MAGIC: Multi-Agent Generative Intention Coordination

David Huk, Oliver Hamelijnck, Dimitris Demiris, Theodoros Damoulas

Unilink Excellence Through Innovation · UNIVERSITY OF WARWICK

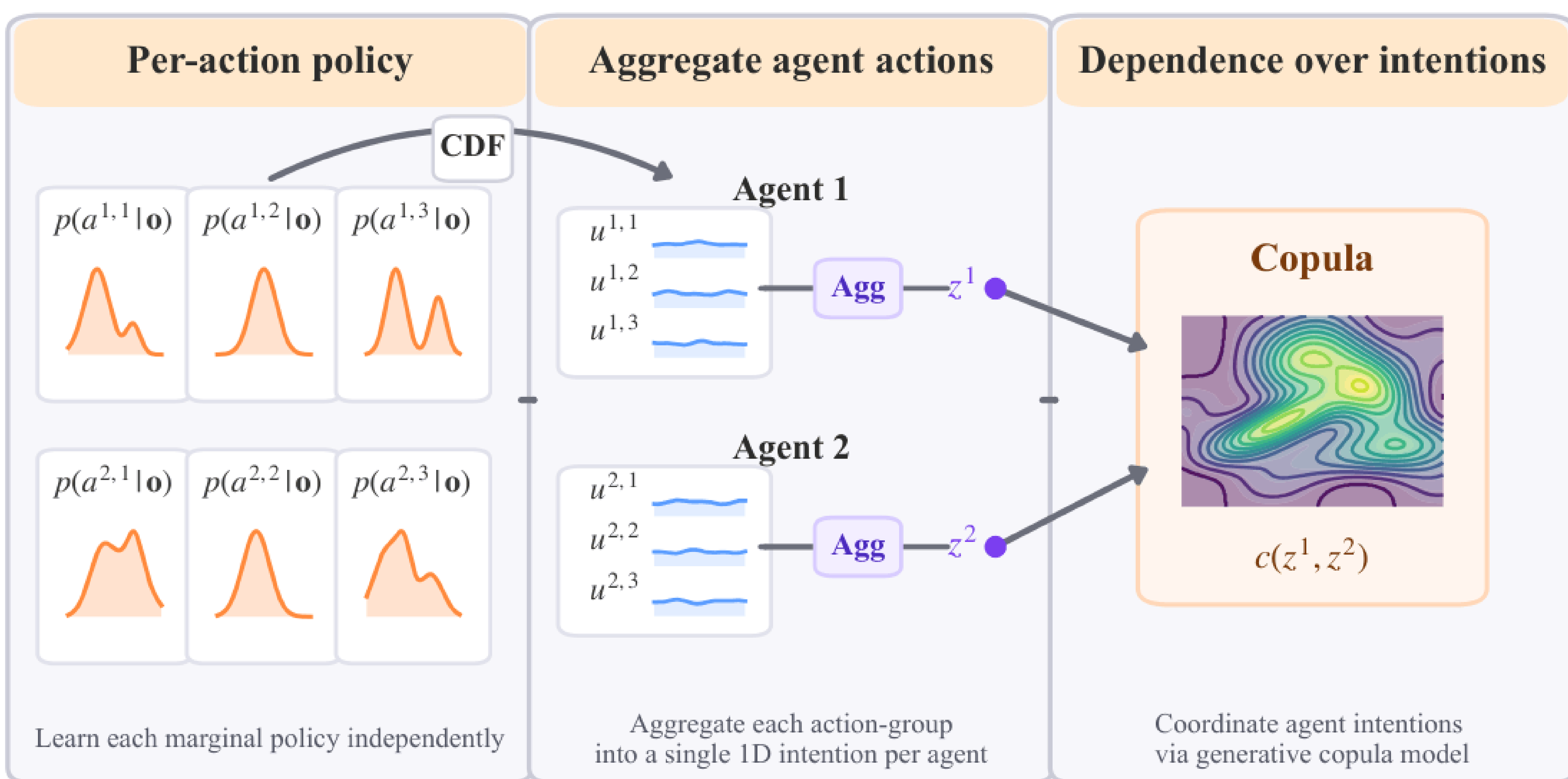
Motivation

- Multi-Agent Imitation Learning (**MAIL**) aims to learn coordinated policies from demonstrations
- But learning dependent policies is often infeasible

Goal: **How can we model coordination among many agents in a scalable way?**

MAGIC follows a **divide-and-conquer** strategy:

1. We learn lightweight **independent** policies.
2. We compress each agent's action distribution into a one-dimensional latent **intention**.
3. We learn a **dependent** generative model over intentions



1. Independent policies

Fit independent conditional densities to each of action.

