

# Marco Canducci

---

**Passaporto:** YB4145242 | **Data di nascita:** 29/10/1991 | **Luogo di nascita:** Rimini, Italia | **Nazionalità:** Italiana | **Sesso:** Maschile | **Telefono:** (+39) 3476817878 (Cellulare) | **Indirizzo e-mail:** [marco.canducci91@gmail.com](mailto:marco.canducci91@gmail.com) | **Indirizzo** 434 King Edwards Wharf, 25 Sheepcote Street, B16 8AB, Birmingham, Regno Unito (Abitazione)

## ● ESPERIENZA LAVORATIVA

---

15/07/2024 - ATTUALE - BIRMINGHAM, REGNO UNITO

### **RICERCATORE UNIVERSITARIO** UK RESEARCH AND INNOVATION

---

Sviluppo di modelli surrogati per la simulazione del processo di solidificazione direzionale di super leghe metalliche, con applicazione alla produzione di componenti di turbine di motori aeromobili. Ricerca avanzata in Computer Vision, Machine Learning, Artificial Intelligence in campi come l'apprendimento non supervisionato, supervisionato e l'auto-apprendimento

**Impresa o settore** Attività professionali, scientifiche e tecniche

15/01/2022 - 15/07/2024 - BIRMINGHAM, REGNO UNITO

### **RICERCATORE UNIVERSITARIO** NATIONAL INSTITUTE FOR HEALTH RESEARCH

---

Sviluppo di strumenti digitali per l'indagine di multi-morbilità in pazienti ricevuti in visite mediche regolare. Interpretazione delle sequenze in termini epidemiologici e contestualizzazione in appartenenza a demografie varie e gerarchiche.

Sviluppo di metodi per la rappresentazione visiva di combinazioni di medicazione per pazienti a rischio di vita in presenza di multi-morbilità complessa.

15/12/2020 - 15/01/2022 - BIRMINGHAM, REGNO UNITO

### **RICERCATORE UNIVERSITARIO** THE ALAN TURING INSTITUTE

---

Stima dell'importanza di "features" in data sets ad alta dimensionalità nell'indagine della progressione della malattia di Cushing, da dati metabolici. Utilizzo e implementazione di tecniche di classificazione e riduzione della dimensionalità con determinazione di tensore metrico discriminante. Formulazione di GMLVQ probabilistico per l'adattamento a data set contenenti errori di misura.

## ● ISTRUZIONE E FORMAZIONE

---

10/11/2017 - 13/07/2022 - BIRMINGHAM, REGNO UNITO

### **DOTTORATO IN COMPUTER SCIENCE-** UNIVERSITY OF BIRMINGHAM

---

**Campi di studio** Matematica e statistica non ulteriormente definite | **Livello EQF:** 8 | **Sito Internet:**

<https://www.birmingham.ac.uk/>

16/12/2013 - 10/03/2017 - BOLOGNA, ITALIA

### **LAUREA MAGISTRALE IN ASTROFISICA E ASTRONOMIA-** UNIVERSITÀ ALMA MATER STUDIORUM DI BOLOGNA

---

**Campi di studio** Scienze naturali, matematiche e statistiche non ulteriormente definite | **Livello EQF:** 7 | **Sito Internet:**

<https://www.unibo.it/it/homepage>

23/10/2010 - 12/12/2013 - BOLOGNA, ITALIA

*Marco Canducci*

● **COMPETENZE LINGUISTICHE**

---

Lingua madre: **ITALIANO**

---

	<b>COMPRENSIONE</b>		<b>ESPRESSIONE ORALE</b>		<b>SCRITTURA</b>
	Ascolto	Lettura	Produzione orale	Interazione orale	
<b>INGLESE</b>	C2	C2	C2	C2	C2

● **COMPETENZE**

---

Python (programmazione informatica) | MATLAB

● **CONFERENZE/SEMINARI**

---

24/06/2019 - 27/06/2019 Pasadena (CA)

**Astroinformatics** Poster

Link <http://131.215.198.54/ai19/talks/Canducci.pdf>

25/11/2021 - 27/11/2021 Manchester (UK)

**Intelligent Data Engineering and Automated Learning** Primo autore e relatore

28/06/2021 - 02/07/2021 Virtual

**The European Astronomical Society (EAS) Annual Meeting** Primo autore e oratore

27/06/2022 - 01/07/2022 Valencia (ES)

**The European Astronomical Society (EAS) Annual Meeting** co-autore

02/09/2022 - 04/09/2022 Zagreb (Croatia)

