Theta} \frac{(\sigma^2_{WLS}(n) - \beta_{WLS}(n)) - \sigma^2_{CR}(n)}{\sigma^2_{CR}(n)}. \tag{2.38}$$

For this analysis note that both $(\sigma^2_{WLS}(n) - \beta_{WLS}(n))$ and $\sigma^2_{CR}(n)$ are functions of the position x_c, Δx, and S/N.

Fig. 2.3 illustrates the worse case discrepancy in Eq. (2.38) for the medium and high S/N regimes where Theorem 2.3 provides an accurate and meaningful prediction of the performance of the WLS method, i.e., for $S/N \in \{120, 230\}$, and across $\Delta x \in [0.05, 0.7]$ arcsec. The discrepancy is quite significant, in the order of 37% and 60% in the range for $\Delta x \in [0.1, 0.3]$ arcsec for S/N 120 and 230, respectively.

To refine the worse case analysis presented in Fig. 2.3, and to evaluate in more detail the sensitivity of the discrepancy indicator given by Eq. (2.38), we evaluate the discrepancy of WLS using the weights associated to x_o^* (the center position of the array) with respect to the CR bound associated to the positions $\{x_o^* - \sigma, x_o^* - 0.8 * \sigma, x_o^* - 0.6 * \sigma, x_o^* - 0.4 * \sigma, x_o^* - 0.2 * \sigma, x_o^*\}$, to study how the discrepancy (measuring the non-optimality of the method) evolves when the adopted position moves far from the prior imposed by the WLS in the center of the array. Fig. 2.4 illustrates this behaviour, where it is possible to see that the discrepancy is

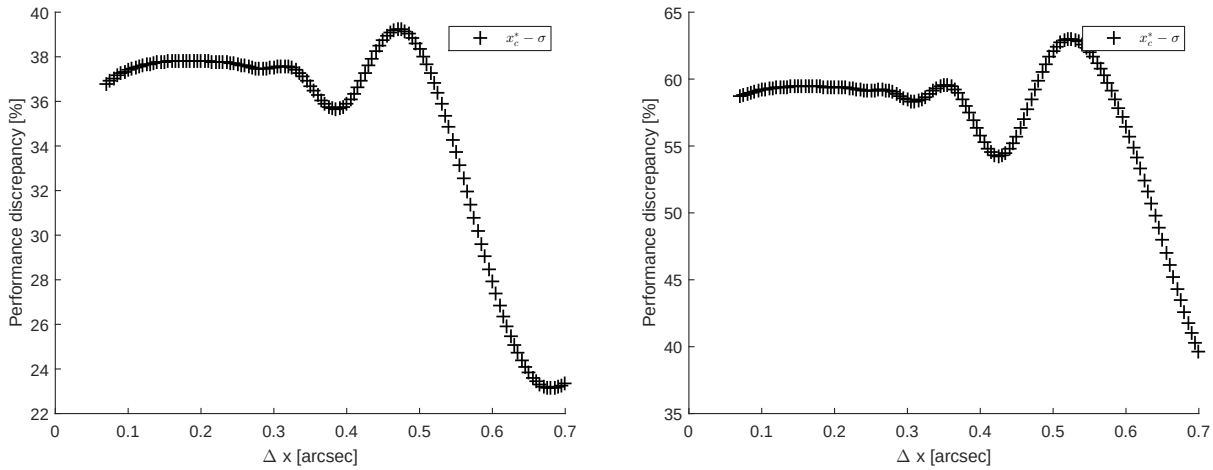

Figure 2.3: Worse case discrepancies in Eq. (2.38) for the WLS estimator using the weights set indexed by the positions $\Theta = \{x_o^* - \sigma, x_o^* - 0.8 * \sigma, x_o^* - 0.6 * \sigma, x_o^* - 0.4 * \sigma, x_o^* - 0.2 * \sigma, x_o^*\}$. Results are reported for two S/N scenarios, namely $\tilde{F} = 20004$ e$^-$ (Left) and $\tilde{F} = 60160$ e$^-$ (Right), and across different pixel sizes.

very sensitive and proportional to the misassumption of the object position, where the worse case discrepancy in the maximum achievable location precision is on the order of 40% for pixel sizes in the range $[0.1, 0.6]$ arcsec for $S/N \sim 120$, and about 60% for pixel sizes in the range $[0.1, 0.6]$ arcsec for $S/N \sim 230$. These worse case scenario happens in both cases when the object is located the farthest from the prior assumption, as it would be expected.

The main conclusion derived form this CR bound analysis is that, independent of the weight selection adopted, as long as the weights are fixed, the WLS estimator is not able to achieve the CR bound in all observational regimes. More precisely, the discrepancy (measuring the non-optimality) in the less favourable case of an hypothetical and feasible position of the object is very significant, in the range of $40\% - 60\%$ for the important regime of high and very high S/N.

2.4.4 Performance analysis of the ML estimator

In this section we perform the same analysis done for the WLS in Sect. 2.4.3, but using the result in Theorem 2.4. In particular, we consider the admissible regime for the variance of the ML estimator given by

$$\left(\sigma_{ML}^2(n) - \beta_{ML}(n), \sigma_{ML}^2(n) + \beta_{ML}(n)\right)$$

in Eq. (2.34), where the nominal value in this case, $\sigma_{ML}^2(n)$ in Eq. (2.35), is precisely the CR bound, while the length of the interval $2\beta_{ML}(n)$ is given by Eqs. (2.43) and (2.84) in closed form.

Considering an object located in the center of the array, i.e., $x_c = x_o$, the performance curves are presented in Fig. 2.5 for $S/N \in \{12, 32, 120, 230\}$ and $\Delta x \in [0.1, 0.65]$ arcsec.

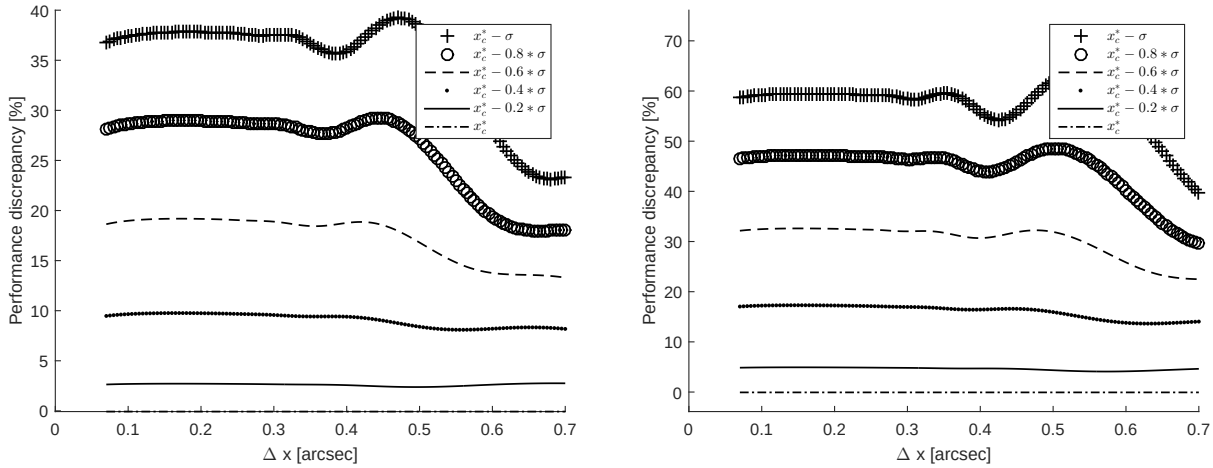

Figure 2.4: Performance discrepancies (measuring the non-optimality) of the WLS estimator using the center position as a prior for the weight selection with respect to the CR bound obtained for the true object positions $\{x_o^* - \sigma, x_o^* - 0.8*\sigma, x_o^* - 0.6*\sigma, x_o^* - 0.4*\sigma, x_o^* - 0.2*\sigma, x_o^*\}$. Results are reported for two S/N scenarios, namely $\tilde{F} = 20004$ e$^-$ (Left) and $\tilde{F} = 60160$ e$^-$ (Right), and across different pixel sizes.

First, we note that there is a significant difference in the predictions of our methodology for the ML estimator in comparison with what we predict in the WLS case. In fact, the results of our approach are very precise for the determination of the variance of the ML estimator in all the regimes, from small to high S/N, which is remarkable. More important it is the fact that, from these findings, we observe that the performance deviation from the MVB in the worse case (small S/N) is very small (see Table 2.1, first raw), while for any practical purposes the variance of the ML estimator achieves the CR limit for all the other regimes, from medium to high S/N, which is a numerical corroboration of the optimality of the ML estimator in astrometry, as predicted theoretically by Theorems 2.2 and 2.4. In Fig. 2.5 we also show the square root of the empirical variance ($Var(\hat{\tau_{ML}})$) with respect to the empirically-determined mean position $\hat{x}_c$ (using the ML estimator), all as derived from the simulations, showing good consistency with our predictions.

Complementing this analysis, we conducted the same comparison considering different positions for the object within the array obtaining the same trends and conclusions. To summarize these results, Fig. 2.6 shows the value $100 \times \frac{\sqrt{\sigma_{ML}^2(n) + \beta_{ML}(n)} - \sigma_{ML}(n)}{\sigma_{ML}(n)}$, which is an indicator of the quality of the estimator (the smaller the better) for different scenarios of the position of the object $x_c \in \Theta = \{x_o^* - \sigma, x_o^* - 0.8*\sigma, x_o^* - 0.6*\sigma, x_o^* - 0.4*\sigma, x_o^* - 0.2*\sigma, x_o^*\}$. In particular, for all the evaluated positions, the relative discrepancy is bounded (relative to the CR bound) by values in the range of 0.025% and 0.012% for pixel resolution in the range $\Delta x \in [0.1, 0.2]$ arcsec for $S/N = 120$ and $S/N = 230$, respectively. Finally, Table 2.1 presents the relative discrepancy for all the range of S/N values considered in this study for the case $\Delta x = 0.2$ arcsec.

We conclude from this analysis that the ML estimator in nearly optimal for the medium, high and very high S/N regimes across pixel resolutions, achieving the MVB for astrometry. This is an interesting result, since it lends further support to the adoption of these type

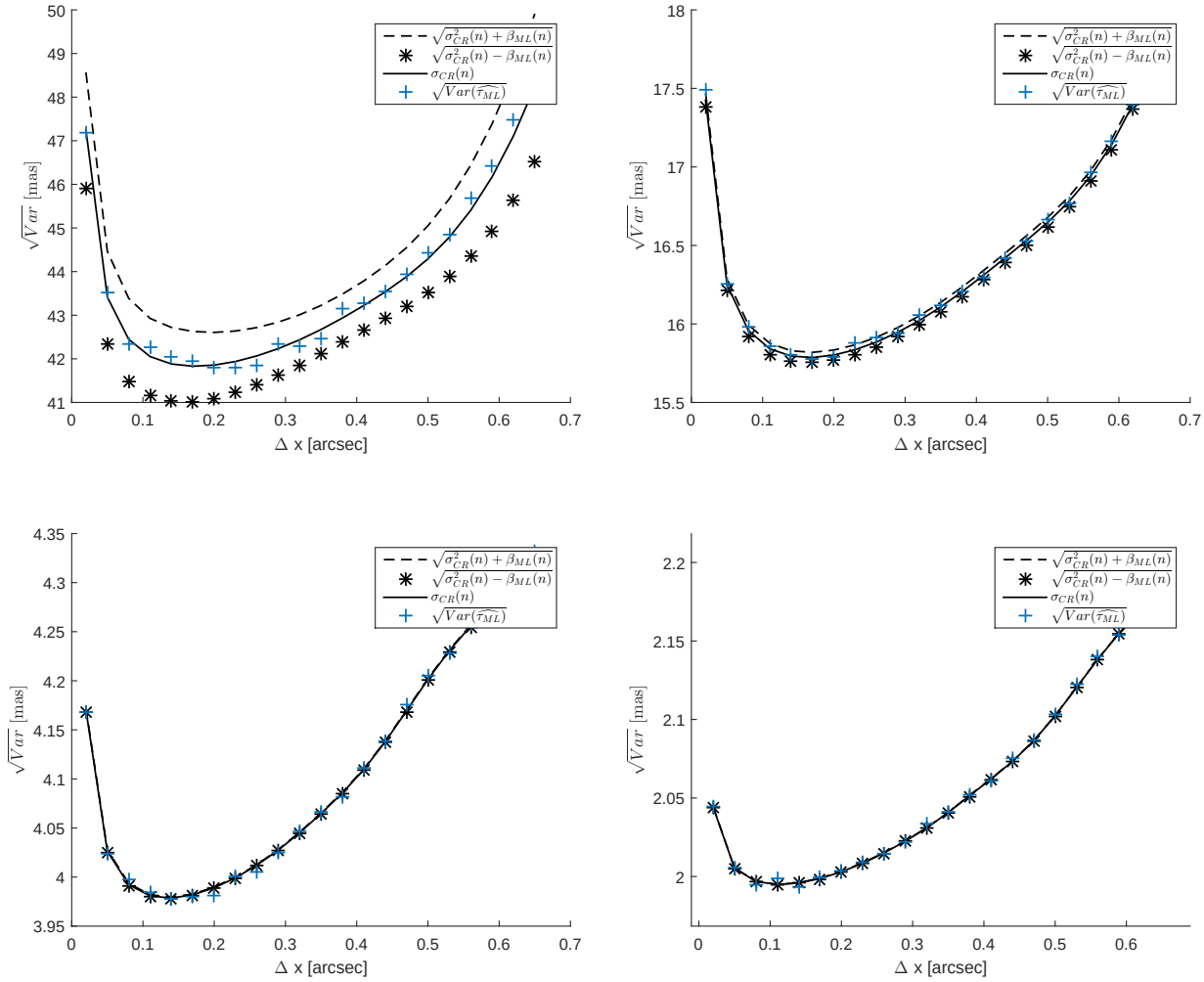

Figure 2.5: Range of the square root of the variance performance (in miliarcsecond=mas) for the ML method in astrometry as predicted by Theorems 2.2 and 2.4, Eq. (2.34). Results are reported for different values of $\tilde{F}$ and across different pixel sizes (top-left to bottom-right): $\tilde{F} \in \{1080, 3224, 20004, 60160\}$ e^-.

Position x_c	S/N=12	S/N=32	S/N=120	S/N=230
x_o	3.8%	0.34%	0.032%	0.010%
$x_o - 0.2 * \sigma$	4.1%	0.27%	0.014%	0.009%
$x_o - 0.4 * \sigma$	4.3%	0.19%	0.022%	0.007%
$x_o - 0.6 * \sigma$	3.8%	0.29%	0.019%	0.009%
$x_o - 0.8 * \sigma$	3.9%	0.30%	0.022%	0.011%
$x_o - \sigma$	3.6%	0.40%	0.019%	0.008%

Table 2.1: Indicators of the performance quality of the ML estimator relative to the Cramér-Rao bound expressed in terms of the indicator $100 \times \frac{\sqrt{\sigma^2_{ML}(n)+\beta_{ML}(n)}-\sigma_{ML}(n)}{\sigma_{ML}(n)}$ from the result in Theorem 2.4. The results are presented for different positions of the object $x_c \in \{x_o^* - \sigma, x_o^* - 0.8*\sigma, x_o^* - 0.6*\sigma, x_o^* - 0.4*\sigma, x_o^* - 0.2*\sigma, x_o^*\}$ (rows) and for different $S/N \in \{12, 32, 120, 230\}$ (columns).

of estimators for very demanding astrometric applications, as has been done in the case of Gaia [59]. We note that [13] reach the same conclusion regarding the optimality of the ML method in comparison with the MVB, through simulations of 2D CCD images using a broad set of Moffat PSF stellar profiles. While their results is purely empirical, it is interesting that they test the ML using a different PSF from ours, and in a 2D scenario, and yet they reach the same conclusions. More recently, [60] have tested (also empirically) the ML method in a 1D scenario (similar to ours), but in the context of a Gaia-like PSF. They find that the ML is unbiased (in agreement with our results, see Fig. 2.1), and, by comparing two implementations of the ML they conclude that they predict self-consistent and reliable results over a broad range of flux, background, and instrument response variations. It would still be quite interesting to compare the performance of those implementations against the CR MVB in order to further test our theoretical predictions.