Independent policy: $p(\boldsymbol{a}_t \mid \boldsymbol{o}_t) \approx \prod_{i=1}^{n} \prod_{j=1}^{d} \widehat{p}^{i,j}(a_t^{i,j} \mid \boldsymbol{o}_t^i)$

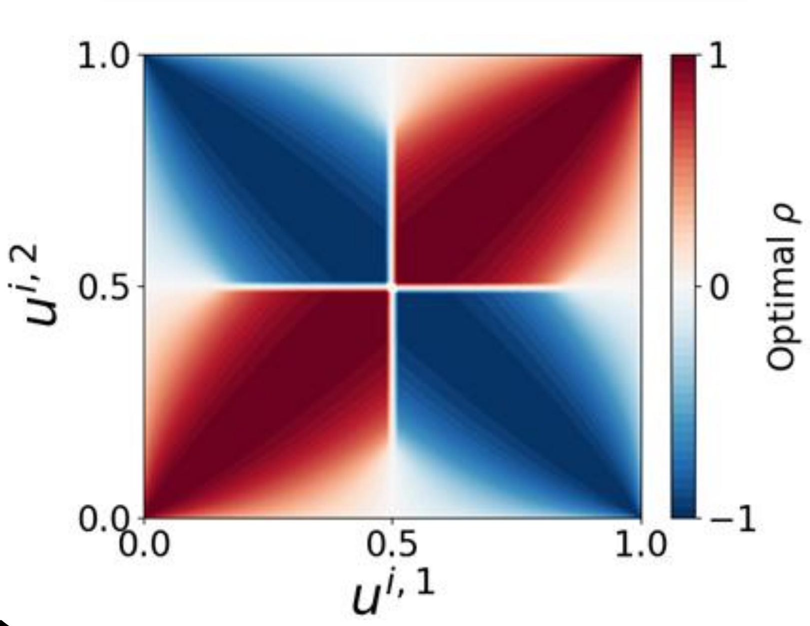
2. Per-agent intentions

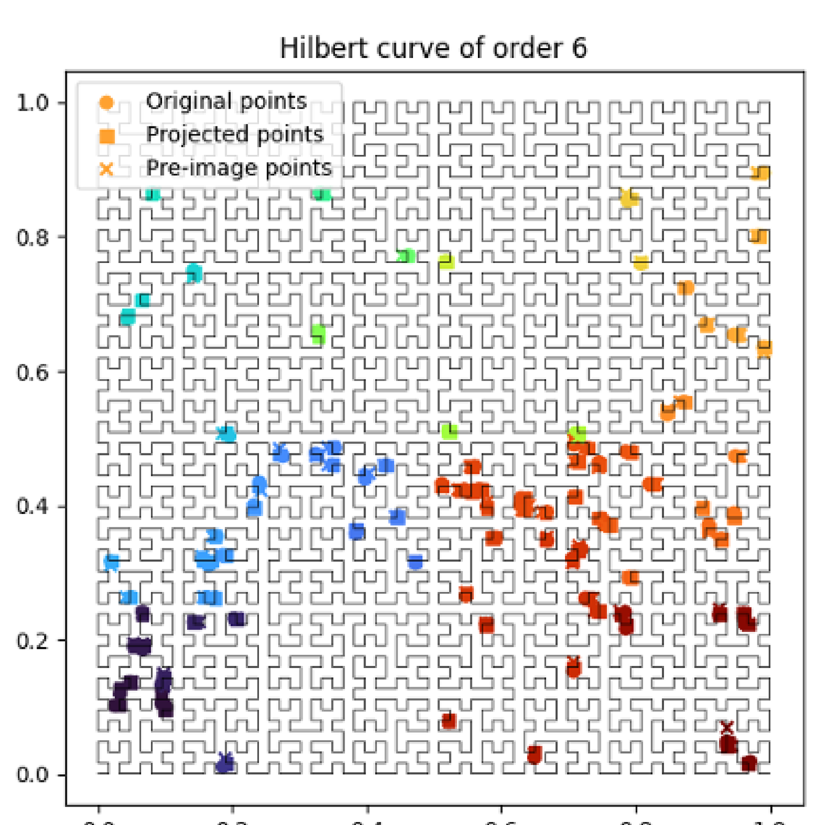
Use 1. to obtain per-action **quantiles**:

$$\boldsymbol{u}^i = (\widehat{P}^{i,1}(a_t^{i,1} \mid \boldsymbol{o}_t^i), \ldots, \widehat{P}^{i,d}(a_t^{i,d} \mid \boldsymbol{o}_t^i)) \in [0,1]^d$$

Compress the quantiles via **intention function**:

$$Z : [0,1]^d \to [0,1]$$





Copula background

In the case of independent data x_1, x_2 we have that their joint distribution factorises as:

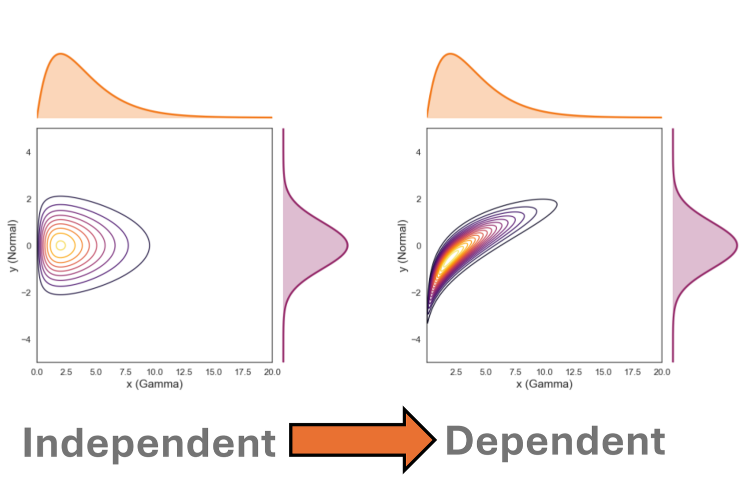
$$f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$$

Now assume that x_1, x_2 are not independent. We then have:

$$f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2) \cdot c(x_1, x_2)$$

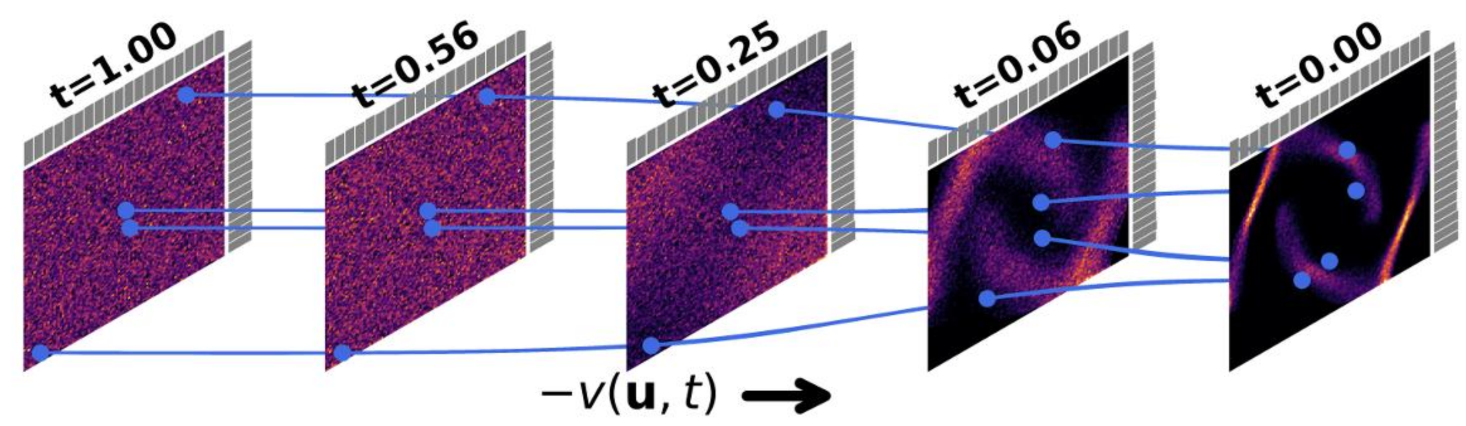
Sklar's Theorem

$$p(\mathbf{x}) = \prod_{i=1}^{d} \underbrace{\{p^i(x^i)\}}_{\text{Product of marginals}} \cdot \underbrace{\mathbf{c}(P^1(x^1), \ldots, P^d(x^d))}_{\text{Copula = dependence}}$$



3. Dependence model

Coordinate intentions via a **generative copula model**.