**ESE Young Endocrinologists and Scientists (EYES)** co-autore

14/11/2022 - 16/11/2022 Harrogate (UK)

**Society for Endocrinology BES** co-autore

10/07/2023 - 15/07/2023 Krakow (HR)

**The European Astronomical Society (EAS) Annual Meeting** co-autore

01/07/2024 - 05/07/2024 Padova (IT)

**The European Astronomical Society (EAS) Annual Meeting** co-autore

09/09/2024 - 10/09/2024 Manchester (UK)

**Pioneering AI in MLTC: Bridging and Practice Conference** poster

16/06/2025 - 20/06/2025 Bologna

**STARS II: Current Challenges, upcoming solutions** Relatore

● **FORMAZIONE AGGIUNTIVA**

---

21/01/2018 - 26/01/2018

*Marco Canducci*

## **Applications of Computer Science Techniques in Galaxy Science**

---

08/06/2018 - 09/06/2018

**SUNDIAL Training School on Professional Software Writing and Project Management**

---

01/10/2018 - 06/10/2018

**Astronomical observing school**

---

04/11/2018 - 08/11/2018

**XXX Canary Islands Winter School of Astrophysics: "Big Data in Astronomy"**

---

11/03/2019 - 14/03/2019

**Data mining astronomical archives and the Virtual Observatory**

---

30/05/2019 - 31/05/2019

**Knowledge Clip Training**

---

### ● **INSEGNAMENTO**

---

07/09/2019 - 15/01/2020

**Mathematical Modelling and Decision Making**

---

University of Birmingham

Module Leader: Prof. Peter Tino

01/09/2020 - 13/01/2026

**Mathematical Foundations of Artificial Intelligence and Machine Learning**

---

University of Birmingham

A.A. 2020/2021, 2021/2022, 2022/2023, 2023/2024, 2024/2025, 2025/2026

Module Leader: Prof. Peter Tino

04/09/2019 - 10/01/2020

**Mathematical Foundations of Computer Science**

---

University of Birmingham

Module Leader: Prof. Achim Jung

01/09/2020 - 04/01/2023

**Algorithms for Data Science**

---

University of Birmingham

Module Leader: Prof. Miqing Li

24/02/2025 - 28/02/2025

**AstroInformatics**

---

EDUCADO Doctoral training school

### ● **ORATORE INVITATO**

---

15/03/2023 - 15/03/2023

**Università Alma Mater Studiorum di Bologna**

---

*Marco Cantini*

Dipartimento di Fisica e Astronomia

Seminar Series

13/01/2025 - 17/01/2025

**Instituto de Astrofísica de Canarias (IAC)**

---

EDUCADO annual and mid-term meeting

11/02/2025 - 11/02/2025

**University of Birmingham**

---

Department of Mathematics, Statistics and Data Science Seminar

04/12/2025 - 04/12/2025

**INAF - Osservatorio astronomico di Padova**

---

Seinar series

## ● **SUPERVISIONE**

---

### **Co-supervisione Studenti di dottorato**

---

Othman Alghamdi (start year 2023)

Meerah Al-Hakbani (start year 2023)

Dario Barone (start year 2024)

### **Co-supervisione Master**

---

Steni Sebastian: "Estimation of Gravitational Wave Compact Binary Population Parameters using Probabilistic Hough Transform." (2020)

### **Co-supervisione Bachelor**

---

Christopher Callum Nailer: "Identifying Filament Structures in an Astronomical Dataset using Probabilistic Hough Transform" (2022)

Jiahao Gai: "Multi-objective optimization in the simplification for multi-morbidity medical records." (2022)

## ● **REVIEWER**

---

### **IEEE Transactions on Pattern Analysis and Machine Intelligence**

---

Link <https://dl.acm.org/journal/itpm>

### **Neural Networks**

---

Link <https://www.sciencedirect.com/journal/neural-networks>

### **Pattern Recognition**

---

Link <https://www.sciencedirect.com/journal/pattern-recognition>

### **AAAI conference**

---

Link <https://aaai.org/conference/aaai/aaai-26/>

### **Astronomy & Astrophysics**

---

Link <https://www.aanda.org/>

Marco Cantucci

## **KDD conference**

Link <https://www.kdd.org/>

- **TERZA MISSIONE**

---

01/07/2024 - 05/07/2024

**Summer School - University of Birmingham**

---

14/07/2025 - 15/07/2025

**Summer School - University of Birmingham**

---

24/08/2020 - 24/08/2024

**Intervista OggiScienza**

---

Intervistatrice: Luisa Alessio

**Link**

[https://oggiscienza.it/2020/08/24/big-data-astronomici-per-lo-studio-delle-galassie-medusa/?fbclid=IwAR1eNMAYM7dznTez0l4ZrPBp02fCrM9Nlj7fy\\_tAM7dBcgELQmqiaHwFCCQ](https://oggiscienza.it/2020/08/24/big-data-astronomici-per-lo-studio-delle-galassie-medusa/?fbclid=IwAR1eNMAYM7dznTez0l4ZrPBp02fCrM9Nlj7fy_tAM7dBcgELQmqiaHwFCCQ)