2.4.5 Comments on an adaptive WLS estimator for astrometry

In Sect. 2.3.1 we presented results that offer a nominal prediction for the variance of the WLS method through Eq. (2.28) which turns out to be very accurate in the regime from medium to high S/N as shown in Sect. 2.4.3. Interestingly, Remark 2 tells us that this nominal values precisely match the CR limit for an optimal selection of weights given in Eq. (2.30), namely $w_{\rm i} \sim 1/\lambda_{\rm i}(x_c)$ for all $\rm i = 1, .., n$ (compare Eqs. (2.29) and (2.35)). As this selection is unfeasible, because it requires the knowledge of x_c (see the expression in Eq. (2.2)), we can approximate this value by a noisy version of it, considering the fact that the expected value of the observations $I_{\rm i}$ that we measure at pixel i is $\lambda_{\rm i}(x_c)$ using our model in Eq. (2.4). Therefore $I_{\rm i}$ can be interpreted as a noisy version of $\lambda_{\rm i}(x_c)$ and

$$\hat{w}_{\rm i}(I_{\rm i}) = \frac{1}{I_{\rm i}} \tag{2.39}$$

can be seen as a noisy version of the ideal weight $\frac{1}{\lambda_{\rm i}(x_c)}$. Adopting this data-driven weighting approach, we would have an adaptive WLS method as the weights are not fixed but instead they are a function of the data $\{I_1 : \rm i = 1, .., n\}$. This selection of weights can be interpreted

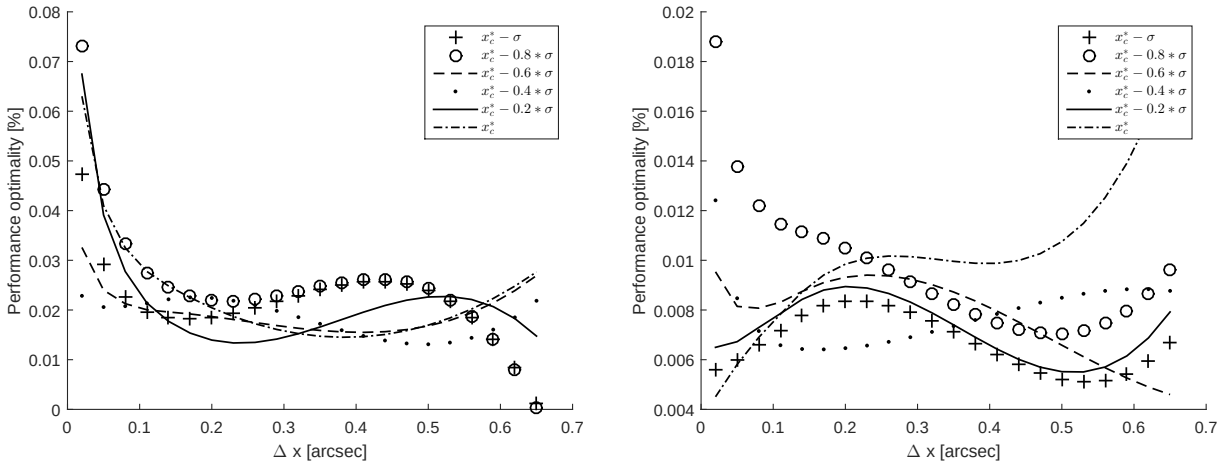

Figure 2.6: Indicator of the performance optimality of the ML indicator (computed as $100 \times \frac{\sqrt{\sigma_{ML}^2(n)+\beta_{ML}(n)}-\sigma_{ML}(n)}{\sigma_{ML}(n)}$) for different positions of the target object $x_c \in \Theta = \{x_o^* - \sigma, x_o^* - 0.8 * \sigma, x_o^* - 0.6 * \sigma, x_o^* - 0.4 * \sigma, x_o^* - 0.2 * \sigma, x_o^*\}$ in the array, as a function of pixel resolution. The left panel shows the case $\tilde{F} = 20004$ e$^-$, the right panel the case $\tilde{F} = 60160$ e$^-$.

as an empirical version of the optimal weights that achieves the CR bound. Then, the problem reduces to solve

$$\tau_{AWLS}(I^n) = \arg\min_{\alpha \in \mathbb{R}} J_{AWLS}(\alpha, I^n) \tag{2.40}$$

where

$$J_{AWLS}(\alpha, I^n) = \sum_{\mathrm{i}=1}^{n} \hat{w}_{\mathrm{i}}(I_{\mathrm{i}})(I_{\mathrm{i}} - \lambda_{\mathrm{i}}(\alpha))^2. \tag{2.41}$$

Fig. 2.7 presents the performance of this scheme for the same regimes we have been exploring in this work, supporting the conjecture that this selection of weights resembles the optimal weight selection and in fact achieves MSE performances that are surprisingly close to the CR bound in all the observational regimes.

Our results in this sub-section show that some commonly adopted weighting schemes, using the analogous to Eq. (2.39), are a very good data-driven choice for methods that employ a WLS scheme, such is the case, e.g., of the well know PSF stellar fitting program (including astrometry) DAOPHOT, described in [9, Eq. (10)].

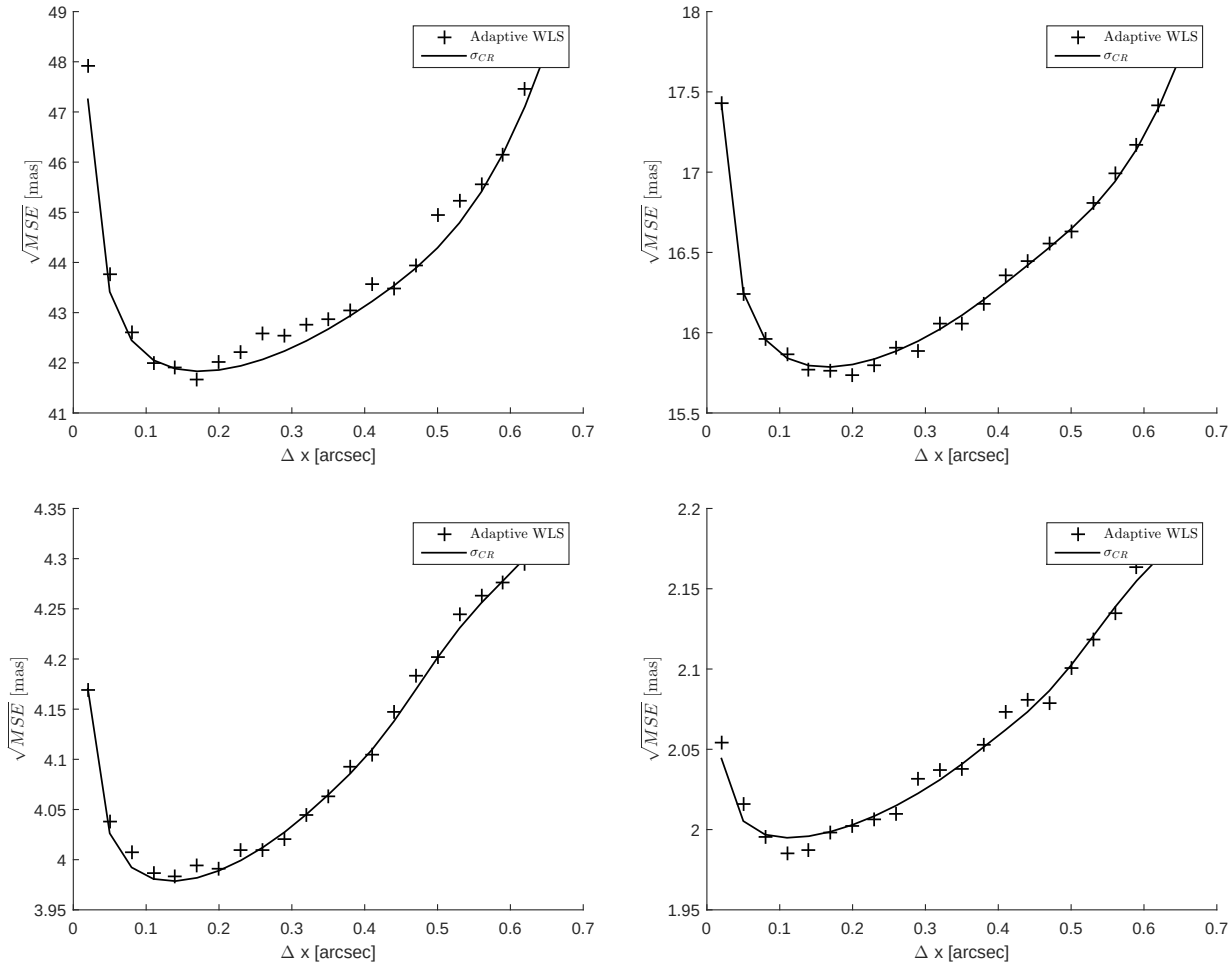

Figure 2.7: Performance comparison between the $\sqrt{MSE}$ of the adaptive WLS estimator and $\sigma_{CR}(n)$, both in mas. Results are reported for different $\tilde{F}$ and across different pixel sizes: (top-left to bottom-right) $\tilde{F} \in \{1080, 3224, 20004, 60160\}$ e^-.

2.5 Conclusions and Final Remarks

We study the performance of the WLS and ML estimators for relative astrometry on digital detectors subject to Poisson noise, in comparison with the best possible attainable precision given by the CR bound. Our study includes analytical results, and numerical simulations under realistic observational conditions to help us to corroborate our theoretical findings.

We extend the work presented in Fessler et al. [50] by giving upper and lower bounds for the variance and the mean of implicit estimators (as is the particular case of the WLS and ML schemes). We verified that the bias of the WLS and ML methods are negligible in all the observational regimes explored in this paper and that the variance of the ML method is close to the CR lower bound in most of the observational regimes explored in this work. We show the suboptimality of the WLS estimator by proving that this estimator can not achieve the CR bound unless you have access to the position (which is impossible in this estimation problem).

2.6 Appendix

2.6.1 Proof of Theorem 2.2

We begin presenting the expressions for $(\epsilon_J(n), \beta_J(n), \sigma_J^2(n))$ to complete the statement of the result.

$$\epsilon_J(n) = \max_{t\in[0,1]} \left| \mathbb{E}_{I^n\sim f_{x_c}} \left\{ \frac{1}{2} \sum_{\mathrm{i}=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial I_{\mathrm{i}} \partial I_j} \tau_J(\bar{I}^n - t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j) \right\} \right|, \tag{2.42}$$

$$\beta_J(n) = \epsilon'_J(n) + 2\delta'_J(n), \tag{2.43}$$

where

$$\epsilon'_J(n) = \max_{t\in[0,1]} \mathbb{E}_{I^n\sim f_{x_c}} \left\{ \left(\frac{1}{2} \sum_{\mathrm{i}=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial I_{\mathrm{i}} \partial I_j} \tau_J(\bar{I}^n - t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j) \right)^2 \right\}, \tag{2.44}$$

$$\delta'_J(n) = \max_{t\in[0,1]} \left| \mathbb{E}_{I^n\sim f_{x_c}} \left\{ (\nabla \tau_J(\bar{I}^n) \cdot (I^n - \bar{I}^n)) \cdot \frac{1}{2} \sum_{\mathrm{i}=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial I_{\mathrm{i}} \partial I_j} \tau_J(\bar{I}^n - t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j) \right\} \right|, \tag{2.45}$$

and, finally,

$$\sigma_J^2(n) = \left[\frac{\partial^2 J(\tau_J(\bar{I}^n), \bar{I}^n)}{\partial \alpha^2} \right]^{-1} \left[\frac{\partial^2 J(\tau_J(\bar{I}^n), \bar{I}^n)}{\partial \alpha \partial I_{\mathrm{i}}} \right] Cov\{I^n\} \left[\frac{\partial^2 J(\tau_J(\bar{I}^n), \bar{I}^n)}{\partial \alpha \partial I_{\mathrm{i}}} \right]^T \left[\frac{\partial^2 J(\tau_J(\bar{I}^n), \bar{I}^n)}{\partial \alpha^2} \right]^{-1}. \tag{2.46}$$

Proof of Theorem 2.2: Using the chain rule in the cost function $J(\alpha, I^n)$ and taking the partial derivative $\frac{\partial}{\partial I_{\mathrm{i}}}$ of both sides in (2.15), we have that

$$0 = \frac{\partial^2}{\partial \alpha^2} J(\tau(I^n), I^n) \frac{\partial}{\partial I_{\mathrm{i}}} \tau(I^n) + \frac{\partial^2}{\partial \alpha \partial I_{\mathrm{i}}} J(\tau(I^n), I^n), \ \mathrm{i} = 1, \ldots, n. \tag{2.47}$$

Thus, we have n equations with one unknown, and it holds for any I^n. Defining the operators

$$\nabla^{20}(\cdot) = \frac{\partial^2}{\partial \alpha^2}, \nabla^{11}(\cdot) = \frac{\partial^2}{\partial \alpha \partial I_{\mathrm{i}}}, \tag{2.48}$$

of dimensions 1×1 and $1 \times n$, respectively, we can express (2.47) in matrix form as

$$0 = \nabla^{20} J(\tau(I^n), I^n) \nabla \tau(I^n) + \nabla^{11} J(\tau(I^n), I^n). \tag{2.49}$$

Assuming that the matrix $\nabla^{20} J(\tau(I^n), I^n)$ is non singular, we can calculate $\nabla \tau(I^n)$ from (2.49)

$$\nabla \tau(I^n) = -[\nabla^{20} J(\tau(I^n), I^n)]^{-1} \nabla^{11} J(\tau(I^n), I^n). \tag{2.50}$$

Finally, using (2.50), evaluating at $\bar{I}^n$, and then replacing in (2.18), we have that

$$\begin{aligned}
\sigma_J^2(n) &= -[\nabla^{20} J(\tau(\bar{I}^n), \bar{I}^n)]^{-1} \nabla^{11} J(\tau(\bar{I}^n), \bar{I}^n) Cov\{I^n\} (-[\nabla^{20} J(\tau(\bar{I}^n), \bar{I}^n)]^{-1} \nabla^{11} J(\tau(\bar{I}^n, \bar{I}^n))^T \\
&= [\nabla^{20} J(\tau(\bar{I}^n), \bar{I}^n)]^{-1} \nabla^{11} J(\tau(\bar{I}^n), \bar{I}^n) Cov\{I^n\} [\nabla^{11} J(\tau(\bar{I}^n), \bar{I}^n))]^T [\nabla^{20} J(\tau(\bar{I}^n), \bar{I}^n)]^{-1}. \\
&= \left[\frac{\partial^2 J(\tau(\bar{I}^n), \bar{I}^n)}{\partial \alpha^2}\right]^{-1} \left[\frac{\partial^2 J(\tau(\bar{I}^n), \bar{I}^n)}{\partial \alpha \partial I_{\mathrm{i}}}\right] Cov\{I^n\} \left[\frac{\partial^2 J(\tau(\bar{I}^n), \bar{I}^n)}{\partial \alpha \partial I_{\mathrm{i}}}\right]^T \left[\frac{\partial^2 J(\tau(\bar{I}^n), \bar{I}^n)}{\partial \alpha^2}\right]^{-1}
\end{aligned} \tag{2.51}$$

Moving into the residual term $\gamma_J(n)$ in (2.18) captured by $\beta_J(n)$, we must consider the variance of the error function $Var\{\mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\}$ and the covariance $Cov\{\nabla \tau_J(\bar{I}^n)(I^n - \bar{I}^n), \mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\}$. For the first, we have that

$$\begin{aligned}
Var\{\mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\} &= Var\left\{\frac{1}{2} \sum_{\mathrm{i}=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \tau}{\partial I_{\mathrm{i}} \partial I_j} (\bar{I}^n + t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j)\right\} \\
&\leq \mathbb{E}\left\{\left(\frac{1}{2} \sum_{\mathrm{i}=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \tau}{\partial I_{\mathrm{i}} \partial I_j} (\bar{I}^n + t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j)\right)^2\right\} \\
&\leq \underbrace{\max_{t \in [0,1]} \mathbb{E}\left\{\left(\frac{1}{2} \sum_{\mathrm{i}=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \tau}{\partial I_{\mathrm{i}} \partial I_j} (\bar{I}^n + t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j)\right)^2\right\}}_{=\epsilon'_J(n)}
\end{aligned} \tag{2.52}$$

On the other hand, for the covariance, using the main assumption in (2.20), it is clear that

$$\mathbb{E}_{I^n \sim f_{x_c}} \left\{\nabla \tau_{ML}(\bar{I}^n) \cdot (I^n - \bar{I}^n)\right\} = \mathbb{E}_{I^n \sim f_{x_c}} \left\{a \sum_{\mathrm{i}=1}^{n} b_{\mathrm{i}} (I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})\right\} = 0. \tag{2.53}$$

From this,

$$\begin{aligned}
& |Cov\{\nabla \tau(\bar{I}^n)(I^n - \bar{I}^n), \mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\}| \\
= \ & |\mathbb{E}\left\{\nabla \tau(\bar{I}^n)(I^n - \bar{I}^n)\left(\mathrm{e}(\bar{I}^n, I^n - \bar{I}^n) - \mathbb{E}\left(\mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\right)\right)\right\}| \\
= \ & \left|\mathbb{E}_{I^n \sim f_{x_c}} \left\{(\nabla \tau(\bar{I}^n) \cdot (I^n - \bar{I}^n)) \cdot \frac{1}{2} \sum_{\mathrm{i}=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial I_{\mathrm{i}} \partial I_j} \tau(\bar{I}^n - t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j)\right\}\right| \\
\leq \ & \underbrace{\max_{t \in [0,1]} \left|\mathbb{E}_{I^n \sim f_{x_c}} \left\{(\nabla \tau(\bar{I}^n) \cdot (I^n - \bar{I}^n)) \cdot \frac{1}{2} \sum_{\mathrm{i}=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial I_{\mathrm{i}} \partial I_j} \tau(\bar{I}^n - t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j)\right\}\right|}_{=\delta'_J(n)}
\end{aligned} \tag{2.54}$$