Sample from the true intention copula via flow-matching.

Theory on MAIL

MAIL strictly in terms the dependence between agents.

Theorem 3.1 (Joint policy copula decomposition). *Let $\boldsymbol{p}(\boldsymbol{a} \mid \boldsymbol{o})$ be a $(n \cdot d)$-dimensional joint policy function with absolutely continuous marginal policies $p^{i,j}(a^{i,j} \mid \boldsymbol{o})$ for $i = 1, \ldots, n$, $j = 1, \ldots, d$. Then there exists a copula with density function $\mathbf{c}(. \mid \boldsymbol{o}) : [0,1]^{n \cdot d} \times \mathbb{R}^r \to \mathbb{R}^+$ such that $\forall \boldsymbol{a} \in \mathbb{R}^{n \cdot d}$:*

$$\boldsymbol{p}(\boldsymbol{a} \mid \boldsymbol{o}) = \prod^{n} \prod^{d} \{p^{i,j}(a^{i,j} \mid \boldsymbol{o})\} \cdot \mathbf{c}(P^{1,1}(a^{1,1} \mid \boldsymbol{o}), \ldots, P^{n,d}(a^{n,d} \mid \boldsymbol{o}) \mid \boldsymbol{o}) \quad (5)$$

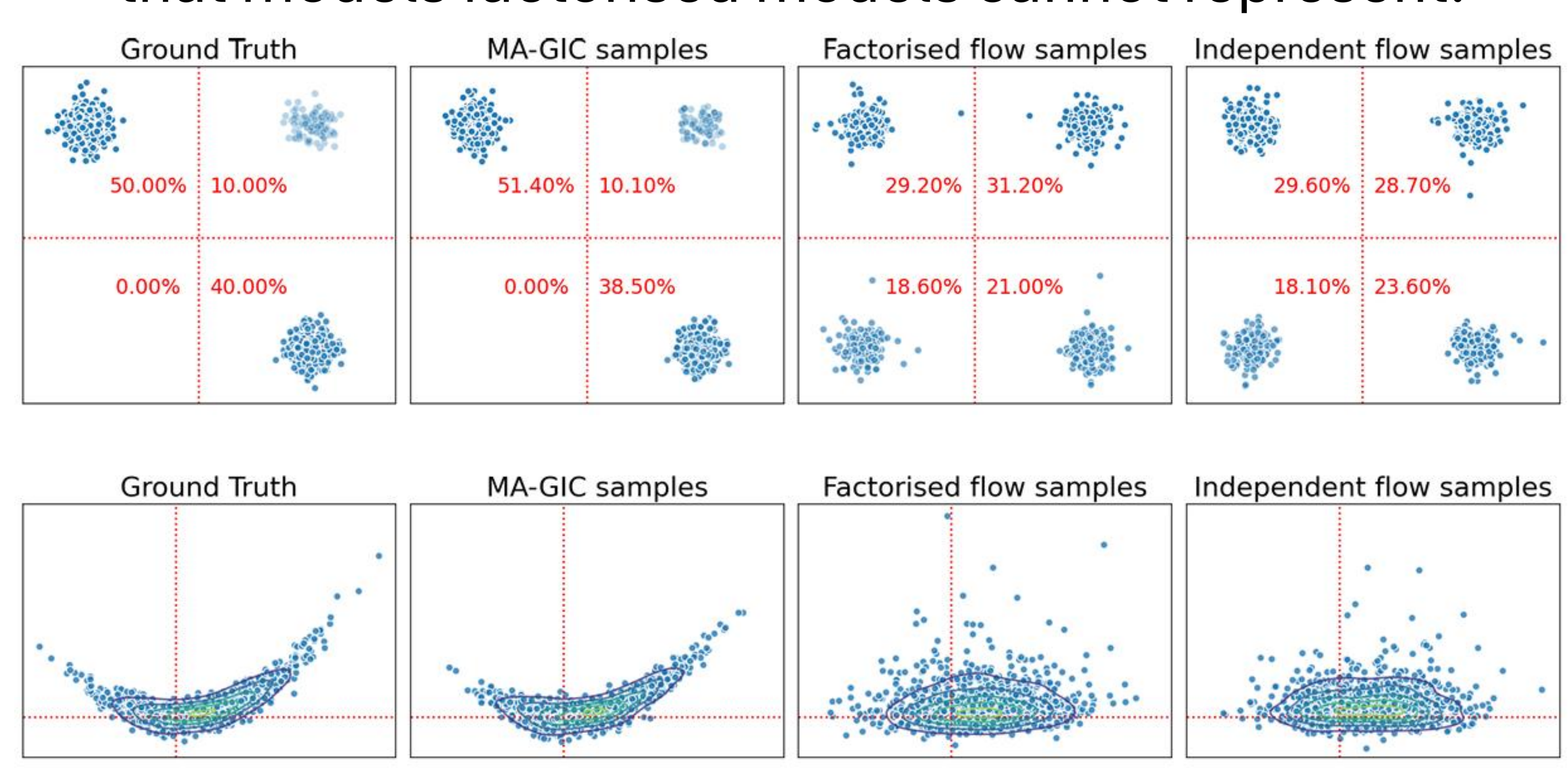
where $P^{i,j}$ are the cumulative distribution functions (CDFs) of marginal policies.