19/09/2018 - 19/09/2018

**"Meet the Experts"**

---

ThinkTank Museum, Birmingham

- **DICHIARAZIONE DI VERIDICITÀ**

---

Le informazioni contenute nel presente 'curriculum vitae et studiorum' sono rese sotto la personale responsabilità del sottoscritto, ai sensi degli articoli 46 e 47 del Decreto del Presidente della Repubblica 28 dicembre 2000, numero 445, e successive modifiche e integrazioni, consapevole della responsabilità penale prevista dall'articolo 76 del medesimo Decreto per le ipotesi di falsità in atti e/o dichiarazioni mendaci

---

*Marco Cantucci*

## 1. Profilo scientifico e linea di ricerca

La mia attività scientifica si colloca all'intersezione tra astrofisica computazionale, modellazione probabilistica e sviluppo di metodologie geometriche per l'analisi di dataset astronomici ad alta dimensionalità e in regime di rumore elevato. Il filo conduttore del mio percorso è lo sviluppo di strumenti metodologici robusti per l'identificazione, l'estrazione e la caratterizzazione di strutture deboli (in particolare strutture filamentarie e manifolds – varietà topologiche) all'interno di distribuzioni complesse di dati osservativi o simulativi.

La mia formazione in Astrofisica presso l'Università di Bologna mi ha fornito solide basi in evoluzione stellare, modelli di popolazioni stellari, dinamica galattica e interpretazione di diagrammi colore–magnitudine. Ho poi conseguito il dottorato in Computer Science presso la University of Birmingham (UK), sviluppando un programma di ricerca originale in ambito probabilistico e geometrico applicato a problemi astrofisici. Questa formazione mi ha permesso di costruire un profilo scientifico orientato alla progettazione e formalizzazione di nuovi metodi per l'analisi, interpretabili e informati dalle conoscenze tecnico-scientifiche d'applicazione. In quanto parte del MSCA-ITN-ETN SUNDIAL, durante il dottorato ho avuto la possibilità di stringere rapporti internazionali con diversi enti Europei e raffinare competenze di ricerca di base e avanzate attraverso le scuole di formazione in astronomia e Computer Science.

Il mio profilo si distingue per una specializzazione metodologica profonda e formalmente strutturata nello sviluppo di algoritmi probabilistici per la scoperta automatica di strutture coerenti in condizioni di segnale debole e rumore elevato. Tale competenza si colloca in modo naturale nel contesto delle grandi survey contemporanee e future, e in particolare nel quadro scientifico di LSST, dove urgono competenze nella separazione probabilistica tra sottostrutture e popolazione di campo e l'analisi di distribuzioni stellari in spazio delle fasi e in spazio fotometrico multibanda.

## 2. Sviluppo metodologico originale

### 2.1 Modellazione probabilistica di manifolds in ambienti rumorosi

Durante il dottorato ho sviluppato un framework probabilistico per il recupero e l'analisi di manifolds (varietà e sottocarietà topologiche) immerse in dataset rumorosi e potenzialmente multi-strutturati. Questo lavoro ha portato alla definizione del metodo 1-DREAM (1D Recovery, Extraction and Analysis of Manifolds in noisy environments) e a una formulazione generale per la modellazione probabilistica di dataset multi-manifold con rumore eterogeneo.

In tali lavori ho contribuito direttamente:

- alla formulazione teorica del modello;
- alla definizione del formalismo probabilistico;
- all'implementazione algoritmica in più linguaggi di programmazione (MATLAB, python, C);
- alla validazione su dataset simulativi e reali;
- al mantenimento del codice con enfasi sul version control tramite piattaforme dedicate (github/gitLab)

Il contributo principale consiste nell'aver costruito un impianto metodologico capace di distinguere strutture geometriche reali da fluttuazioni stocastiche in regimi ad alta contaminazione. La modellazione probabilistica di dataset multi-manifold fornisce uno strumento generale per trattare simultaneamente più componenti strutturate immerse in un fondo diffuso, condizione tipica dei campi stellari galattici.