Finally, replacing (2.52) and (2.54) in the definition of $\gamma_J(n)$, we have that:

$$\begin{aligned} |\gamma_J(n)| &\leq Var\{\mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\} + 2|Cov\{\nabla\tau(\bar{I}^n)(I^n - \bar{I}^n), \mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\}| \\ &\leq \epsilon'_J(n) + 2\delta'_J(n) = \beta_J(n). \end{aligned} \tag{2.55}$$

For the bias expression of the result in (2.22), using the hypothesis in (2.21), we can take expectation at both sides of (2.17) to obtain that

$$\begin{aligned} |\mathbb{E}_{I^n\sim f_{x_c}}\{\tau(I^n)\} - x_c| &= \left|\mathbb{E}_{I^n\sim f_{x_c}}\left\{a\sum_{\mathrm{i}=1}^{N} b_{\mathrm{i}}(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}}) + \mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\right\}\right| \\ &= \left|\mathbb{E}_{I^n\sim f_{x_c}}\left\{\mathrm{e}(\bar{I}^n, I^n - \bar{I}^n)\right\}\right| \\ &= \left|\mathbb{E}_{I^n\sim f_{x_c}}\left\{\frac{1}{2}\sum_{\mathrm{i}=1}^{n}\sum_{j=1}^{n}\frac{\partial^2}{\partial I_{\mathrm{i}}\partial I_j}\tau(\bar{I}^n - t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j)\right\}\right| \\ &\leq \underbrace{\max_{t\in[0,1]}\left|\mathbb{E}_{I^n\sim f_{x_c}}\left\{\frac{1}{2}\sum_{\mathrm{i}=1}^{n}\sum_{j=1}^{n}\frac{\partial^2}{\partial I_{\mathrm{i}}\partial I_j}\tau(\bar{I}^n - t(I^n - \bar{I}^n))(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})(I_j - \bar{I}_j)\right\}\right|}_{=\epsilon_J(n)} \end{aligned} \tag{2.56}$$

2.6.2 Proof of Theorem 2.3

Proof: The proof and, in particular, the derivation of $\sigma^2_{WLS}(n)$, $\beta_{WLS}(n)$ and $\epsilon_{WLS}(n)$ simply reduces to an straightforward application of Theorem 2.2. For that we need to first validate the assumptions of Theorem 2.2. If we begin with Eq. (2.50)

$$\nabla\tau(I^n) = -[\nabla^{20}J(\tau_J(I^n), I^n)]^{-1}\nabla^{11}J(\tau_J(I^n), I^n), \tag{2.57}$$

and then we calculate the gradient terms in the RHS of (2.57) for our WLS context, it follows that

$$\begin{aligned} \nabla^{20}J_{WLS}(\alpha, I^n) &= \frac{\partial^2}{\partial\alpha^2}J_{WLS}(\alpha, I^n) \\ &= 2\sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}}\left(\left(\frac{\partial\lambda_{\mathrm{i}}(\alpha)}{\partial\alpha}\right)^2 + (\lambda_{\mathrm{i}}(\alpha) - I_{\mathrm{i}})\frac{\partial^2\lambda_{\mathrm{i}}(\alpha)}{\partial\alpha^2}\right), \end{aligned} \tag{2.58}$$

$$\nabla^{11}J_{WLS}(\alpha, I^n) = \left(\frac{\partial^2}{\partial\alpha\partial I_1}J_{WLS}(\alpha, I^n), \ldots, \frac{\partial^2}{\partial\alpha\partial I_n}J_{WLS}(\alpha, I^n)\right)^T \tag{2.59}$$

$$= -2\left(w_1\frac{\partial\lambda_1(\alpha)}{\partial\alpha}, \ldots, w_n\frac{\partial\lambda_n(\alpha)}{\partial x_c}\right)^T. \tag{2.60}$$

Following (2.57), we need to evaluate Eqs. (2.58) and (2.60) at $\alpha = \tau_{WLS}(\bar{I}^n)$. For that, we have the following[9]

$$\tau_{WLS}(\bar{I}^n) = \underset{\alpha\in\mathbb{R}}{\operatorname{argmin}}\sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}}(\lambda_{\mathrm{i}}(x_c) - \lambda_{\mathrm{i}}(\alpha))^2. \tag{2.61}$$

Then we will use the following result:

[9] considering that $\bar{I}_{\mathrm{i}} = E\{I_{\mathrm{i}}\} = \lambda_{\mathrm{i}}(x_c)$.

Proposition 2.5 *Under the assumption of a Gaussian PSF, $\tau_{WLS}(\bar{I}^n) = x_c$.*

Notice that this proposition is the second assumption used in Theorem 2.2. Using this proposition, we obtain that

$$\begin{aligned}
\nabla^{20} J_{WLS}(\tau_{WLS}(\bar{I}^n), \bar{I}^n) &= 2\sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}} \left(\left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \right)^2 \Bigg|_{\alpha=x_c} + (\lambda_{\mathrm{i}}(x_c) - \lambda_{\mathrm{i}}(x_c)) \left. \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \right|_{\alpha=x_c} \right), \\
&= 2\sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \right)^2 \Bigg|_{\alpha=x_c}, \qquad (2.62) \\
\nabla^{11} J_{WLS}(\tau_{WLS}(\bar{I}^n), \bar{I}^n) &= -2 \left(w_1 \frac{\partial \lambda_1(\alpha)}{\partial \alpha}, \ldots, w_n \frac{\partial \lambda_n(\alpha)}{\partial \alpha} \right)^T \Bigg|_{\alpha=\tau_{WLS}(\bar{I}^n)}, \\
&= -2 \left(w_1 \frac{\partial \lambda_1(\alpha)}{\partial \alpha}, \ldots, w_n \frac{\partial \lambda_n(\alpha)}{\partial \alpha} \right)^T \Bigg|_{\alpha=x_c}. \qquad (2.63)
\end{aligned}$$

Finally, applying (2.62) and (2.63) in (2.50) we have that

$$\begin{aligned}
\nabla \tau_{WLS}(\bar{I}^n) \cdot (I^n - \bar{I}^n) &= -[\nabla^{20} J(\tau_{WLS}(\bar{I}^n), \bar{I}^n)]^{-1} [\nabla^{11} J(\tau_{WLS}(\bar{I}^n), \bar{I}^n)](I^n - \bar{I}^n) \\
&= \underbrace{\left[\sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \right)^2 \Bigg|_{\alpha=x_c} \right]^{-1}}_{a} \cdot \sum_{j=1}^{n} \underbrace{w_j \left. \frac{\partial \lambda_j(\alpha)}{\partial \alpha} \right|_{\alpha=x_c}}_{b_j} (I_j - \mathbb{E}(I_j)), \qquad (2.64)
\end{aligned}$$

which offers the decomposition needed for the application of Theorem 2.2 (Eq. (2.20)). For the value of $\sigma^2_{WLS}(n)$ in (2.46), since the observations are independent and follow a Poisson distribution, we have that

$$Cov\{I_{\mathrm{i}}, I_j\} = \begin{cases} Var\{I_{\mathrm{i}}\} = \lambda_{\mathrm{i}}(x_c), & \text{if } \mathrm{i} = j, \\ 0 & \sim . \end{cases} \qquad (2.65)$$

Then if we replace (2.62), (2.63) and (2.65) in (2.46), we have that

$$\sigma^2_{WLS}(n) = \frac{\sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}}^2 \lambda_{\mathrm{i}}(x_c) \left. \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \right)^2 \right|_{\alpha=x_c}}{\left(\sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}} \left. \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \right)^2 \right|_{\alpha=x_c} \right)^2}. \qquad (2.66)$$

On the other hand, the expression for $\beta_{WLS}(n)$ and $\epsilon_{WLS}(n)$ can be determined from the evaluation of (2.43) and (2.42), respectively. Looking at them, the problem reduces to determine the key term $\frac{\partial^2 \tau_{WLS}}{\partial I_{\mathrm{i}} \partial I_j}(\bar{I}^n + t(I^n - \bar{I}^n))$. For that, if we use [50, Eq. (17)] we can obtain

the following identity[10]

$$\frac{\partial^2 \tau_{WLS}}{\partial I_{\mathrm{i}} \partial I_j}(\bar{I}^n + t(I^n - \bar{I}^n)) = \frac{-1}{\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot (\lambda_{\mathrm{i}}(\alpha) - (\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})))2w_{\mathrm{i}} + 2w_{\mathrm{i}} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right]^2} \cdot \left[\left[\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^3 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^3} \cdot (\lambda_{\mathrm{i}}(\alpha) - (\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})))2w_{\mathrm{i}} + 6w_{\mathrm{i}} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right] \cdot \frac{\left(2w_j \frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)}{\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot (\lambda_{\mathrm{i}}(\alpha) - (\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})))2w_{\mathrm{i}} - 2w_{\mathrm{i}} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right]} - \left(2w_j \frac{\partial^2 \lambda_j(\alpha)}{\partial \alpha^2}\right)\right] \cdot \left(2w_{\mathrm{i}} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right) - \left(2w_{\mathrm{i}} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2}\right) \cdot \left(2w_j \frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)\right]\Bigg|_{\alpha = \tau_{WLS}(\bar{I}^n + t(I^n - \bar{I}^n))}, \tag{2.67}$$

which concludes the result.

Proof of Proposition 2.5

Proof: Using the function $h(\alpha) = \sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}}(\lambda_{\mathrm{i}}(x_c) - \lambda_{\mathrm{i}}(\alpha))^2$, we need to show that the minimum is reached only at $\alpha = x_c$. From this, we have that $h(\alpha) \geq 0$ and it achieves its minimum at x_c. To prove uniqueness, let us assume that there is another position $x_c^* \neq x_c$ at which h is zero. Then

$$\begin{aligned} h(x_c^*) &= \sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}}(\lambda_{\mathrm{i}}(x_c) - \lambda_{\mathrm{i}}(x_c^*))^2 = 0 \\ &\Leftrightarrow \lambda_{\mathrm{i}}(x_c) = \lambda_{\mathrm{i}}(x_c^*), \quad \forall \mathrm{i} \in \{1, \ldots, n\}. \end{aligned} \tag{2.68}$$

The last identity is not possible, because if we use a Gaussian PSF there is at least one $\mathrm{i} \in \{1, \ldots, n\}$ such that $\lambda_{\mathrm{i}}(x_c) \neq \lambda_{\mathrm{i}}(x_c^*)$.

Proof of Eq. (2.67)

Proof: Recall [50, Eq. (17)] and considering $J_{WLS}(\alpha, I^n)$ as the cost function we have that

$$\begin{aligned} \frac{\partial^2 \tau_{WLS}}{\partial I_{\mathrm{i}} \partial I_j}(\bar{I}^n + t(I^n - \bar{I}^n)) = &\left[-\frac{\partial^2 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2}\right]^{-1} \Bigg(\Bigg[\frac{\partial^3 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^3} \cdot \frac{\partial \tau_{WLS}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_j} + \\ &\frac{\partial^3 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2 \partial I_j}\Bigg] \cdot \frac{\partial \tau_{WLS}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_{\mathrm{i}}} \\ &+ \frac{\partial^3 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2 \partial I_{\mathrm{i}}} \cdot \frac{\partial \tau_{WLS}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_j} + \frac{\partial^3 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha \partial I_{\mathrm{i}} \partial I_j}\Bigg)\Bigg|_{\alpha = \tau_{WLS}(\bar{I}^n + t(I^n - \bar{I}^n))}, \end{aligned} \tag{2.69}$$

where from the definition of $J_{ML}(\alpha, I^n)$ we have that

$$\frac{\partial^2 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2} = \sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot (\lambda_{\mathrm{i}}(\alpha) - (\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})))2w_{\mathrm{i}} + 2w_{\mathrm{i}} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2, \tag{2.70}$$

[10] The derivation of this identity is presented in Appendix 2.6.2.

$$\frac{\partial^3 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^3} = \sum_{\mathrm{i}=1}^{n} \frac{\partial^3 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^3} \cdot (\lambda_{\mathrm{i}}(\alpha) - (\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})))2w_{\mathrm{i}} + 6w_{\mathrm{i}} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}, \tag{2.71}$$

$$\frac{\partial^3 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2 \partial I_{\mathrm{i}}} = -\left(2w_{\mathrm{i}} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2}\right), \tag{2.72}$$

$$\frac{\partial^3 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha \partial I_{\mathrm{i}} \partial I_j} = 0. \tag{2.73}$$

Concerning $\frac{\partial \tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_{\mathrm{i}}}$ it is just the i-th component of the gradient in Eq. (2.57), then we use (2.58) and (2.60)

$$\begin{aligned}\frac{\partial \tau_{WLS}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_{\mathrm{i}}} &= \frac{-\frac{\partial^2 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n)))}{\partial \alpha \partial I_{\mathrm{i}}}}{\frac{\partial^2 J_{WLS}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2}} \\ &= \frac{\left(2w_{\mathrm{i}} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)}{\sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot (\lambda_{\mathrm{i}}(\alpha) - (\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})))2w_{\mathrm{i}} + 2w_{\mathrm{i}} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2}.\end{aligned} \tag{2.74}$$

Finally, replacing (2.70), (2.71), (2.72), (2.73) and (2.74) in (2.69), and evaluating in $\alpha = \tau_{WLS}(\bar{I}^n + t(I^n - \bar{I}^n))$ we obtain the desired result

$$\begin{aligned}\frac{\partial^2 \tau_{WLS}}{\partial I_{\mathrm{i}} \partial I_j}(\bar{I}^n + t(I^n - \bar{I}^n)) = &\frac{-1}{\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot (\lambda_{\mathrm{i}}(\alpha) - (\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})))2w_{\mathrm{i}} + 2w_{\mathrm{i}} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right]^2}. \\ &\left[\left[\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^3 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^3} \cdot (\lambda_{\mathrm{i}}(\alpha) - (\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})))2w_{\mathrm{i}} + 6w_{\mathrm{i}} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right] \cdot \right. \\ &\frac{\left(2w_j \frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)}{\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot (\lambda_{\mathrm{i}}(\alpha) - (\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})))2w_{\mathrm{i}} - 2w_{\mathrm{i}} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right]} - \left(2w_j \frac{\partial^2 \lambda_j(\alpha)}{\partial \alpha^2}\right)\Bigg] \cdot \\ &\left.\left(2w_{\mathrm{i}} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right) - \left(2w_{\mathrm{i}} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2}\right) \cdot \left(2w_j \frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)\right]\Bigg|_{\alpha = \tau_{WLS}(\bar{I}^n + t(I^n - \bar{I}^n))}.\end{aligned} \tag{2.75}$$

2.6.3 Proof of Theorem 2.4

Proof: Again the proof and the derivation of $\sigma^2_{ML}(n)$, $\beta_{ML}(n)$ and $\epsilon_{ML}(n)$ reduce to apply Theorem 2.2. First, we need to validate the assumption of Theorem 2.2. Beginning with the

equality in Eq. (2.57), it follows that

$$\begin{aligned}
\nabla^{20} J_{ML}(\alpha, I^n) &= \frac{\partial^2}{\partial \alpha^2} J_{ML}(\alpha, I^n) \\
&= -\sum_{\mathrm{i}=1}^{n} I_{\mathrm{i}} \frac{1}{\lambda_{\mathrm{i}}^2(\alpha)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \right)^2 + \sum_{\mathrm{i}=1}^{n} \left(I_{\mathrm{i}} \frac{1}{\lambda_{\mathrm{i}}(\alpha)} - 1 \right) \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2}, \quad (2.76) \\
\nabla^{11} J_{ML}(\alpha, I^n) &= \left(\frac{\partial^2}{\partial \alpha \partial I_1} J_{ML}(\alpha, I^n), \ldots, \frac{\partial^2}{\partial \alpha \partial I_n} J_{ML}(\alpha, I^n) \right)^T \\
&= \left(\frac{1}{\lambda_1(\alpha)} \frac{\partial \lambda_1(\alpha)}{\partial \alpha}, \ldots, \frac{1}{\lambda_n(\alpha)} \frac{\partial \lambda_n(\alpha)}{\partial \alpha} \right)^T. \quad (2.77)
\end{aligned}$$

For evaluating these two expression at $\alpha = \tau_{ML}(\bar{I}^n)$ as required in (2.57), we use that[11]

$$\tau_{ML}(\bar{I}^n) = \operatorname*{argmin}_{\alpha \in \mathbb{R}} \sum_{\mathrm{i}=1}^{n} -\lambda_{\mathrm{i}}(x_c) \ln(\lambda_{\mathrm{i}}(\alpha)) + \lambda_{\mathrm{i}}(\alpha). \quad (2.78)$$

Then we will use the following result

Proposition 2.6 *Under the assumption of a Gaussian PSF, $\tau_{ML}(\bar{I}^n) = x_c$.*

