Unveiling the star formation history of the Upper Scorpius association through its kinematics

Vito Squicciarini [ID],[1,2]★ Raffaele Gratton,[2] Mariangela Bonavita [ID][2,3,4] and Dino Mesa[2]

[1]*Department of Physics and Astronomy 'Galileo Galilei', University of Padova, Via dell'Osservatorio 3, I-35122 Padova, Italy*
[2]*INAF – Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, I-35122 Padova, Italy*
[3]*Scottish Universities Physics Alliance (SUPA), Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK*
[4]*Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK*

ABSTRACT

Stellar associations can be discerned as overdensities of sources not only in the physical space but also in the velocity space. The common motion of their members, gradually eroded by the galactic tidal field, is partially reminiscent of the initial kinematic structure. Using recent data from *Gaia* EDR3, combined with radial velocities from GALAH and APOGEE, we traced back the present positions of stars belonging to Upper Scorpius (USCO), a subgroup of Scorpius–Centaurus, the nearest OB association. About one half of the subgroup (the 'clustered' population) appears composed of many smaller entities, which were in a more compact configuration in the past. The presence of a kinematic duality is reflected into an age spread between this younger clustered population and an older diffuse population, in turn confirmed by a different fraction f_D of disc-bearing stars ($f_D = 0.24 \pm 0.02$ versus $f_D = 0.10 \pm 0.01$). Star formation in USCO appears to have lasted more than 10 Myr and proceeded in small groups that, after a few Myr, dissolve in the field of the older population but retain for some time memory of their initial structure. The difference of ages inferred through isochrones and kinematics, in this regard, could provide a powerful tool to quantify the time-scale of gas removal.

Key words: stars: formation – stars: kinematics and dynamics – open clusters and associations: individual: Upper Scorpius.

1 INTRODUCTION

The discipline of star formation, in all of its aspects, has always been of paramount importance in astrophysics: Lying at the interception of several fields of studies, it stretches from the grandest scale of galactic evolution, passing through the intricate web of the interstellar medium, to the dusty scenery of primordial discs, where stellar furnaces mildly begin to shimmer and glow. The presence of active stellar formation in our Galaxy (Evans 1999) provides the opportunity to gaze at the process while it unfolds before our eyes: During the last years, several surveys have been performed to study the environment of galactic star-forming regions (e.g. Churchwell et al. 2009), the formation of multiple systems (e.g. Kraus et al. 2011), or even to directly test the predictions from models of planetary formation (e.g. Janson et al. 2021).

It has long been known (Ambartsumian 1954) that, following the collapse and fragmentation of gigantic structures called *molecular clouds*, a plethora of stars ($N = 10$–10^5) begins to form; initially concealed by the same dusty envelope which they are born from, they rapidly (2–7 Myr; Kim et al. 2021) divest themselves of it by means of harsh stellar winds and ionizing radiation, mostly originating from massive stars; HII regions, the impressive product of the irremediable alteration of the original cloud, are ruthlessly sculpted by the injection of energy and momentum from exploding supernovae (Barnes et al.

2020); after just a few Myr, the region is virtually devoid of its original gas reservoir (Kroupa, Aarseth & Hurley 2001). The abrupt change within stellar natal environment is proven by observations showing that, while at $t < 5$ Myr stars are often still embedded in their parent cloud, after 10 Myr only ~ 10 per cent of stars are found in bound clusters (Lada & Lada 2003). What we witness as an *association* for just – in cosmic terms – the blink of an eye (10–100 Myr; Moraux 2016), is therefore a young system, still dwelt by bright ephemeral OB stars and regions of active stellar formation (Kroupa et al. 2001).

The kinematic signature of members of clusters and associations, first recognized in the Hyades cluster and in Ursa Major already in the nineteenth century (Proctor 1869), is slowly eroded as the galactic differential rotation and tides spread the stars, turning them into moving groups or streams (Larson 2002). Given that the observed densities of associations are too low to give rise to significant close encounters and scatterings, their initial velocity structure can largely be conserved over the time-scale of several Myr (Wright & Mamajek 2018), as the above-mentioned perturbation induced by the galactic tidal field is expected to begin dominating on time-scales of $\sim 10^7$ yr (Wright & Mamajek 2018). If this is the case, we might think to use our present knowledge of an association to delve into its past. The first attempt in this direction was done by Blaauw (1946), Blaauw (1964), who devised a simple linear expansion model, where all the sibling stars move away at a constant pace from their natal position. By tracing back their motion, it is, in principle, possible to obtain an estimate of the association age in a way that is independent of stellar evolution models. However, a quantitative assessment of the

★ E-mail: vito.squicciarini@inaf.it

expansion model has long been considered elusive due to difficulties, on the one hand, in distinguishing real members from interlopers and, on the other hand, to obtain precise measurements of stellar distances and motions. This is why the idea has largely been shelved for decades (Brown, Dekker & de Zeeuw 1997), even though the notion of OB associations as the inflated outcome of compact clusters kept being popular (Lada & Lada 2003).