*Marco Cantucci*

Parallelamente, ho lavorato su tecniche di apprendimento discriminativo in spazi ad alta dimensionalità, sviluppando approcci probabilistici per la stima della rilevanza delle feature e per la riduzione dimensionale informata da errori di misura. Questa esperienza si rivela cruciale nell'analisi di dataset astronomici complessi, dove incertezze osservative e covarianze non trascurabili giocano un ruolo determinante.

Infine, l'introduzione dell'errore di misura e l'incertezza sui dati è fonte di incertezza sul risultato. Risulta naturale, attraverso l'utilizzo di tecniche probabilistiche, l'inclusione di tali misure nei modelli ideati. Tale inclusione può essere consistentemente trattata per avere risultati robusti sul processo di inferenza. Questo filone di ricerca è in sviluppo nei lavori di dottorato di cui sono co-supervisore. In questi lavori, la modellazione topologica utilizza il formalismo della trasformata di Hough, reinterpretata per l'assunzione di un modello generativo. I dati vengono utilizzati come evidenze incerte sul supporto del modello.

## 2.2 Applicazione a stellar streams e strutture filamentarie

Le metodologie sviluppate sono state applicate a diversi contesti astrofisici:

- caratterizzazione della struttura Fornax-Eridanus e delle connessioni filamentarie tra ammassi;
- studio di filamenti associati a cluster di galassie e loro ruolo nell'evoluzione ambientale;
- identificazione e analisi morfologica di stellar streams (es. Jhelum,  $\omega$ -Centauri);
- rilevazione e caratterizzazione di cavità e strutture nel mezzo interstellare simulato.

In ciascun caso il mio contributo ha riguardato sia lo sviluppo dell'algoritmo sia la sua applicazione e interpretazione fisica, con un'attenzione particolare alla traduzione dei risultati metodologici in insight astrofisici quantitativi. Questa integrazione tra formalismo matematico e interpretazione fisica costituisce un elemento distintivo del mio percorso.

Nei lavori dedicati all'identificazione di strutture filamentarie tramite versioni probabilistiche e iterative della Hough Transform e tramite approcci ispirati all'intelligenza collettiva (es. applicazione a Jhelum e a strutture filamentarie su larga scala), ho contribuito allo sviluppo di metodi capaci di:

- rilevare strutture localmente lineari in presenza di forte rumore,
- operare in regimi di contaminazione elevata,
- quantificare la significatività statistica della rilevazione.

Sebbene tali metodologie siano state applicate anche a strutture su larga scala, la loro natura è intrinsecamente scalabile e direttamente trasferibile allo studio di:

- code mareali di ammassi globulari,
- code mareali di ammassi aperti,
- strutture di dissoluzione in open clusters dinamicamente evoluti.

Gli ammassi aperti, in particolare, rappresentano sistemi più fragili e dinamicamente meno legati rispetto ai globulari, e quindi più suscettibili a stripping mareale e dispersione nel disco galattico. L'identificazione delle loro code e delle sottostrutture associate richiede strumenti sensibili a segnali deboli e a morfologie non ideali: esattamente il regime in cui le metodologie da me sviluppate sono state progettate per operare.

Attualmente sto conducendo uno studio sistematico delle regioni circostanti ammassi globulari utilizzando dati astrometrici di Gaia, con l'obiettivo di identificare e caratterizzare eventuali

*Marco Cantucci*

strutture mareali deboli. A tal fine sto applicando e adattando in modo estensivo i modelli sviluppati in precedenza (in particolare approcci basati su Hough Transform probabilistica e strategie ispirate all'ant-colony optimization) per costruire una procedura uniforme di ricerca delle code mareali attorno all'intera popolazione di ammassi noti. Questo approccio risulta in una stima probabilistica dell'appartenza di ciascuna stella alla struttura statisticamente rilevata (membership probabilistica).

La modifica delle metodologie già descritte permette di inserire l'importanza di ciascuna stella nel processo di voto. Questa è stimata dalla compatibilità con il diagramma colore-magnitudine dell'ammasso globulare. L'obiettivo è definire un framework metodologico coerente e statisticamente robusto, in grado di fornire stime quantitative della significatività delle strutture rilevate e di essere scalato, in prospettiva, ai dataset più profondi e complessi che saranno disponibili con LSST.