Notice again, that this proposition is the second assumption used in Theorem 2.2. From this proposition, it follows that

$$\begin{aligned}
\nabla^{20} J(\tau_{ML}(\bar{I}^n), \bar{I}^n) &= -\sum_{\mathrm{i}=1}^{n} \lambda_{\mathrm{i}}(x_c) \frac{1}{\lambda_{\mathrm{i}}^2(x_c)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \right)^2 \Bigg|_{\alpha = x_c} + \sum_{\mathrm{i}=1}^{n} \left(\frac{\lambda_{\mathrm{i}}(x_c)}{\lambda_{\mathrm{i}}(x_c)} - 1 \right) \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \Bigg|_{\alpha = x_c}, \\
&= -\sum_{\mathrm{i}=1}^{n} \frac{1}{\lambda_{\mathrm{i}}(x_c)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \right)^2 \Bigg|_{\alpha = x_c}, \quad (2.79) \\
\nabla^{11} J(\tau_{ML}(\bar{I}^n), \bar{I}^n) &= \left(\frac{\partial^2}{\partial \alpha \partial I_1} J(\alpha, I^n), \ldots, \frac{\partial^2}{\partial \alpha \partial I_n} J(\alpha, I) \right)^T \Bigg|_{\alpha = \tau_{ML}(\bar{I}^n)} \\
&= \left(\frac{1}{\lambda_1(\alpha)} \frac{\partial \lambda_1(\alpha)}{\partial \alpha}, \ldots, \frac{1}{\lambda_n(\alpha)} \frac{\partial \lambda_n(\alpha)}{\partial \alpha} \right)^T \Bigg|_{\alpha = x_c}. \quad (2.80)
\end{aligned}$$

Finally, we apply (2.79) and (2.80) in (2.50) to obtain that

$$\begin{aligned}
\nabla \tau_{ML}(\bar{I}^n) \cdot (I^n - \bar{I}^n) &= -[\nabla^{20} J(\tau_{ML}(\bar{I}^n), \bar{I}^n)]^{-1} [\nabla^{11} J(\tau_{ML}(\bar{I}^n), \bar{I}^n)](I^n - \bar{I}^n) \\
&= \underbrace{-\left[\sum_{\mathrm{i}=1}^{n} \frac{1}{\lambda_{\mathrm{i}}(x_c)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \right)^2 \Bigg|_{\alpha = x_c} \right]^{-1}}_{a} \cdot \sum_{j=1}^{n} \underbrace{\frac{1}{\lambda_j(x_c)} \frac{\partial \lambda_j(\alpha)}{\partial \alpha} \Bigg|_{\alpha = x_c}}_{b_j} (I_j - \mathbb{E}(I_j)) \quad (2.81)
\end{aligned}$$

that shows that the sufficient condition in (2.20) of Theorem 2.2 is satisfied. For computing the value $\sigma_{ML}^2(n)$ in (2.46), we have that

$$Cov\{I_{\mathrm{i}}, I_j\} = \begin{cases} Var\{I_{\mathrm{i}}\} = \lambda_{\mathrm{i}}(x_c), & \text{if } \mathrm{i} = j, \\ 0 & \sim . \end{cases} \quad (2.82)$$

[11] Considering that $\bar{I}_{\mathrm{i}} = E\{I_{\mathrm{i}}\} = \lambda_{\mathrm{i}}(x_c)$.

since the observations are independent and follow a Poisson distribution. Then, replacing (2.79), (2.80) and (2.82) in (2.46) we have that

$$\sigma^2_{ML}(n) \quad = \quad \frac{1}{\sum_{\mathrm{i}=1}^{n} \frac{1}{\lambda_{\mathrm{i}}(x_c)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2 \Big|_{\alpha=x_c}}, \tag{2.83}$$

which resolves the identity in (2.35). Finally $\beta_{ML}(n)$ and $\epsilon_{ML}(n)$ comes from evaluating (2.43) and (2.42) in this ML context. For that we only need to determine $\frac{\partial^2 \tau_{ML}}{\partial I_{\mathrm{i}} \partial I_j}(\bar{I}^n + t(I^n - \bar{I}^n))$. Using [50, Eq. (17)], we can obtain the following identity[12]

$$\begin{aligned}
&\frac{\partial^2 \tau_{ML}}{\partial I_{\mathrm{i}} \partial I_j}(\bar{I}^n + t(I^n - \bar{I}^n)) = \frac{-1}{\left[\sum\limits_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}(\alpha)} - \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^2(\alpha)} \cdot \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right]^2} \cdot \\
&\left[\left[\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^3 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^3} \cdot \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}(\alpha)} - 3 \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^2(\alpha)} + 2 \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^3(\alpha)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^3\right] \cdot \right. \\
&\frac{\left(-\frac{1}{\lambda_j(\alpha)} \frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)}{\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}(\alpha)} - \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^2(\alpha)} \cdot \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right]} + \frac{\partial^2 \lambda_j(\alpha)}{\partial \alpha^2} \frac{1}{\lambda_j(\alpha)} - \frac{1}{\lambda_j(\alpha)} \left(\frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)^2\Bigg] \cdot \\
&\left.\left(-\frac{1}{\lambda_{\mathrm{i}}(\alpha)} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right) + \left(\frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \frac{1}{\lambda_{\mathrm{i}}(\alpha)} - \frac{1}{\lambda_{\mathrm{i}}(\alpha)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right) \cdot \left(-\frac{1}{\lambda_j(\alpha)} \frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)\right]\Bigg|_{\alpha=\tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))},
\end{aligned} \tag{2.84}$$

which concludes the result.

Proof of Proposition 2.6

Proof: Let us consider the function $g_n : \mathbb{R}_+^n \to \mathbb{R}$ given by

$$g_n(y_1, ..., y_n)_{\lambda_{\mathrm{i}}^n} = \sum_{\mathrm{i}=1}^{n} -\lambda_{\mathrm{i}} \ln(y_{\mathrm{i}}) + y_{\mathrm{i}}. \tag{2.85}$$

We note that

$$\min_{y_1^n \in \mathbb{R}_+^n} g_n(y_1, ..., y_n)_{\lambda_{\mathrm{i}}^n} = \sum_{\mathrm{i}=1}^{n} \min_{y_{\mathrm{i}} \in \mathbb{R}_+} g_1(y_{\mathrm{i}})_{\lambda_{\mathrm{i}}}, \tag{2.86}$$

where applying first order condition $y_{\mathrm{i}} = \lambda_{\mathrm{i}}$, $\forall \mathrm{i} \in \{1, ..., n\}$. Returning to our problem in (2.78) where $\lambda_{\mathrm{i}} = \bar{I}_{\mathrm{i}} = \lambda_{\mathrm{i}}(x_c)$ and $y_{\mathrm{i}} = \lambda_{\mathrm{i}}(\alpha)$, it is clear, considering the Gaussian profile in PSF, that

$$\lambda_{\mathrm{i}}(\alpha) = \lambda_{\mathrm{i}}(x_c) \quad \forall \mathrm{i} \in \{1, ..., n\} \text{ if } \alpha = x_c, \tag{2.87}$$

which concludes the result.

[12] The derivation of this result is presented in Appendix 2.6.3.

Proof of Eq. (2.84)

Proof: Recall [50, Eq. (17)] and considering $J_{ML}(\alpha, I^n)$ as the cost function we have that:

$$\frac{\partial^2 \tau_{ML}}{\partial I_{\mathrm{i}} \partial I_j}(\bar{I}^n + t(I^n - \bar{I}^n)) = \left[-\frac{\partial^2 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2}\right]^{-1} \left(\left[\frac{\partial^3 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^3} \cdot \frac{\partial \tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_j} + \frac{\partial^3 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2 \partial I_j}\right] \cdot \frac{\partial \tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_{\mathrm{i}}} + \frac{\partial^3 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2 \partial I_{\mathrm{i}}} \cdot \frac{\partial \tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_j} + \frac{\partial^3 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha \partial I_{\mathrm{i}} \partial I_j}\right)\Bigg|_{\alpha = \tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))}, \tag{2.88}$$

where from the definition of $J_{ML}(\alpha, I^n)$ we have that

$$\frac{\partial^2 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2} = \sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}(\alpha)} - \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^2(\alpha)} \cdot \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2, \tag{2.89}$$

$$\frac{\partial^3 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^3} = \sum_{\mathrm{i}=1}^{n} \frac{\partial^3 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^3} \cdot \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}(\alpha)} - 3\frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^2(\alpha)} + 2\frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^3(\alpha)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^3, \tag{2.90}$$

$$\frac{\partial^3 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2 \partial I_{\mathrm{i}}} = \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot \frac{1}{\lambda_{\mathrm{i}}(\alpha)} - \frac{1}{\lambda_{\mathrm{i}}^2(\alpha)} \cdot \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2, \tag{2.91}$$

$$\frac{\partial^3 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha \partial I_{\mathrm{i}} \partial I_j} = 0. \tag{2.92}$$

Concerning $\frac{\partial \tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_{\mathrm{i}}}$ it is just the i-th component of the gradient in Eq. (2.57), then we use (2.76) and (2.77)

$$\begin{aligned} \frac{\partial \tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))}{\partial I_{\mathrm{i}}} &= \frac{-\frac{\partial^2 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n)))}{\partial \alpha \partial I_{\mathrm{i}}}}{\frac{\partial^2 J_{ML}(\alpha, \bar{I}^n + t(I^n - \bar{I}^n))}{\partial \alpha^2}} \\ &= \frac{\left(-\frac{1}{\lambda_{\mathrm{i}}(\alpha)} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)}{\sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}(\alpha)} - \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^2(\alpha)} \cdot \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2}. \end{aligned} \tag{2.93}$$

Finally, replacing (2.89), (2.90), (2.91), (2.92) and (2.93) in (2.88), and evaluating in $\alpha = \tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))$ we obtain the desired result

$$\begin{aligned} \frac{\partial^2 \tau_{ML}}{\partial I_{\mathrm{i}} \partial I_j}(\bar{I}^n + t(I^n - \bar{I}^n)) = & \frac{-1}{\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}(\alpha)} - \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^2(\alpha)} \cdot \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right]^2} \cdot \\ & \left[\left[\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^3 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^3} \cdot \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}(\alpha)} - 3\frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha} \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^2(\alpha)} + 2\frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^3(\alpha)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^3\right] \cdot \right. \\ & \frac{\left(-\frac{1}{\lambda_j(\alpha)} \frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)}{\left[\sum_{\mathrm{i}=1}^{n} \frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \cdot \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}(\alpha)} - \frac{\bar{I}_{\mathrm{i}} + t(I_{\mathrm{i}} - \bar{I}_{\mathrm{i}})}{\lambda_{\mathrm{i}}^2(\alpha)} \cdot \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right]} + \frac{\partial^2 \lambda_j(\alpha)}{\partial \alpha^2} \frac{1}{\lambda_j(\alpha)} - \frac{1}{\lambda_j(\alpha)} \left(\frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)^2\Bigg] \cdot \\ & \left.\left(-\frac{1}{\lambda_{\mathrm{i}}(\alpha)} \frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right) + \left(\frac{\partial^2 \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha^2} \frac{1}{\lambda_{\mathrm{i}}(\alpha)} - \frac{1}{\lambda_{\mathrm{i}}(\alpha)} \left(\frac{\partial \lambda_{\mathrm{i}}(\alpha)}{\partial \alpha}\right)^2\right) \cdot \left(-\frac{1}{\lambda_j(\alpha)} \frac{\partial \lambda_j(\alpha)}{\partial \alpha}\right)\right]\Bigg|_{\alpha = \tau_{ML}(\bar{I}^n + t(I^n - \bar{I}^n))}. \end{aligned} \tag{2.94}$$

Chapter 3

Detection in Hypothesis Testing

In the simplest hypothesis testing problem the statistician has to decide on the basis of a sample of size n between the null ($H0$) and alternative ($H1$) hypothesis of which only one is true. Often this task reduces to find a test with a minimal probability of error of type II given prescribed probability of error of type I. The asymptotic performance of the type II error is characterized by the Stein's Lemma [40]. We consider a new dimension to this problem by assuming that the statistician does not have direct access to the data; rather, he/she can be informed about them with a rate constraint, meaning lossy version of the data. Then we are interested in the non-asymptotic rate of convergence of the type II error in this new setting. First of all we formalize the hypothesis testing problem with rate constraint. Then we present some classical results from the simple hypothesis testing by giving the non-asymptotic performance of the type II error given a prescribed type I error. Finally we move to the challenging problem where we have rate restrictions and we present an achievable rate of convergence for the second kind of error.

3.1 Preliminaries

For the rest of this work, we restrict our attention to finite alphabets $\mathbb{X}$ and $\mathbb{Y}$. Denoting $\mathbb{P}(\mathbb{X})$ the family of probability measures on $\mathbb{X}$. The measure and the product measure of the random variables X and Y taking values in $\mathbb{X}$ and $\mathbb{Y}$ will be denoted $\mu_X \in \mathbb{P}(\mathbb{X})$, $\mu_Y \in \mathbb{P}(\mathbb{Y})$ and $\mu_{X,Y} \in \mathbb{P}(\mathbb{X} \times \mathbb{Y})$, respectively. $X_1^n = (X_1, ..., X_n)$ and $Y_1^n = (Y_1, ..., Y_n)$ denote the finite block samples with joint product measure $\mu_{X_1^n, Y_1^n} \triangleq \mu_{X,Y}^n$ ($n-$fold distribution).
Consider a bivariate hypothesis testing defined as

$$\theta = 0 \rightarrow H_0 : (X, Y) \sim \mu_{X,Y}(0) \triangleq \mu_{X,Y} \in \mathbb{P}(\mathbb{X} \times \mathbb{Y})$$
$$\theta = 1 \rightarrow H_1 : (X, Y) \sim \mu_{X,Y}(1) \triangleq \tilde{\mu}_{X,Y} \in \mathbb{P}(\mathbb{X} \times \mathbb{Y})$$

with $0 < D(\mu_{X,Y}(0) || \mu_{X,Y}(1)) < \infty$.

3.2 Inference Problem

The problem to tackle consists on a non-asymptotoc analysis of a particular case named *Test against independence with a fixed rate constraint on one of the modalities.* In particular, our null and alternative hypothesis are described as follows:

$$\theta = 0 \to H_0 : (X, Y) \sim \mu_{X,Y}(0) \triangleq \mu_{X,Y} = \mu_X \cdot \mu_{Y|X} \in \mathbb{P}(\mathbb{X} \times \mathbb{Y})$$
$$\theta = 1 \to H_1 : (\tilde{X}, Y) \sim \mu_{X,Y}(1) \triangleq \tilde{\mu}_{X,Y} = \mu_X \cdot \mu_Y \in \mathbb{P}(\mathbb{X} \times \mathbb{Y}),$$

where $\mu_Y = \sum_{x \in \mathbb{X}} \mu_X(x)\mu_{Y|X}(\cdot|x)$. Consider that the statistician observes Y_1^n samples directly and can be informed about X_1^n samples indirectly, via a fixed rate encoding functions $f_n(\cdot)$ of rate R in bits per sample. More precisely, given the joint vector $(X_1, ..., X_n, Y_1, ..., Y_n)$ an encoding-decoding rule (f_n, ϕ_n) of length n and rate R can be represented by two functions:

- $f_n : \mathbb{X}^n \to \{1, ..., 2^{nR}\}$, (Encoder)
- $\phi_n : \{1, ..., 2^{nR}\} \times \mathbb{Y}^n \to \Theta = \{0, 1\}$, (Decoder).

Now we will define the operational problem following the classical definitions in hypothesis testing [61].