Corollary 3.2. *Under Assumptions 1 and 2, for $\boldsymbol{u} = (\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n) \in [0,1]^{n \cdot d}$, the following holds:*

1. *$\boldsymbol{p}(\boldsymbol{a}|\boldsymbol{o})$ admits an independent (2) policy $\iff$ $\mathbf{c}(\boldsymbol{u}) = 1$ a.e. in $\boldsymbol{u}$.*
2. *$\boldsymbol{p}(\boldsymbol{a}|\boldsymbol{o})$ admits a factorised (3) or projected (4) policy $\iff$ $\mathbf{c}(\boldsymbol{u}) = \prod_{i=1}^{n} c^i(\boldsymbol{u}^i)$ a.e. in $\boldsymbol{u}$.*
3. *$\boldsymbol{p}(\boldsymbol{a}|\boldsymbol{o})$ always admits a copula policy representation (6) [1].*

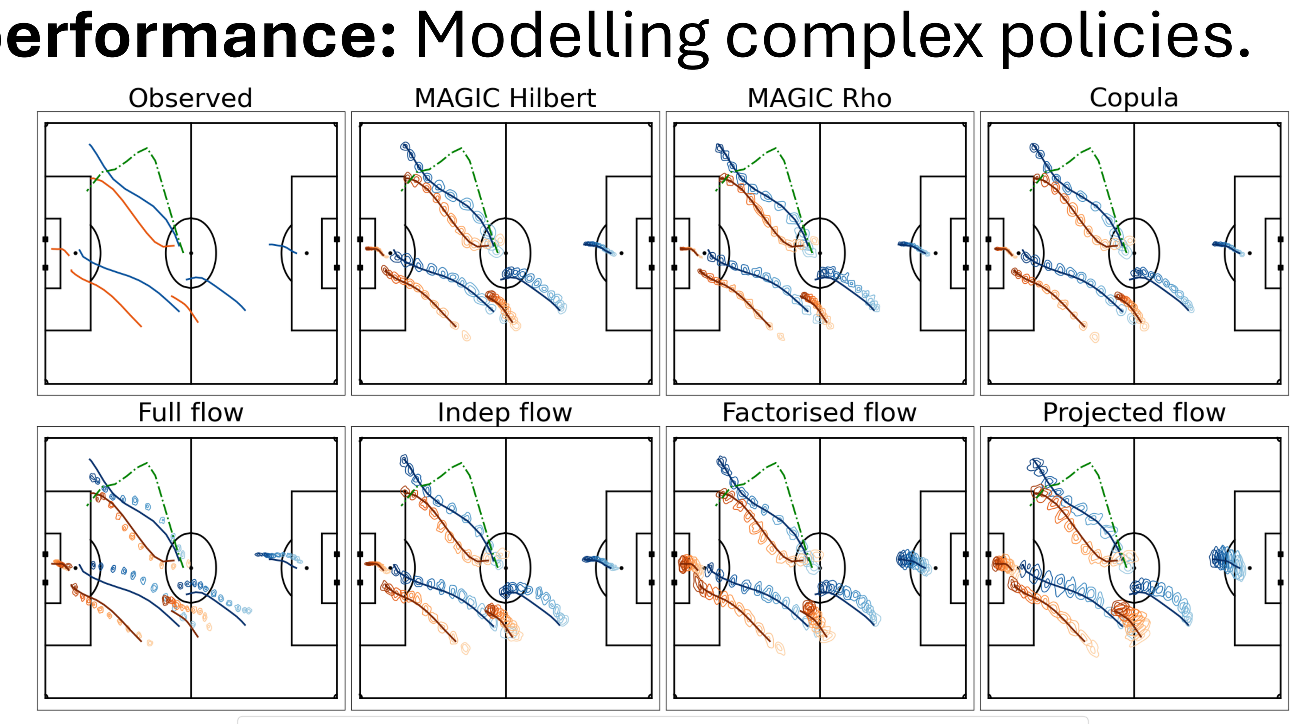
Results

✓ **Policy flexibility:** MAGIC learn dependent policies that models factorised models cannot represent.



✓ **Imitation performance:** Modelling complex policies.

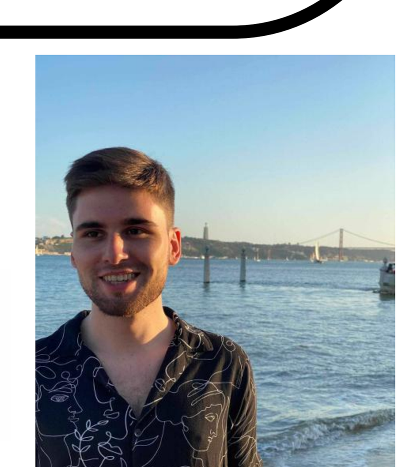
Trajectories of 44D movements of players in two football teams.



Model	Overall RMSE	Overall MAE	Overall Energy Score (44D)	Per-Player Mean RMSE	Per-Player Mean MAE	Per-Player Mean Energy Score