### **3. Analisi esplorativa di dati**

Un elemento trasversale del mio percorso scientifico è la capacità di affrontare dataset complessi in modo esplorativo ma formalmente rigoroso, con l'obiettivo di identificare strutture latenti, sottospazi informativi e regolarità non immediatamente evidenti. Questa attitudine si è tradotta in una linea di ricerca coerente, documentata dalle pubblicazioni e rafforzata da una formazione interdisciplinare che integra astrofisica, modellazione probabilistica e machine learning avanzato.

Più recentemente, ho esteso questo approccio allo studio di rappresentazioni apprese tramite modelli neurali strutturati, in particolare attraverso l'utilizzo di Graph Attention AutoEncoder per l'estrazione di rappresentazioni latenti in dati organizzati su grafi. In questo contesto, l'attenzione è rivolta alla capacità del modello di apprendere embedding interpretabili, capaci di preservare relazioni topologiche e strutturali tra oggetti. Questo tipo di approccio è direttamente rilevante per l'analisi di sistemi stellari, dove relazioni di prossimità, coerenza dinamica o similitudine fotometrica possono essere naturalmente rappresentate come grafi.

Parallelamente, sto conducendo un'attività esplorativa nell'ambito della Computer Vision applicata a immagini industriali, con l'obiettivo di stimare sottospazi latenti che distinguano in modo robusto le proprietà principali delle immagini (texture, morfologia, difetti strutturali). Sebbene tale linea sia attualmente in fase preliminare e non ancora oggetto di pubblicazioni, essa rappresenta un'estensione naturale del mio interesse per l'identificazione di sottospazi informativi rilevanti e per la separazione tra segnale strutturato e variazione non significativa.

Nel complesso, queste attività testimoniano una competenza consolidata nell'analisi esplorativa avanzata dei dati, intesa non come fase preliminare puramente descrittiva, ma come processo strutturato di costruzione di rappresentazioni interpretabili e statisticamente fondate.

L'interpretazione dei dati attraverso uno spazio latente che ne catturi i processi generativi fondamentali è anche alla base del progetto con cui ho partecipato al recente bando MSCA-European Postdoctoral Fellowship 2025. Il progetto propone l'omogeneizzazione dello spazio dei parametri delle fotosfere stellari stimati da survey multiple attraverso tecniche probabilistiche. Queste coinvolgono la rappresentazione latente di processi stocastici condivisi. Il progetto ha ricevuto una valutazione di 90/100 ma non è stato finanziato per budget insufficiente.

*Marco Cantucci*

#### **4. Esperienza in ambienti interdisciplinari e sviluppo software**

La mia attività presso UK Research and Innovation, il National Institute for Health Research e il Turing Institute mi ha permesso di consolidare competenze nella gestione di dataset complessi, nella progettazione di strumenti digitali e nello sviluppo di modelli scalabili.

Sebbene parte di tale esperienza sia maturata in ambito biomedico e industriale, il nucleo metodologico – modellazione probabilistica, gestione del rumore, apprendimento su dataset ad alta dimensionalità, sviluppo di strumenti interpretabili – è direttamente trasferibile all’astrofisica delle grandi survey.

Ho inoltre maturato esperienza:

- nella supervisione di studenti di dottorato e master su tematiche astrofisiche computazionali;
- nell’insegnamento continuativo di fondamenti matematici dell’IA e del machine learning;
- nella revisione per riviste di primo piano in ambito sia astrofisico sia metodologico;
- nella collaborazione con diversi gruppi internazionali (tra cui “Millennium Nucleus for Galaxies-MINGAL”, Valparaíso, Chile; “Kaptein Astronomical Institute”, Groningen, NL; “Instituto de Astrofisica de Canarias-IAC”; “University of Ghent”, BE; “EDUCADO” ITN).

Questi elementi testimoniano maturità scientifica, autonomia e capacità di contribuire alla crescita metodologica di un gruppo di ricerca.

---

#### **5. Prospettive di sviluppo**

Nei prossimi anni intendo sviluppare:

- framework probabilistici scalabili per la rilevazione di filamenti deboli in survey profonde;
- metodi per la classificazione automatica dello stato dinamico di ammassi;
- strumenti interpretabili per la quantificazione dell’incertezza nelle strutture identificate;
- integrazione tra modellazione geometrica e informazioni multi-banda.

L’obiettivo è contribuire alla costruzione di strumenti robusti che permettano di massimizzare il ritorno scientifico dei dati LSST, combinando rigore metodologico e interpretazione astrofisica.

*Marco Cantucci*