Definition 3.1 *Given $(X_1, Y_1), ..., (X_n, Y_n) \sim \mu_{X,Y}(\theta)$, and given a pair of encoder-decoder (f_n, ϕ_n) of rate R and blocklength n, we define the type I error as*

$$\begin{aligned} P_0(f_n, \phi_n, \mu_X) &\triangleq \mathbb{P}(\phi_n(f_n(X_1^n), Y_1^n) \neq 0 | \theta = 0) \\ &= \mu_{X,Y}^n(A^c(f_n, \phi_n)) \end{aligned} \tag{3.1}$$

where $\mu_{X,Y}^n$ is the $n-$fold distribution ($\mu_{X,Y}^n \triangleq \prod_{\mathrm{i}=1}^{n} \mu_{X,Y}$) and

$$A(f_n, \phi_n) \triangleq \{(x_1, ..., x_n, y_1, ..., y_n) \in \mathbb{X}^n \times \mathbb{Y}^n | \phi_n(f_n(x_1, ..., x_n), y_1, ..., y_n) = 0\}. \tag{3.2}$$

In particular, we can express $A(f_n, \phi_n)$ as

$$A(f_n, \phi_n) = \bigcup_{(\mathrm{i}, y_1, ..., y_n) \in \phi_n^{-1}(\{0\})} f_n^{-1}(\{\mathrm{i}\}) \times \{(y_1, ..., y_n)\}. \tag{3.3}$$

Similarly the type II error is defined as

$$\begin{aligned} P_1(f_n, \phi_n, \mu_X) &\triangleq \mathbb{P}(\phi_n(Y_1^n, f_n(X_1^n)) = 0 | \theta = 1) \\ &= \tilde{\mu}_{X,Y}^n(A(f_n, \phi_n)). \end{aligned} \tag{3.4}$$

and $\tilde{\mu}_{X,Y}^n \triangleq \prod_{\mathrm{i}=1}^{n} \tilde{\mu}_{X,Y}$

Definition 3.2 *We say that a pair of encoder-decoder (f_n, ϕ_n) of length n and rate R operates at type I error $\epsilon > 0$ if for $\mu_X \in \mathbb{P}(\mathbb{X})$*

$$P_0(f_n, \phi_n, \mu_X) \leq \epsilon. \tag{3.5}$$

With this definition we introduce the optimal type II error of rate R and type I error ϵ (function of $\mu_X \in \mathbb{P}(\mathbb{X})$) as the solution of

$$\beta_n(\epsilon, R, \mu_X) \triangleq \min_{\substack{(f_n,\phi_n) \text{ of blocklength } n \text{ and size } R, \\ \text{operating at type I error } \epsilon}} \{P_1(f_n, \phi_n, \mu_X)\}. \tag{3.6}$$

Alternatively, we can recover $\beta_n(\epsilon, R, \mu_X)$ in two-stages as

$$\beta_n(\epsilon, R, \mu_X) = \min_{f_n} \left\{ \beta_n(\epsilon, R, \mu_X, f_n) \middle| \ \frac{1}{n} \log(|f_n(\mathbb{X}^n)|) \leq R \right\}, \tag{3.7}$$

where

$$\beta_n(\epsilon, R, \mu_X, f_n) \triangleq \min\{\tilde{\mu}_{f_n(X_1^n)Y_1^n}(A) |\ A \subseteq f_n(\mathbb{X}^n) \times \mathbb{Y}^n \text{ and } \mu_{f_n(X_1^n)Y_1^n}(A) \geq 1 - \epsilon\} \tag{3.8}$$

Note that this definition follows closely the one adopted by Ahlswede et al. [39] and it represents the optimum for the families of encoder-decoder given a distribution μ_X. In what follows we are interested in the asymptotic error exponent (assuming for a second that this limit exists):

$$\xi(\epsilon, R, \mu_X) \triangleq \lim_{n\to\infty} -\frac{1}{n} \log(\beta_n(\epsilon, R, \mu_X)). \tag{3.9}$$

The following result shows a single letter closed-form characterization for $\xi(\epsilon, R, \mu_X)$.

Theorem 3.3 *(Ahlswede & Csiszár. [39])* $\forall \epsilon > 0$,

$$\xi(\epsilon, R, \mu_X) = \max_{\substack{U: U \to X \to Y \\ I(U;X) \leq R \ \ |\mathbb{U}| \leq |\mathbb{X}|+1}} I(U;Y). \tag{3.10}$$

Interestingly, this result connects this problem of test of independences with communication constraint with the noisy lossy source coding problem using the log-loss (or cross entropy) as the distortion metric [62]. This is precisely (in its asymptotic regime when $n \lim \infty$, see Eq. 3.9) the problem of information bottle-neck [63].

Definition 3.4 *Using the same reasoning that in the problem of universal source coding, given (f_n, ϕ_n) of length n and size R operating at type I error $\epsilon > 0$ for μ_X, we can define the discrepancy in error exponent as*

$$EO(f_n, \phi_n, \mu_X) = \xi(\epsilon, R, \mu_X) - \left(-\frac{1}{n} \log(P_1(f_n, \phi_n, \mu_X)) \right). \tag{3.11}$$

Adopting $EO(f_n, \phi_n, \mu_X)$ as a performance indicator, we can introduce weak and strong notions of universal source coding for this type of rate constrained hypothesis testing problem. Before this, we extend Definition 3.2 to a family $\mathcal{F} \subseteq \mathbb{P}(\mathbb{X})$.

Definition 3.5 *We say that (f_n, ϕ_n) of length n and size R operates at type 1 error $\epsilon > 0$ for the family $\mathcal{F} \subseteq \mathbb{P}(\mathbb{X})$, if $\forall \mu_X \in \mathcal{F}$*

$$P_0(f_n, \phi_n, \mu_X) \leq \epsilon. \tag{3.12}$$

3.3 Non asymptotic analysis for the Error Exponent: no-rate constraint case

Here we present an analysis of the discrepancy in Eq. (3.11) for the case of a known $\mu_X \in \mathbb{P}(\mathbb{X})$ with finite n and no-rate constraint (i.e, $R > H(X)$). In other words, we present a non-asymptotic analysis of the error exponent of the classical Stein's Lemma [40]. We begin with some definitions:

Definition 3.6 *Given $(X_1, Y_1), ..., (X_n, Y_n) \sim \mu_{X,Y}(\theta)$, and given a decoder (test) $\phi_n : \mathbb{X}^n \times \mathbb{Y}^n \to \Theta = \{0, 1\}$ of blocklength n, the type I error is*

$$P_0(\phi_n, \mu_X) \triangleq \mathbb{P}(\phi_n(X_1^n, Y_1^n) \neq 0 | \theta = 0) \tag{3.13}$$
$$= \mu_{X,Y}^n(A^c(\phi_n)), \tag{3.14}$$

where

$$A(\phi_n) \triangleq \{(x_1, ..., x_n, y_1, ..., y_n) \in \mathbb{X}^n \times \mathbb{Y}^n | \phi_n(x_1, ..., x_n, y_1, ..., y_n) = 0\}. \tag{3.15}$$

Similarly the type II error is defined as

$$P_1(\phi_n, \mu_X) \triangleq \mathbb{P}(\phi_n(X_1^n, Y_1^n) = 0 | \theta = 1) \tag{3.16}$$
$$= \tilde{\mu}_{X,Y}^n(A(\phi_n)). \tag{3.17}$$

Definition 3.7 *We say that ϕ_n of length n operates at type I error with no-rate restrictions $\epsilon > 0$ if for $\mu_X \in \mathbb{P}(\mathbb{X})$*

$$P_0(\phi_n, \mu_X) \leq \epsilon. \tag{3.18}$$

Consequently, the operational type II error is defined as

$$\beta_n(\epsilon, \mu_X) \triangleq \min_{\substack{(\phi_n) \text{ of blocklength } n \\ \text{operating at type I error } \epsilon}} \{P_1(\phi_n, \mu_X)\}, \tag{3.19}$$

and the asymptotic error exponent is expressed by (considering that the limit exists)

$$\xi(\epsilon, \mu_X) = \lim_{n \to \infty} \frac{-1}{n} \log(\beta_n(\epsilon, \mu_X)). \tag{3.20}$$

A single letter characterization for $\xi(\epsilon, \mu_X)$ is presented in the following result.

Theorem 3.8 *(Stein [40])* $\forall \epsilon > 0$,

$$\xi(\epsilon, \mu_X) = D(\mu_{X,Y} || \tilde{\mu}_{X,Y}). \tag{3.21}$$

The following two results (theorems 3.9 and 3.10) present an upper (achievability argument) and lower (converse argument) bound for the discrepancy between the asymptotic closed-form expression $\xi(\epsilon, \mu_X)$ and its non-asymptotic counterpart $\frac{-1}{n} \log(\beta_n(\epsilon, \mu_X))$ measured by

$$D(\mu_{X,Y} || \tilde{\mu}_{X,Y}) - \left(-\frac{1}{n} \log(\beta_n(\epsilon, \mu_X)) \right). \tag{3.22}$$

3.3.1 Achievability (Upper bound) and Converse (lower bound) for the discrepancy

Theorem 3.9 *Achievability: Given a fixed $\mu_X \in \mathbb{P}(\mathbb{X})$ and $\epsilon \in (0,1)$ then*

$$\xi(\epsilon, \mu_X) - \left(-\frac{1}{n}\log(\beta_n(\epsilon, \mu_X))\right) \leq -\sqrt{\frac{\sigma^2(\mu_{X,Y}, \tilde{\mu}_{X,Y})}{n}}\Phi^{-1}(\epsilon) + O\left(\frac{1}{n}\right) \quad \forall n \geq 1, \quad (3.23)$$

where $\sigma^2(\mu_{X,Y}, \tilde{\mu}_{X,Y}) \triangleq Var_{X,Y\sim\mu_{X,Y}}\left(\log\left(\frac{\mu_{X,Y}(X,Y)}{\tilde{\mu}_{X,Y}(X,Y)}\right)\right)$ and $\Phi^{-1}(\cdot)$ denotes the inverse of the cumulative distribution function of the standard normal distribution.

Similar, for the converse we have that

Theorem 3.10 *Under the same hypothesis than Theorem 3.9, for any $0 < \epsilon < 1/2$ we have that*

$$\xi(\epsilon, \mu_X) - \left(-\frac{1}{n}\log(\beta_n(\epsilon, \mu_X))\right) \geq \frac{-\log\left(\frac{1}{1-\epsilon-\tilde{\delta}_n(\epsilon)}\right)}{n} - \tilde{\delta}_n(\epsilon) \quad (3.24)$$

where $\tilde{\delta}_n(\epsilon) = \sqrt{\frac{2\ln(1/\epsilon)}{n}} \cdot \sup_{(x,y)\in\mathbb{X}\times\mathbb{Y}}\left|\log\left(\frac{\mu_{X,Y}(\{(x,y)\})}{\tilde{\mu}_{X,Y}(\{(x,y)\})}\right)\right|$.

The proof are presented in Appendix 3.6.1 and 3.6.2, respectively. Theorems 3.9 and 3.10 show a non-asymptotic performance of the type II error for a prescribed type I error. It is well known that the type II error decreases as an exponential rate, this theorems give more precise asymptotic description of the behavior of this probability of error and the rate of which converges to the limit given by $\xi(\epsilon, \mu_X)$. Both bounds are in the same order of magnitude ($O(1/\sqrt{n})$) and, consequently, they are consistent and we can say that they offer optimal rate of convergences to the limit. A corollary of this theorems is when we make the blocklength tends to infinity where we recover the Stein's Lemma.

3.4 Discrepancy analysis: Rate constraint case

Here we consider the more challenging case with a rate constraint. We follow closely the arguments presented in the non-asymptotic case. The following result offers an upper bound for the discrepancy in Eq. (3.11), in the case when we have a rate restriction.

Theorem 3.11 *Given $\mu_X \in \mathbb{P}(\mathbb{X})$ and $\epsilon > 0$, we have that $\forall\gamma > 0$, there exists a scheme $\{(f_n, \phi_n)_{n\geq 1}\}$, operating at type I error $\epsilon > 0$, such that*

$$\xi(\epsilon, R, \mu_X) - \left(-\frac{1}{n}\log(P_1(f_n, \phi_n, \mu_X))\right) \leq \gamma + \delta_n \quad \forall n \geq 1 \quad (3.25)$$

where δ_n is $O(1/\sqrt{n})$. Furthermore, from Eq. (3.25) it follows that eventually in n

$$\xi(\epsilon, R, \mu_X) - \left(\frac{-1}{n}\log(\beta_n(\epsilon, R, \mu_X))\right) \leq O\left(\frac{\log(n)}{n^{1/3}}\right). \quad (3.26)$$

Proof. For the proof of this result we use the following Lemma that is an extension of the achievability and converse argument presented in Section 3.3.

Lemma 3.4.1 Given $\mu_X \in \mathbb{P}(\mathbb{X})$ and an arbitrary function $\tilde{f}_l : \mathbb{X}^l \to \{1, ..., 2^{lR}\}$ there exist a scheme $\{(f_n, \phi_n), n \geq 1\}$ operating at type I error $\epsilon > 0$ of rate R, such that

$$\xi(\epsilon, R, \mu_X) - \left(-\frac{1}{n}\log(P_1(f_n, \phi_n, \mu_X))\right) \leq \left[\xi(\epsilon, R, \mu_X) - \frac{1}{l}D(\mu_{\tilde{f}_l(X^l),Y^l}||\tilde{\mu}_{\tilde{f}_l(X^l),Y^l})\right] + \tilde{\delta}_{n,l}(\epsilon) \tag{3.27}$$

with

$$\tilde{\delta}_{n,l}(\epsilon) = \sqrt{\frac{2\ln(1/\epsilon)}{nl}} \cdot \sup_{(z,y_1,...,y_l)\in\tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l} \left|\log\left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}\right)\right|. \tag{3.28}$$

The proof of Lemma 3.4.1 follows closely the converse argument of Section **??** and is presented in Appendix 3.6.5.

The result in Eq. (3.25) follows from Lemma 3.4.1 and the fact that $\forall\gamma > 0$ we can always find l^* and f^* function of γ such that (see [39], their equation (2.6) and Theorem 3),

$$\xi(\epsilon, R, \mu_X) - \gamma < \frac{1}{l^*}D(\mu_{\tilde{f}_l^*(X^{l^*}),Y^{l^*}}||\tilde{\mu}_{\tilde{f}_l^*(X^{l^*}),Y^{l^*}}) < \xi(\epsilon, R, \mu_X) \tag{3.29}$$

Note that Eq. (3.25) tells us that we can construct an scheme operating at type I error ϵ that has a discrepancy with respect to $(\xi(\epsilon, R, \mu_X) - \gamma)$ that goes to zero at a rate $O\left(\frac{1}{\sqrt{n}}\right)$ as along as we tolerate an offset $\gamma > 0$.
From this result, we can address the case where we make $\gamma = 0$ using the following inequality: $\forall l > 0$ and $\tilde{f}_l : \mathbb{X}^l \to \{1, ..., 2^{lR}\}$ we have from Lemma 3.4.1 and the definition of $\beta_n(\epsilon, R, \mu_X)$ in Eq. (3.6), that:

$$\begin{aligned}
\xi(\epsilon, R, \mu_X) - \left(-\frac{1}{n}\log(\beta_n(\epsilon, R, \mu_X))\right) &\leq \xi(\epsilon, R, \mu_X) - \left(-\frac{1}{n}\log(P_1(f_n, \phi_n, \mu_X))\right) \\
&\leq \xi(\epsilon, R, \mu_X) - \frac{1}{l}D(\mu_{\tilde{f}_l(X^l),Y^l}||\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}) + \tilde{\delta}_{n,l}(\epsilon) \\
&= \left(\max_{\substack{U:U\to X\to Y \\ I(U;X)\leq R \ |\mathbb{U}|\leq|\mathbb{X}|+1}} I(U;Y) - \frac{1}{l}I(\tilde{f}_l(X^l), Y^l)\right) + \tilde{\delta}_{n,l}(\epsilon)
\end{aligned} \tag{3.30}$$

First, we note that the bound in Eq. (3.30) is valid for an arbitrary $\tilde{f}_l$, in particular we can optimize it by the supremum. Second, we know the rate of convergence of $\tilde{\delta}_{n,l}(\epsilon)$ from Eq. (3.28). Then the optimal upper bound obtained from Eq. (3.30) reduces to the analysis of:

$$\max_{\substack{U:U\to X\to Y \\ I(U;X)\leq R \ |\mathbb{U}|\leq|\mathbb{X}|+1}} I(U;Y) - \max_{\tilde{f}_l:\mathbb{X}^l\to\{1,...,2^{lR}\}} \frac{1}{l}I(\tilde{f}_l(X^l), Y^l). \tag{3.31}$$

In other words, this is precisely the non-asymptotic analysis of the information bottleneck problem [63]. It is well-known that this coding problem can be presented as a classical rate-distortion (fixed-rate) source coding problem using the log-loss as the distortion function [68]. More precisely, Eq. (3.31) can be described as

$$\min_{\tilde{f}_l:\mathbb{X}^l\to\{1,...,2^{lR}\}} \frac{1}{l}H(Y^l|\tilde{f}_l(X^l)) - \min_{\substack{U:U\to X\to Y \\ I(U;X)\leq R \ |\mathbb{U}|\leq|\mathbb{X}|+1}} H(Y|U) \tag{3.32}$$

Now, let us consider a family of probability distributions $\mu_\lambda \in \mathbb{P}(\mathbb{Y})$, indexed with a parameter $\lambda \in \Lambda$ where Λ is some alphabet. Given a sequence of parameters $\lambda^n = (\lambda_1, ..., \lambda_n) \in \Lambda^n$ the product probability distribution in $\mathbb{P}(\mathbb{Y})$ is defined as

$$\mu_{\lambda^n}(\{(y_1, ..., y_n)\}) \triangleq \prod_{\mathrm{i}=1}^{n} \mu_{\lambda_\mathrm{i}}(\{y_\mathrm{i}\}) \tag{3.33}$$

Let $\rho(\lambda^n, Y^n) : \Lambda^n \times \mathbb{Y}^n \to \mathbb{R}^+ \cup \{0\}$ denote the logarithmic loss distortion given by:

$$\rho(\lambda^n, y^n) \triangleq -\frac{1}{n} \log \mu_{\lambda_1^n}(\{(y_1, ..., y_n)\}) = \sum_{\mathrm{i}=1}^{n} -\frac{1}{n} \log \mu_{\lambda_\mathrm{i}}(\{y_\mathrm{i}\}). \tag{3.34}$$

Where by construction $\rho(\lambda^n, y^n)$ is additive. Then the following lemma holds:

Lemma 3.4.2 [68] Let (X^l, Y^l) be a random vector with known joint distribution. For any, fixed rate encoding function $\tilde{f}_l : \mathbb{X}^l \to \{1, ..., 2^{lR}\}$ and decoding function $g : \{1, ..., 2^{lR}\} \to \Lambda^n$ such that $g(\tilde{f}_l(X^l)) = \lambda^l$ it follows that

$$\mathbb{E}[\rho(g(u), Y^l) | \tilde{f}_l(X^l) = u] \geq \frac{1}{l} H(Y^l | \tilde{f}_l(X^l) = u). \tag{3.35}$$

Noting that averaging Eq. (3.35) with repect to X^l we get that

$$\mathbb{E}[\rho(g(\tilde{f}_l(X^l)), Y^l)] \geq \frac{1}{l} H(Y^l | \tilde{f}_l(X^l)). \tag{3.36}$$

The term in the LHS of Eq. (3.36) is a case of noisy rate distortion under the logarithmic loss, the encoder $\tilde{f}_l$ and decoder g[1]. It is convenient to define a new distortion function $\tilde{\rho}(x^l, \lambda^l) : \mathbb{X}^l \times \Lambda^l \to \mathbb{R} \cup \{0\}$ as

$$\tilde{\rho}(x^l, \lambda^l) \triangleq \mathbb{E}[\rho(\lambda^l, Y^l) | X^l = x^l]. \tag{3.37}$$

Notice that by this definition $\tilde{\rho}(x^l, \lambda^l) = \sum_{\mathrm{i}=1}^{l} \tilde{\rho}(x_\mathrm{i}, \lambda_\mathrm{i})$ is additive and $\lambda_\mathrm{i} = g_\mathrm{i}(\tilde{f}_l(x^l))$ where g_i is the ith component of function g. Then using $\tilde{f}_l$ as the encoder and g_i as the decoder we recover the classical rate distortion problem [69]. Then we use [[70], Theorem 3] to obtain that

$$\frac{1}{l} H(Y^l | \tilde{f}_l(X^l)) \overset{\text{from lemma 3.4.2}}{\leq} \mathbb{E}_{X \sim \mu_X^l}[\tilde{\rho}(X^l, \lambda^l)] \overset{\text{from [[70], Theorem 3]}}{\leq} D(R) - \frac{\partial}{\partial R} D(R) \frac{\log l}{2l} + o\left(\frac{\log l}{l}\right) \tag{3.38}$$

where $D(R)$ is the noisy rate-distortion function, that precisely correspond in this context to

$$\min_{\substack{U: U \to X \to Y \\ I(U;X) \leq R \ |\mathbb{U}| \leq |\mathbb{X}|+1}} H(Y|U)$$

. Returning to Eq. (3.31), we have from Eq. (3.38) that

$$\max_{\substack{U: U \to X \to Y \\ I(U;X) \leq R \ |\mathbb{U}| \leq |\mathbb{X}|+1}} I(U;Y) - \max_{\tilde{f}_l : \mathbb{X}^l \to \{1, ..., 2^{lR}\}} \frac{1}{l} I(\tilde{f}_l(X^l), Y^l) \leq -\frac{\partial}{\partial R} D(R) \frac{\log l}{2l} + o\left(\frac{\log l}{l}\right), \tag{3.39}$$

[1] see Appendix (pending)

which from Eq. (3.30) implies that

$$\xi(\epsilon, R, \mu_X) - \left(-\frac{1}{n}\log(\beta_n(\epsilon, R, \mu_X))\right) \leq -\frac{\partial}{\partial R}D(R)\frac{\log l}{2l} + \tilde{\delta}_{n,l}(\epsilon) + o\left(\frac{\log l}{l}\right). \tag{3.40}$$

If we look closely to Eq. (3.40) there is a compromise between $\tilde{\delta}_{n,l}(\epsilon)$ and $\frac{\log l}{l}$ function of l. First, we need to give an explicit dependence of l in the term presented in $\tilde{\delta}_{n,l}(\epsilon)$, for this, we use the following lemma

Proposition 3.12 *Consider two arbitrary probability measures $\mu_X, v_X \in \mathbb{P}(\mathbb{X})$, an arbitrary encoder $f_n : \mathbb{X} \to \{1, ..., n\}$. and its induced partition of $\mathbb{X}$ given by:*

$$\pi_n = \{A_{\mathrm{i},n} \triangleq f_n^{-1}(\{\mathrm{i}\}) : \mathrm{i} \in \{1, ..., n\}\} \subset \mathcal{B}(X) \tag{3.41}$$

where $\mathcal{B}(X)$ denotes the power set of $\mathbb{X}$, then

$$\sup_{A\in\pi_n} \frac{\mu_X(A)}{v_X(A)} \leq \sup_{x\in\mathbb{X}} \frac{\mu_X(\{x\})}{v_X(\{x\})} \tag{3.42}$$

Proof. Given $A \in \pi_n$ we note that

$$\frac{\mu_X(A)}{v_X(A)} = \frac{\sum_{j=1}^{|A|} \mu_X(\{j : j \in A\})}{\sum_{j=1}^{|A|} v_X(\{j : j \in A\})} \tag{3.43}$$

Then, given a collection of positive numbers $\{a_{\mathrm{i}} : \mathrm{i} \in \{1, ..., n\}\}$ and $\{b_{\mathrm{i}} : \mathrm{i} \in \{1, ..., n\}\}$, we use the following calculus inequality

$$\frac{\sum_{\mathrm{i}=1}^{n} a_{\mathrm{i}}}{\sum_{\mathrm{i}=1}^{n} b_{\mathrm{i}}} \leq \max_{\mathrm{i}} \left\{\frac{a_{\mathrm{i}}}{b_{\mathrm{i}}}\right\} \tag{3.44}$$

Finally, since A is arbitrary and the positiveness of the probability measure we conclude the desired result. □

Now, define

$$\pi_{l,R} = \{A_{\mathrm{i},l,R} \triangleq \tilde{f}_l^{-1}(\{\mathrm{i}\}) : \mathrm{i} \in \{1, ..., 2^{lR}\}\} \subset \mathcal{B}(X^l) \tag{3.45}$$

Notice that by Eq. (3.45), $\tilde{f}_l$ induces a partition over $\mathbb{X}^l$, then using Proposition 3.12, we get

from $\tilde{\delta}_{n,l}(\epsilon)$ the following upper bound

$$\begin{aligned}
\tilde{\delta}_{n,l}(\epsilon) &= \sqrt{\frac{2\ln(1/\epsilon)}{nl}} \cdot \sup_{(z,y_1,...,y_l)\in \tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l} \left| \log\left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}\right)\right|. \\
&= \sqrt{\frac{2\ln(1/\epsilon)}{nl}} \cdot \sup_{(A_{l,R},y_1,...,y_l)\in \pi_{l,R}\times\mathbb{Y}^l} \left| \log\left(\frac{\mu^l_{X,Y}(A_{l,R}\times\{(y_1,...,y_l)\})}{\tilde{\mu}^l_{X,Y}(A_{l,R}\times\{(y_1,...,y_l)\})}\right)\right| \\
&\leq \sqrt{\frac{2\ln(1/\epsilon)}{nl}} \cdot \sup_{(x_1,...x_l,y_1,...,y_l)\in \mathbb{X}^l\times\mathbb{Y}^l} \left| \log\left(\frac{\mu^l_{X,Y}(\{(x_1,...,x_l,y_1,...,y_l)\})}{\tilde{\mu}^l_{X,Y}(\{(x_1,...,x_l,y_1,...,y_l)\})}\right)\right| \\
&\leq \sqrt{\frac{2l\ln(1/\epsilon)}{n}} \cdot \sup_{(x,y)\in \mathbb{X}\times\mathbb{Y}} \left| \log\left(\frac{\mu_{X,Y}(\{(x,y)\})}{\tilde{\mu}_{X,Y}(\{(x,y)\})}\right)\right|. \qquad (3.46)
\end{aligned}$$

Then the problem reduces to minimize the RHS of Eqs. (3.40) and (3.46) function of l. For that if we propose $l = n^\alpha$ with $\alpha < 1$, the optimum value of α is the consequence of assuming that the two components has the same asymptotic behaviour as a function of n, this reduces to the following matching correlations. $\frac{\log n^\alpha}{n^\alpha} = \sqrt{\frac{n^\alpha}{n}}$, implying that $\alpha^* = 1/3$. Finally replacing in Eq. (3.40) we get the desired result

$$\xi(\epsilon, R, \mu_X) - \left(-\frac{1}{n}\log(\beta_n(\epsilon, R, \mu_X))\right) \leq O\left(\frac{\log(n)}{n^{1/3}}\right). \qquad (3.47)$$

□

3.5 Conclusions and Final Remarks

The results provided in Section 3.3 and 3.4 provided concrete non-asymptotic finite sampling description of the performance of an hypothesis test where we have a rate constraint in one of the sources. For the proof many arguments are used based on the use of sofisticated results from concentration inequalities. For the classical problem we obtain rate of convergences to the limit that are optimal of the order $O(1/\sqrt{n})$. For the more challenging rate constraint problem, we obtain a result that provide a rate of convergence to the asymptotic limit that goes as $O(\ln(n)/n^{1/3})$. Thus we found a discrepancy with the classical setting presented in Section 3.3, which rises the problem of finding a converse result to resolve if this $O(\ln(n)/n^{1/3})$ rate of convergence to the limit is optimal.

3.6 Appendix

3.6.1 Proof of Theorem 3.9

Proof. The proof is based on an achievability argument [64], meaning that we are going to prove that there is a scheme $\{\phi_n\}_{n\geq 1}$ operating at type I error ϵ. For that we consider the

optimal Neyman-Pearson family parametrized by $t \geq 0$, in particular we consider

$$A_{n,t} \triangleq \left\{ (x_1, ..., x_n, y_1, ..., y_n) \in \mathbb{X}^n \times \mathbb{Y}^n \middle| \frac{\mu^n_{X,Y}(\{(x_1, ..., x_n, y_1, ..., y_n)\})}{\tilde{\mu}^n_{X,Y}(\{(x_1, ..., x_n, y_1, ..., y_n)\})} > \mathrm{e}^{nt} \right\}, \tag{3.48}$$

an its induced test denoted by $\phi_{n,t}(\cdot) : \mathbb{X}^n \times \mathbb{Y}^n \mapsto \{0, 1\}$[2]. Choosing this collection, the probability of type I is given by

$$\begin{aligned} P_0(\phi_{n,t}, \mu_X) &= \mathbb{P}(\phi_{n,t}(X^n_1, Y^n_1) \neq 0 | \theta = 0) \\ &= \mu^n_{X,Y} \left(\left\{ (x_1, ..., x_n, y_1, ..., y_n) \in \mathbb{X}^n \times \mathbb{Y}^n \middle| \frac{\mu^n_{X,Y}(\{(x_1, ..., x_n, y_1, ..., y_n)\})}{\tilde{\mu}^n_{X,Y}(\{(x_1, ..., x_n, y_1, ..., y_n)\})} \leq \mathrm{e}^{nt} \right\} \right). \end{aligned} \tag{3.49}$$

Due to the construction of the test it follows directly that

$$P_1(\phi_{n,t}, \mu_X) \leq \mathrm{e}^{-nt}, \tag{3.50}$$

and consequently, we are looking for

$$t^*_n(\epsilon) \triangleq \sup\{t : \mu^n_{X,Y}(A^c_{n,t}) \leq \epsilon\}. \tag{3.51}$$

For this, it is useful to consider $\tilde{t}_n(\epsilon) = D(\mu_{X,Y}||\tilde{\mu}_{X,Y}) + \sqrt{\frac{\sigma^2(\mu_{X,Y}, \tilde{\mu}_{X,Y})}{n}} \Phi^{-1} \left(\epsilon - \frac{6\kappa(\mu_{X,Y}, \tilde{\mu}_{X,Y})}{\sigma^2(\mu_{X,Y}, \tilde{\mu}_{X,Y})^{3/2}\sqrt{n}} \right)$ where $\kappa(\mu_{X,Y}, \tilde{\mu}_{X,Y})$ is the absolute third moment associated to the random variable $\log \left(\frac{\mu_{X,Y}(X,Y)}{\tilde{\mu}_{X,Y}(X,Y)} \right)$. Choosing this $\tilde{t}_n(\epsilon)$, and applying the Berry-Esséen theorem [65] it follows that:

$$\begin{aligned} &P_0(\phi_{n,t}, \mu_X) \\ &= \mathbb{P} \left(\frac{1}{n} \sum_{\mathrm{i}=1}^{n} \log \left(\frac{\mu_{X,Y}(X_\mathrm{i}, Y_\mathrm{i})}{\tilde{\mu}_{X,Y}(X_\mathrm{i}, Y_\mathrm{i})} \right) \leq \tilde{t}_n(\epsilon) \middle| \theta = 0 \right) \\ &\overset{(1)}{\leq} \frac{C\kappa(\mu_{X,Y}, \tilde{\mu}_{X,Y})}{\sqrt{n}\sigma^2(\mu_{X,Y}, \tilde{\mu}_{X,Y})^{3/2}} + \Phi \left((\tilde{t}_n(\epsilon) - D(\mu_{X,Y}||\tilde{\mu}_{X,Y})) \sqrt{\frac{n}{\sigma^2(\mu_{X,Y}, \tilde{\mu}_{X,Y})}} \right) \\ &\overset{(2)}{=} \frac{C\kappa(\mu_{X,Y}, \tilde{\mu}_{X,Y})}{\sqrt{n}\sigma^2(\mu_{X,Y}, \tilde{\mu}_{X,Y})^{3/2}} + \epsilon - \frac{6\kappa(\mu_{X,Y}, \tilde{\mu}_{X,Y})}{\sqrt{n}\sigma^2(\mu_{X,Y}, \tilde{\mu}_{X,Y})^{3/2}} \\ &\overset{(3)}{\leq} \epsilon, \end{aligned} \tag{3.52}$$

where for (1) we applied the Berry-Esséen theorem (the statement of the theorem is presented in Appendix 3.6.3), for (2) we replaced the value of $\tilde{t}_n(\epsilon)$ and for (3) we use the fact that constant C provided is less than 6 (see Appendix 3.6.3). Finally, replacing this value of $\tilde{t}_n(\epsilon)$ in Eq. (3.50) and using a Taylor approximation for $\Phi^{-1}(\cdot)$ of first order around ϵ we have that

[2]Meaning that $\phi_{n,t}(x_1, ..., x_n, y_1, ..., y_n) = 0$ if $(x_1, ..., x_n, y_1, ..., y_n) \in A_{n,t}$

$$
\begin{aligned}
\frac{\log(\beta_n(\epsilon,\mu_X))}{n} &\leq \frac{\log(P_1(\phi_{n,t},\mu_X))}{n} \\
&\leq -D(\mu_{X,Y}||\tilde{\mu}_{X,Y}) - \sqrt{\frac{\sigma^2(\mu_{X,Y},\tilde{\mu}_{X,Y})}{n}}\Phi^{-1}\left(\epsilon - \frac{6\kappa(\mu_{X,Y},\tilde{\mu}_{X,Y})}{\sigma^2(\mu_{X,Y},\tilde{\mu}_{X,Y})^{3/2}\sqrt{n}}\right) \\
&= -D(\mu_{X,Y}||\tilde{\mu}_{X,Y}) - \sqrt{\frac{\sigma^2(\mu_{X,Y},\tilde{\mu}_{X,Y})}{n}}\left[\Phi^{-1}(\epsilon) - \left.\frac{\mathrm{d}\Phi^{-1}(x)}{\mathrm{d}x}\right|_{x=\xi}\frac{6\kappa(\mu_{X,Y},\tilde{\mu}_{X,Y})}{\sigma^2(\mu_{X,Y},\tilde{\mu}_{X,Y})^{3/2}\sqrt{n}}\right] \\
&\overset{(1)}{\leq} -D(\mu_{X,Y}||\tilde{\mu}_{X,Y}) - \sqrt{\frac{\sigma^2(\mu_{X,Y},\tilde{\mu}_{X,Y})}{n}}\left[\Phi^{-1}(\epsilon) - \left.\frac{\mathrm{d}\Phi^{-1}(x)}{\mathrm{d}x}\right|_{x=\epsilon}\frac{6\kappa(\mu_{X,Y},\tilde{\mu}_{X,Y})}{\sigma^2(\mu_{X,Y},\tilde{\mu}_{X,Y})^{3/2}\sqrt{n}}\right] \\
&= -D(\mu_{X,Y}||\tilde{\mu}_{X,Y}) - \sqrt{\frac{\sigma^2(\mu_{X,Y},\tilde{\mu}_{X,Y})}{n}}\Phi^{-1}(\epsilon) + O\left(\frac{1}{n}\right), \qquad (3.53)
\end{aligned}
$$

where for (1) we consider $\xi \in \left(\epsilon - \frac{6\kappa(\mu_{X,Y},\tilde{\mu}_{X,Y})}{\sigma^2(\mu_{X,Y},\tilde{\mu}_{X,Y})^{3/2}\sqrt{n}}, \epsilon\right)$ and the monotonicity of $\Phi^{-1}(\cdot)$. □