MAGIC Hilbert	$0.4271_{\pm 0.15}$	$0.2997_{\pm 0.10}$	$2.8953_{\pm 0.33}$	$0.5845_{\pm 0.15}$	$0.4782_{\pm 0.11}$	$0.4814_{\pm 0.09}$
MAGIC ρ	$\mathbf{0.3928}_{\pm 0.14}$	$\mathbf{0.2690}_{\pm 0.09}$	$\mathbf{2.5176}_{\pm 0.32}$	$\mathbf{0.5354}_{\pm 0.15}$	$\mathbf{0.4310}_{\pm 0.10}$	$0.4359_{\pm 0.08}$
Copula	$0.3951_{\pm 0.14}$	$0.2720_{\pm 0.09}$	$2.6217_{\pm 0.30}$	$0.5387_{\pm 0.15}$	$0.4351_{\pm 0.10}$	$\mathbf{0.4341}_{\pm 0.08}$
Full flow	$5.6424_{\pm 2.31}$	$2.5868_{\pm 0.69}$	$19.0765_{\pm 30.08}$	$7.7923_{\pm 1.72}$	$4.1062_{\pm 0.53}$	$3.6326_{\pm 0.51}$
Indep flow	$0.6494_{\pm 0.20}$	$0.4638_{\pm 0.14}$	$3.2245_{\pm 0.65}$	$0.8847_{\pm 0.25}$	$0.7349_{\pm 0.19}$	$0.5958_{\pm 0.12}$
Factorised flow	$0.6453_{\pm 0.18}$	$0.4676_{\pm 0.14}$	$3.5609_{\pm 0.51}$	$0.8857_{\pm 0.22}$	$0.7418_{\pm 0.18}$	$0.6627_{\pm 0.11}$
Projected flow	$0.7316_{\pm 0.21}$	$0.5274_{\pm 0.16}$	$4.2080_{\pm 0.62}$	$1.0029_{\pm 0.25}$	$0.8330_{\pm 0.21}$	$0.7736_{\pm 0.15}$

Paper and more!

Have a Question? Ask me!



https://warwick.ac.uk/fac/sci/statistics/staff/research_students/huk/