3.6.2 Proof of Theorem 3.10

Proof. Let us consider the set

$$
A^c_{n,\delta} \triangleq \left\{(x_1,...,x_n,y_1,...,y_n) \in \mathbb{X}^n \times \mathbb{Y}^n \left|\left|\frac{1}{n}\log\left(\frac{\mu^n_{X,Y}(\{(x_1,...,x_n,y_1,...,y_n)\})}{\tilde{\mu}^n_{X,Y}(\{(x_1,...,x_n,y_1,...,y_n)\})}\right) - D(\mu_{X,Y}||\tilde{\mu}_{X,Y})\right| \geq \delta\right.\right\}, \qquad (3.54)
$$

we have the following:

Lemma 3.6.1 (Chapter 11, Sect 11.8 [66]) For an arbitrary set $B_n \subseteq \mathbb{X}^n \times \mathbb{Y}^n$ and its induced test ϕ_n[3] such that operates at type I error (i.e. $\mu^n_{X,Y}(B^c_n) \leq \epsilon$) then

$$
\tilde{\mu}^n_{X,Y}(B_n) > (1-\epsilon-\delta)2^{-n(D(\mu_{X,Y}||\tilde{\mu}_{X,Y})+\delta)}. \qquad (3.55)
$$

By construction, it is clear that there exists $\delta > 0$ such that $A^c_{n,\delta}$ operates at type I error ϵ. In fact we consider

$$
\delta^*_n(\epsilon) \triangleq \sup\{\delta : \mu^n_{X,Y}(A^c_{n,\delta}) \leq \epsilon\}. \qquad (3.56)
$$

To characterize $\delta^*_n(\epsilon)$ we use that:

$$
\frac{1}{n}\log\left(\frac{\mu^n_{X,Y}(X_1,...,X_n,Y_1,...,Y_n)}{\tilde{\mu}^n_{X,Y}(X_1,...,X_n,Y_1,...,Y_n)}\right) = \sum_{(x,y)\in\mathbb{X}\times\mathbb{Y}} \hat{\mu}_n(x,y)\log\left(\frac{\mu_{X,Y}(\{(x,y)\})}{\tilde{\mu}_{X,Y}(\{(x,y)\})}\right), \qquad (3.57)
$$

where $\hat{\mu}_n(x,y)$ denotes the empirical distribution induced by $(X_1,...,X_n,Y_1,...,Y_n)$. Then we can use the bounded difference inequality [67] to bound $\mu^n_{X,Y}(A^c_{n,\delta})$. For that we need to

[3] Meaning that $\phi_n(x_1,...,x_n,y_1,...,y_n) = 0$ if $(x_1,...,x_n,y_1,...,y_n) \in B_n$

obtain an expression for the bounded difference[4] given by:

$$\sup_{\mathrm{i}\in\{1,\dots,n\}} \sup_{(X_1,Y_1),\dots,(X_n,Y_n),(\tilde{X}_\mathrm{i},\tilde{Y}_\mathrm{i})\in\mathbb{X}\times\mathbb{Y}} \left| \sum_{(x,y)\in\mathbb{X}\times\mathbb{Y}} \hat{\mu}_n(x,y)\log\left(\frac{\mu_{X,Y}(\{(x,y)\}}{\tilde{\mu}_{X,Y}(\{(x,y)\})}\right) - \sum_{(x,y)\in\mathbb{X}\times\mathbb{Y}} \tilde{\mu}_n(x,y)\log\left(\frac{\mu_{X,Y}(\{(x,y)\})}{\tilde{\mu}_{X,Y}(\{(x,y)\})}\right)\right|$$
$$\leq \frac{2}{n}\cdot \sup_{(x,y)\in\mathbb{X}\times\mathbb{Y}} \left|\log\left(\frac{\mu_{X,Y}(\{(x,y)\})}{\tilde{\mu}_{X,Y}(\{(x,y)\})}\right)\right|, \tag{3.58}$$

where $\tilde{\mu}_n(x,y)$ denotes the empirical distribution induced by $(X_1,...,\tilde{X}_\mathrm{i},...,X_n,Y_1,...,\tilde{Y}_\mathrm{i},...,Y_n)$. Then the bounded difference inequality [67] (Sec 2.3) tell us that:

$$\underbrace{\mathbb{P}\left(\left|\frac{1}{n}\log\left(\frac{\mu^n_{X,Y}(X_1,...,X_n,Y_1,...,Y_n)}{\tilde{\mu}^n_{X,Y}(X_1,...,X_n,Y_1,...,Y_n)}\right) - D(\mu_{X,Y}||\tilde{\mu}_{X,Y})\right| \geq \delta_n \middle| \theta = 0\right)}_{\mu^n_{X,Y}(A^c_{n,\delta})}$$
$$\leq \exp\left(\frac{-n\delta_n^2}{2\left(\sup_{(x,y)\in\mathbb{X}\times\mathbb{Y}}\left|\log\left(\frac{\mu_{X,Y}(\{(x,y)\})}{\tilde{\mu}_{X,Y}(\{(x,y)\})}\right)\right|\right)^2}\right). \tag{3.59}$$

Using this bound, a lower bound for $\delta^*_n(\epsilon)$ is given by the solution of:

$$\exp\left(\frac{-n(\tilde{\delta}_n(\epsilon))^2}{2\left(\sup_{(x,y)\in\mathbb{X}\times\mathbb{Y}}\left|\log\left(\frac{\mu_{X,Y}(\{(x,y)\})}{\tilde{\mu}_{X,Y}(\{(x,y)\})}\right)\right|\right)^2}\right) = \epsilon \tag{3.60}$$

which give us

$$\delta^*_n(\epsilon) \geq \tilde{\delta}_n(\epsilon) = \sqrt{\frac{2\ln(1/\epsilon)}{n}}\cdot \sup_{(x,y)\in\mathbb{X}\times\mathbb{Y}}\left|\log\left(\frac{\mu_{X,Y}(\{(x,y)\})}{\tilde{\mu}_{X,Y}(\{(x,y)\})}\right)\right|. \tag{3.61}$$

Finally, using $\tilde{\delta}_n(\epsilon)$ in Eq. (3.61) replacing in Eq. (3.55) and taking logarithm we have that for all B_n with $\mu^n_{X,Y}(B^c_n) \leq \epsilon$,

$$\log(\tilde{\mu}_{X,Y}(B_n)) > \log(1-\epsilon-\tilde{\delta}_n(\epsilon)) - n(D(\mu_{X,Y}||\tilde{\mu}_{X,Y}) + \tilde{\delta}_n(\epsilon)). \tag{3.62}$$

Since B_n is arbitrary, we conclude that

$$\xi(\epsilon,R,\mu_X) - \left(-\frac{1}{n}\log(P_1(f_n,\phi_n,\mu_X))\right) \geq \frac{\log(1-\epsilon-\tilde{\delta}_n(\epsilon))}{n} - \tilde{\delta}_n(\epsilon). \tag{3.63}$$

□

[4]For completeness the theorem is presented in Appendix 3.6.4

3.6.3 Berry-Esséen theorem

There exits a positive constant C such that if $X_1, X_2, ..., X_n$ are i.i.d. random variables with $\mathbb{E}_{X_1\sim\mu_{X_1}}(X_1) = 0$, $\mathbb{E}_{X_1\sim\mu_{X_1}}(X_1^2) = \sigma^2 > 0$ and $\mathbb{E}_{X_1\sim\mu_{X_1}}(|X_1|^3) = \rho < \infty$ and if we define

$$Y_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

the sample mean, and $F_n(\cdot)$ the cumulative distribution function of

$$\frac{Y_n\sqrt{n}}{\sigma}$$

and $\Phi(\cdot)$ the cumulative distribution function of the standard normal distribution, then for all x and n,

$$|F_n(x) - \Phi(x)| \leq \frac{C\rho}{\sigma^3\sqrt{n}}$$

That is: given a sequence of independent and identically distributed random variables, each having mean zero and positive variance, if additionally the third absolute moment is finite, then the cumulative distribution functions of the standardized sample mean and the standard normal distribution differ by no more than the specified amount. Note that the approximation error for all n (and hence the limiting rate of convergence for indefinite n sufficiently large) is bounded by the order of $n^{-1/2}$. The best estimate of C up to date is $0.40973 \leq C \leq 0.4748$ [71].

3.6.4 Bounded difference inequality

Let A be some set, and assume a function $g : A^n \to \mathbb{R}$ satisfies the bounded difference assumption

$$\sup_{x_1,...,x_n,\tilde{x}_\mathrm{i}\in A} |g(x_1, ..., x_n) - g(x_1, ..., x_{\mathrm{i}-1}, \tilde{x}_\mathrm{i}, x_{\mathrm{i}+1}, ..., x_n)| \leq c_\mathrm{i}, \quad \forall \mathrm{i} \in \{1, ..., n\}$$

In other words, we assume that if we change the ith variable of g while keeping all the others fixed, then the value of the function does not change by more than c_i. Consider $X_1, ..., X_n$ are independent random variables taking values in A. Under the bounded difference assumption, for all $t > 0$,

$$\mathbb{P}(g(X_1, ..., X_n) - \mathbb{E}_{X_1,...,X_n\sim\mu_{X_1,...,X_n}}(g(X_1, ..., X_n)) \geq t) \leq \mathrm{e}^{-2t^2/\sum_{\mathrm{i}=1}^{n} c_\mathrm{i}^2}$$

and

$$\mathbb{P}(\mathbb{E}_{X_1,...,X_n\sim\mu_{X_1,...,X_n}}(g(X_1, ..., X_n)) - g(X_1, ..., X_n) \geq t) \leq \mathrm{e}^{-2t^2/\sum_{\mathrm{i}=1}^{n} c_\mathrm{i}^2}$$

3.6.5 Proof of Lemma 3.4.1

Proof. For an arbitrary encoder f_n let us consider the family of optimal Neyman-Pearson tests:

$$B_{n,t}(f_n) \triangleq \left\{ (x_1, ..., x_n, y_1, ..., y_n) \middle| \frac{\mu_{f_n(X^n),Y^n}(\{(z = f_n(x_1, ..., x_n), y_1, ..., y_n)\})}{\tilde{\mu}_{f_n(X^n),Y^n}(\{(z = f_n(x_1, ..., x_n), y_1, ..., y_n)\})} > \mathrm{e}^{nt} \right\}, \tag{3.64}$$

parametrized by t and its induced test denoted by $\phi_{n,t}(\cdot) : \{1, ..., 2^{nR}\} \times \mathbb{Y}^n \mapsto \{0, 1\}$ [5]. Using this set, the probability of type I is given by

$$\begin{aligned} P_0(f_n, \phi_{n,t}, \mu_X) &= \mathbb{P}(\phi_{n,t}(f_n(X_1^n), Y_1^n) \neq 0 | \theta = 0) \\ &= \mu_{X,Y}^n \left(\left\{ (x_1, ..., x_n, y_1, ..., y_n) \middle| \frac{\mu_{f_n(X^n),Y^n}(\{(z = f_n(x_1, ..., x_n), y_1, ..., y_n)\})}{\tilde{\mu}_{f_n(X^n),Y^n}(\{(z = f_n(x_1, ..., x_n), y_1, ..., y_n)\})} \leq \mathrm{e}^{nt} \right\} \right). \end{aligned} \tag{3.65}$$

We also obtain an upper bound for the type II error by:

$$P_1(f_n, \phi_{n,t}, \mu_X) \leq \mathrm{e}^{-nt}, \tag{3.66}$$

then the problem to tackle consist on

$$t_n^*(\epsilon) = \sup_{f_n \text{ encoder of rate } R} \sup_t \{t : \mu_{X,Y}^n(B_{n,t}^c(f_n)) \leq \epsilon\}. \tag{3.67}$$

Note that f_n in general is an arbitrary function that breaks the i.i.d assumption then determining $t_n^*(\epsilon)$ is not a feasible task. At this point it is convenient to derive a lower bound for $t_n^*(\epsilon)$. For this we use $\tilde{f}_l$ (of rate R) to construct an i.i.d setting of blocklength l. More precisely we construct f_n applying the function $\tilde{f}_l$ to every sub-block of length l, k times such that $n = kl$

$$\tilde{f}_{n,l}(x_1, ..., x_l, x_{l+1}, ..., x_{2l}, ..., x_{l(k-1)+1}, ..., x_{kl}) = (\tilde{f}_l(x_1, ..., x_l), \tilde{f}_l(x_{l+1}, ..., x_{2l}), ..., \tilde{f}_l(x_{l(k-1)+1}, ..., x_{kl})). \tag{3.68}$$

Recall $B_{n,t}(\tilde{f}_{n,l})$, in this case it is convenient to re-parametrize t as $t = \frac{1}{l} D(\mu_{\tilde{f}_l(X^l),Y^l} || \tilde{\mu}_{\tilde{f}_l(X^l),Y^l}) - \delta$ for $\delta > 0$. The type I error can be upper bounded as follows

$$\begin{aligned} \mu_{X,Y}^n \left(B_{n,t}^c(\tilde{f}_{n,l})) \right) &= \mathbb{P} \left(\frac{1}{k} \log \left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}^k(Z_1, ..., Z_k, Y_1, ..., Y_{kl})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}^k(Z_1, ..., Z_k, Y_1, ..., Y_{kl})} \right) \leq D(\mu_{\tilde{f}_l(X^l),Y^l} || \tilde{\mu}_{\tilde{f}_l(X^l),Y^l}) - l\delta \middle| \theta = 0 \right) \\ &\leq \mathbb{P} \left(\left| \frac{1}{k} \log \left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}^k(Z_1, ..., Z_k, Y_1, ..., Y_{kl})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}^k(Z_1, ..., Z_k, Y_1, ..., Y_{kl})} \right) - D(\mu_{\tilde{f}_l(X^l),Y^l} || \tilde{\mu}_{\tilde{f}_l(X^l),Y^l}) \right| \geq l\delta \middle| \theta = 0 \right) \end{aligned} \tag{3.69}$$

with $Z_k = \tilde{f}_l(X_{l(k-1)+1}, ..., X_{kl})$. Notice that

$$\frac{1}{k} \log \left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}^k(Z_1, ..., Z_k, Y_1, ..., Y_{kl})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}^k(Z_1, ..., Z_k, Y_1, ..., Y_{kl})} \right) = \sum_{(z, y_1, ..., y_l) \in \tilde{f}_l(\mathbb{X}^l) \times \mathbb{Y}^l} \hat{\mu}_k(z, y_1, ..., y_l) \log \left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z, y_1, ..., y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z, y_1, ..., y_l)\})} \right) \tag{3.70}$$

[5] Meaning that $\phi_{n,t}(f_n(x_1, ..., x_n), y_1, ..., y_n) = 0$ if $(x_1, ..., x_n, y_1, ..., y_n) \in B_{n,t}(f_n)$

and

$$\sup_{\substack{(Z_1,Y_1,...,Y_l),...,\\(Z_k,Y_{l(k-1)+1},...,Y_{lk}),\\(Z_{i'},Y_{l(i'-1)+1},...,Y_{li'})\\ \in\tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l}} \left| \sum_{\substack{(z,y_1,...,y_l)\\ \in\tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l}} \hat{\mu}_k(z,y_1,...,y_l)\log\left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}\right)... \right. \tag{3.71}$$

$$\left. - \sum_{\substack{(z,y_1,...,y_l)\\ \in\tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l}} \hat{\mu}_k(z,y_1,...,y_l)\log\left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}\right)\right|$$

$$\leq \frac{2}{k}\cdot \sup_{(z,y_1,...,y_l)\in\tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l}\left|\log\left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}\right)\right| \quad \forall \mathrm{i}\in\{1,...,k\}. \tag{3.72}$$

Again, using bounded difference inequality [67] we get

$$\mathbb{P}\left(\left|\frac{1}{k}\log\left(\frac{\mu^k_{\tilde{f}_l(X^l),Y^l}(Z_1,...,Z_k,Y_1,...,Y_{kl})}{\tilde{\mu}^k_{\tilde{f}_l(X^l),Y^l}(Z_1,...,Z_k,Y_1,...,Y_{kl})}\right) - D(\mu_{\tilde{f}_l(X^l),Y^l}||\tilde{\mu}_{\tilde{f}_l(X^l),Y^l})\right| \geq l\delta \middle| \theta=0\right) \tag{3.73}$$

$$\leq \exp\left(\frac{-k(l\delta_n(l))^2}{2\sup\limits_{(z,y_1,...,y_l)\in\tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l}\left|\log\left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}\right)\right|^2}\right). \tag{3.74}$$

Finally a lower bound for $t^*_n(\epsilon)$ can be determined as function of $\tilde{\delta}_{n,l}(\epsilon))$ that is the solution of the following equality.

$$\exp\left(\frac{-k(l\tilde{\delta}_{n,l}(\epsilon))^2}{2\sup\limits_{(z,y_1,...,y_l)\in\tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l}\left|\log\left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}\right)\right|^2}\right) = \epsilon. \tag{3.75}$$

More precisely, we have that

$$t^*_n(\epsilon) \geq \frac{1}{l}D(\mu_{\tilde{f}_l(X^l),Y^l}||\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}) - \sqrt{\frac{2\log(1/\epsilon)}{nl}}\cdot \sup_{(z,y_1,...,y_l)\in\tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l}\left|\log\left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}\right)\right| \tag{3.76}$$

Finally, replacing the bound of Eq. (3.76) in Eq. (3.66) and taking logarithm we have that:

$$\xi(\epsilon,R,\mu_X) - \left(-\frac{1}{n}\log(P_1(\tilde{f}_{n,l},\phi_n,\mu_X))\right) \leq \left[\xi(\epsilon,R,\mu_X) - \frac{1}{l}D(\mu_{\tilde{f}_l(X^l),Y^l}||\tilde{\mu}_{\tilde{f}_l(X^l),Y^l})\right] + \tilde{\delta}_{n,l}(\epsilon), \tag{3.77}$$

with

$$\tilde{\delta}_{n,l}(\epsilon) = \sqrt{\frac{2\ln(1/\epsilon)}{nl}}\cdot \sup_{(z,y_1,...,y_l)\in\tilde{f}_l(\mathbb{X}^l)\times\mathbb{Y}^l}\left|\log\left(\frac{\mu_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}{\tilde{\mu}_{\tilde{f}_l(X^l),Y^l}(\{(z,y_1,...,y_l)\})}\right)\right|, \tag{3.78}$$

which concludes the result. □

Chapter 4

Conclusion

We summarize our main findings for the estimation and detection problems in sections 4.1 and 4.2 respectively.

4.1 Astrometry

We study the performance of the WLS and ML estimators for relative astrometry on digital detectors subject to Poisson noise, in comparison with the best possible attainable precision given by the CR bound. Our study includes analytical results, and numerical simulations under realistic observational conditions to help us to corroborate our theoretical findings.

By generalizing the technical result presented in [50] we are able to obtain, for the first time, close-form expressions for the variance and the mean of implicit estimators (as is the particular case of the WLS and ML schemes), which can be computed directly from the data (see Theorem 2.2, in particular Eqs. (2.22) and (2.23), and Appendix A). When specifying this result to astrometry with digital detectors, we are able to bound both the bias and the variance of the relative position of a celestial source on a CCD array as a function of all the relevant parameters of the problem (see Eqs. (2.27) and (2.28) or Eqs. (2.33) and (2.34) for the WLS and ML estimators respectively). We verified that the bias of the WLS and ML methods are negligible in all the observational regimes explored in this paper (see Fig. 2.1).

A careful analysis of our predictions confirms earlier results by [22] (for the LS method) in that the WLS method is, in general, sub-optimal (in comparison with the MVB given by the CR result), specially at high and very high S/N (see the two bottom panels on Fig. 2.2). However, a judicious data-driven selection of weights (called "adaptive" WLS method by us, Sect. (2.4.5)), improves the performance of the WLS substantially (see Fig. 2.7). This is an interesting result, given the widespread use and simple numerical implementation of the WLS method.

The ML method is found to have both a smaller bias than the WLS method (compare left and right panels of Fig. 2.1), although the bias on both methods is already quite small), and

a tight correspondence to the MVB throughout the entire range of S/N regimes explored in this paper (Fig. 2.5). Therefore, the ML estimator for astrometry is consistently optimal, and should be the estimator of choice for high-precision applications.

This paper, along with [22], completes an in-depth study of the performance of commonly used estimators in astrometry using PIDs, and sets the stage for the development of codes that could efficiently implement astrometric ML estimators on 2D detectors, incorporating also the simultaneous measurement of fluxes, as explored in [60].

4.2 Hypothesis testing

This work characterizes an achievable bound for the optimal error exponent in the bivariate hypothesis testing setting with communication constraint. The main technical challenge is the fact that there exists many mathematical difficulties when dealing with the likelihood of the ratio in the rate constraint problem. Developing bounds to the induced measure given by the encoder function is not an easy task due to the freedom of such function. To address this technical issue, this work extends the approach of Zhang et al. [70] to the case of noisy rate distortion theory and obtain fundamental bounds via concentration inequalities such as the bounded difference inequality and the Berry-Esséen theorem.

The achievable bound seems to offer a reasonable velocity of convergence to the limit, however, we do not know if this rate is optimal without a converse result. As a future work we want to find a converse bound for this problem, the main technical challenge of the converse falls into analysing the function f (enconders), because this function breaks the i.i.d assumption of one of the sources, and consequently, we can not use classical concentration arguments (based of sum of independent measurements) as used in the unconstrained case. One attractive path of the future research in this direction is to find a way to connect this likelihood ratio to an additive distortion measure, in this way, we can use Zhang's theorem [70] to derive a converse tight bound.

Bibliography

[1] S. Reffert, "Astrometric measurement techniques," *New Astronomy Reviews*, vol. 53, no. 11, pp. 329–335, 2009.

[2] E. Høg, "Astrometric accuracy during the past 2000 years," *arXiv preprint arXiv:1707.01020*, 2017.

[3] G. F. Benedict, B. E. McArthur, E. P. Nelan, and T. E. Harrison, "Astrometry with hubble space telescope fine guidance sensors—a review," *Publications of the Astronomical Society of the Pacific*, vol. 129, no. 971, p. 012001, 2016.

[4] G. Collaboration, T. Prusti, J. de Bruijne, *et al.*, "Astronomy and astrophysics, 595," 2016.

[5] W. F. van Altena, *Astrometry for Astrophysics: Methods, Models, and Applications.* Cambridge University Press, 2013.

[6] C. Cacciari, E. Pancino, and M. Bellazzini, "Gaia," *Astronomische Nachrichten*, vol. 337, no. 8-9, pp. 899–903, 2016.

[7] M. Lattanzi, "Astrometric cosmology," *Memorie della Societa Astronomica Italiana*, vol. 83, p. 1033, 2012.

[8] I. R. King, "Accuracy of measurement of star images on a pixel array," *Publications of the Astronomical Society of the Pacific*, pp. 163–168, 1983.

[9] P. B. Stetson, "Daophot: A computer program for crowded-field stellar photometry," *Publications of the Astronomical Society of the Pacific*, pp. 191–222, 1987.

[10] C. Alard and R. H. Lupton, "A method for optimal image subtraction," *The Astrophysical Journal*, vol. 503, no. 1, p. 325, 1998.

[11] J.-F. Lee and W. Van Altena, "Theoretical studies of the effects of grain noise on photographic stellar astrometry and photometry," *The Astronomical Journal*, vol. 88, pp. 1683–1689, 1983.

[12] R. C. Stone, "A comparison of digital centering algorithms," *The Astronomical Journal*, vol. 97, pp. 1227–1237, Apr. 1989.

[13] M. Vakili and D. W. Hogg, "Do fast stellar centroiding methods saturate the cram\'{e}r-rao lower bound?," *arXiv preprint arXiv:1610.05873*, 2016.

[14] W. Van Altena and L. Auer, "Digital image centering, i," in *Image Processing Techniques in Astronomy*, pp. 411–418, Springer, 1975.

[15] L. Auer and W. Van Altena, "Digital image centering, ii," *The Astronomical Journal*, vol. 83, pp. 531–537, 1978.

[16] L. Chun-Lin, "Maximum likelihood estimation of digital image center," in *Developments in Astrometry and their Impact on Astrophysics and Geodynamics*, pp. 113–116, Springer, 1993.

[17] R. Gray and L. D. Davisson, *Introduction to Statistical Signal Processing.* Cambridge Univ Press, 2004.

[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley Interscience, New York, second ed., 2006.

[19] L. Lindegren, "High-accuracy positioning: astrometry," *ISSI Scientific Reports Series*, vol. 9, pp. 279–291, 2010.

[20] R. A. Mendez, J. F. Silva, and R. Lobos, "Analysis and interpretation of the cramér-rao lower-bound in astrometry: One-dimensional case," *Publications of the Astronomical Society of the Pacific*, vol. 125, no. 927, pp. pp. 580–594, 2013.

[21] R. A. Méndez, J. F. Silva, R. Orostica, and R. Lobos, "Analysis of the cramér-rao bound in the joint estimation of astrometry and photometry," *Publications of the Astronomical Society of the Pacific*, vol. 126, no. 942, pp. 798–810, 2014.

[22] R. A. Lobos, J. F. Silva, R. A. Mendez, and M. Orchard, "Performance analysis of the least-squares estimator in astrometry," *Publications of the Astronomical Society of the Pacific*, vol. 127, no. 957, p. 1166, 2015.

[23] L. Lindegren, "Photoelectric astrometry-a comparison of methods for precise image location," in *IAU Colloq. 48: Modern Astrometry*, vol. 1, pp. 197–217, 1978.

[24] P. Jakobsen, P. Greenfield, and R. Jedrzejewski, "The cramer-rao lower bound and stellar photometry with aberrated hst images," *Astronomy and astrophysics*, vol. 253, no. 1, pp. 329–332, 1992.

[25] T. Zaccheo, R. Gonsalves, S. Ebstein, and P. Nisenson, "Estimating the cramer-rao bound for restored astronomical observations," *The Astrophysical Journal*, vol. 439, pp. L43–L45, 1995.

[26] H.-M. Adorf, "Limits to the precision of joint flux and position measurements on array data," in *Astronomical Data Analysis Software and Systems V*, vol. 101, p. 13, 1996.

[27] U. Bastian, "The maximum reachable astrometric precision - The Cramer-Rao Limit."

2004BASNOCODE, April 2004.

[28] C. Radhakrishna Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bulletin of the Calcutta Mathematical Society*, vol. 37, no. 3, pp. 81–91, 1945.

[29] H. Cramér, "A contribution to the theory of statistical estimation," *Scandinavian Actuarial Journal*, vol. 1946, no. 1, pp. 85–94, 1946.

[30] S. Bouquillon, R. Mendez, M. Altmann, T. Carlucci, C. Barache, F. Taris, A. Andrei, and R. Smart, "Characterizing the astrometric precision limit for moving targets observed with digital-array detectors," *Astronomy & Astrophysics*, vol. 606, p. A27, 2017.

[31] M. Altmann, S. Bouquillon, F. Taris, I. A. Steele, R. L. Smart, A. H. Andrei, C. Barache, T. Carlucci, and S. G. Els, "Gbot-ground based optical tracking of the gaia satellite," in *Proc. SPIE*, vol. 9149, p. 25, 2014.

[32] J. Zhang, Y. C. Hao, L. Wang, and Y. Long, "Theoretic studies of full constraints on a star tracker's influential error sources for in-orbit calibration," *Publications of the Astronomical Society of the Pacific*, vol. 128, no. 961, p. 035003, 2016.

[33] C. A. Lemon, M. W. Auger, R. G. McMahon, and S. E. Koposov, "Gravitationally lensed quasars in gaia: I. resolving small-separation lenses," *Monthly Notices of the Royal Astronomical Society*, vol. 472, no. 4, pp. 5023–5032, 2017.

[34] H. L. Van Trees, *Detection, estimation, and modulation theory.* John Wiley & Sons, 2004.

[35] A. Echeverria, J. F. Silva, R. A. Mendez, and M. Orchard, "Analysis of the bayesian cramér-rao lower bound in astrometry-studying the impact of prior information in the location of an object," *Astronomy & Astrophysics*, vol. 594, p. A111, 2016.

[36] D. Michalik, L. Lindegren, D. Hobbs, and A. G. Butkevich, "Gaia astrometry for stars with too few observations. a bayesian approach," *Astronomy & Astrophysics*, vol. 583, p. A68, 2015.

[37] D. Michalik and L. Lindegren, "Quasars can be used to verify the parallax zero-point of the tycho-gaia astrometric solution," *Astronomy & Astrophysics*, vol. 586, p. A26, 2016.

[38] H. So, Y. Chan, K. Ho, and Y. Chen, "Simple formulae for bias and mean square error computation [dsp tips and tricks]," *Signal Processing Magazine, IEEE*, vol. 30, no. 4, pp. 162–165, 2013.

[39] R. Ahlswede and I. Csiszár, "Hypothesis testing with communication constraints," *IEEE transactions on information theory*, vol. 32, no. 4, pp. 533–542, 1986.

[40] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, pp. 493–507, 1952.

[41] T. Han, "Hypothesis testing with multiterminal data compression," *IEEE transactions on information theory*, vol. 33, no. 6, pp. 759–772, 1987.

[42] V. Y. Tan *et al.*, "Asymptotic estimates in information theory with non-vanishing error probabilities," *Foundations and Trends® in Communications and Information Theory*, vol. 11, no. 1-2, pp. 1–184, 2014.

[43] V. Kostina, *Lossy data compression: nonasymptotic fundamental limits.* PhD thesis, Princeton University, 2013.

[44] I. Kontoyiannis and S. Verdú, "Optimal lossless data compression: Non-asymptotics and asymptotics," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 777–795, 2014.

[45] M. Hayashi, "Second-order asymptotics in fixed-length source coding and intrinsic randomness," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4619–4637, 2008.

[46] M. Kendall, A. Stuart, J. Ord, and S. Arnold, *Vol. 2A: Classical inference and the linear model.* London [etc.]: Arnold [etc.], 1999.

[47] S. M. Kay, "Fundamentals of statistical signal processing, volume i: Estimation theory (v. 1)," *PTR Prentice-Hall, Englewood Cliffs*, 1993.

[48] R. A. Bradley and J. J. Gart, "The asymptotic properties of ml estimators when sampling from associated populations," *Biometrika*, vol. 49, no. 1/2, pp. 205–214, 1962.

[49] B. Hoadley, "Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case," *The Annals of mathematical statistics*, pp. 1977–1991, 1971.

[50] J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography," *Image Processing, IEEE Transactions on*, vol. 5, no. 3, pp. 493–506, 1996.

[51] J. R. Janesick, *Scientific charge-coupled devices*, vol. 83. SPIE press, 2001.

[52] S. B. Howell, *Handbook of CCD astronomy*, vol. 5. Cambridge University Press, 2006.

[53] J. R. Janesick, *Photon transfer.* SPIE press San Jose, 2007.

[54] I. S. McLean, *Electronic imaging in astronomy: detectors and instrumentation.* Springer Science & Business Media, 2008.

[55] I. R. King, "The profile of a star image," *Publications of the Astronomical Society of the Pacific*, pp. 199–201, 1971.

[56] O. Bendinelli, G. Parmeggiani, A. Piccioni, and F. Zavatti, "The newton-gauss regularized method-application to point-spread-function determination in ccd frames," *The*

Astronomical Journal, vol. 94, pp. 1095–1100, 1987.

[57] R. A. Méndez, E. Costa, M. H. Pedreros, M. Moyano, M. Altmann, and C. Gallart, "Proper motions of local group dwarf spheroidal galaxies i: First ground-based results for fornax," *Publications of the Astronomical Society of the Pacific*, vol. 122, no. 893, pp. 853–875, 2010.

[58] F. R. Chromey, *To measure the sky: an introduction to observational astronomy*. Cambridge University Press, 2010.

[59] L. Lindegren, "A general maximum-likelihood algorithm for model fitting to ccd sample data," *Gaia DPAC public document GAIA-C3-TN-LU-LL-078 (November 2008)*, 2008.

[60] M. Gai, D. Busonero, and R. Cancelliere, "Performance of an algorithm for estimation of flux, background, and location on one-dimensional signals," *Publications of the Astronomical Society of the Pacific*, vol. 129, no. 975, p. 054502, 2017.

[61] M. Kendall, A. Stuart, K. J. Ord, and S. Arnold, "Kendall's advanced theory of statistics: Volume 2a–classical inference and and the linear model (kendall's library of statistics)," *A Hodder Arnold Publication,*, 1999.

[62] Y. Shkel, M. Raginsky, and S. Verdú, "Universal lossy compression under logarithmic loss," in *Information Theory (ISIT), 2017 IEEE International Symposium on*, pp. 1157–1161, IEEE, 2017.

[63] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," *arXiv preprint physics/0004057*, 2000.

[64] S. Watanabe, "Neyman-pearson test for zero-rate multiterminal hypothesis testing," in *Information Theory (ISIT), 2017 IEEE International Symposium on*, pp. 116–120, IEEE, 2017.

[65] K. L. Chung, *A course in probability theory*. Academic press, 2001.

[66] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[67] L. Devroye and G. Lugosi, *Combinatorial methods in density estimation*. Springer Science & Business Media, 2012.

[68] T. A. Courtade and T. Weissman, "Multiterminal source coding under logarithmic loss," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 740–761, 2014.

[69] T. Berger, "Rate-distortion theory," *Encyclopedia of Telecommunications*, 1971.

[70] Z. Zhang, E.-H. Yang, and V. K. Wei, "The redundancy of source coding with a fidelity criterion. 1. known statistics," *IEEE Transactions on Information Theory*, vol. 43, no. 1, pp. 71–91, 1997.

[71] V. Y. Korolev and I. G. Shevtsova, "On the upper bound for the absolute constant in the berry–esseen inequality," *Theory of Probability & Its Applications*, vol. 54, no. 4, pp. 638–658, 2010.