Ultraprecise astrometry from *Gaia* (Gaia Collaboration et al. 2016), covering, in its last release, almost two billion sources all over the sky, providing reliable optical photometry up to $G \approx 21$ mag and, most importantly, supplying the astronomical community with astrometry and proper motions of unprecedented accuracy, has been revolutionizing our knowledge of the Galaxy. An extensive search for new members of the known associations has been undertaken, reaching, for the nearest ones, the hydrogen-burning limit (Gagné & Faherty 2018), and opening up a kaleidoscope of opportunities for kinematic studies. Especially when combined with radial velocities (RVs) from external catalogues, *Gaia* is able to delve deeper than ever into the core of association architectures, unearthing exquisite fragments of their history. Some examples include the Gamma Velorum cluster, showing two distinct kinematic components (Jeffries et al. 2014), Taurus (Kraus et al. 2017), Cygnus OB2 and its complex substructures (Wright et al. 2016), and even smaller structures like the TWA moving group (Ducourant et al. 2014). OB associations, in particular, are spectacularly confirming the expectation that their kinematic substructure is reminiscent of its initial structure (e.g. Larson 1981; Wright et al. 2016). The same complexity emerges when studying the geometry and the internal motion of molecular clouds (Falgarone, Phillips & Walker 1991; Hacar et al. 2013): Starting from structure analysis of prestellar cores (Ladjelate et al. 2020) and from the variegated shapes taken by filaments and filamentary networks in the early phase of stellar formation (Hacar et al. 2018; Hoemann, Heigl & Burkert 2021) that give rise to distinct stellar populations (e.g. Arzoumanian et al. 2019), it is natural to think that the complex, fractal structure is inherited by young stars (Elmegreen & Elmegreen 2001; Gutermuth et al. 2008).

In this regard, the nearest OB association to the Sun, Scorpius–Centaurus (Sco-Cen), naturally stands as an ideal benchmark. Spanning an enormous area of approximately $80° \times 40°$ in the sky, this very young ($t < 20$ Myr; Pecaut, Mamajek & Bubar 2012) region comprises a few regions still actively forming stars (Wilking, Gagné & Allen 2008). The exquisite concoction of closeness, youth and low extinction, allowing detection of members down to the brown dwarf regime, has made it the target of many studies in the last decades (e.g. de Zeeuw et al. 1999; Mamajek, Meyer & Liebert 2002) involving binaries (e.g. Janson et al. 2013), primordial discs (e.g. Carpenter et al. 2006), high mass (Chen et al. 2012), Sun-like (e.g. Pecaut et al. 2012) and low-mass stars (e.g. Lodieu 2013), and even young planets (e.g. Janson et al. 2021).

The star formation history of Sco-Cen is closely related to its spatial structure. Sco-Cen is classically divided into three main subgroups (Blaauw 1946; de Zeeuw et al. 1999): going towards lower galactic longitude, Upper Scorpius (USCO), Upper Centaurus-Lupus (UCL), and Lower Centaurus-Crux (LCC). Both density and age have a spatial gradient, with USCO being more compact and younger than UCL and LCC (Pecaut & Mamajek 2016). This intriguing observation led Preibisch & Zinnecker (1999) to put forward the idea of a triggered star formation, where the process, started in LCC, gradually expanded eastward by means of supernova shocks, causing star formation in USCO after some Myr. The same USCO is thought to have triggered a minor burst of star formation in the ρ Ophiuchi complex, which is still ongoing (Wilking et al. 2008). However, the

picture is complicate, and the three subgroups appear composed of many smaller entities, each bearing a peculiar mark while being conditioned by feedback from the surrounding environment (Wright & Mamajek 2018). This tension between nature and nurture has given renovated impulse to the idea of investigating the substructure of the three subgroups to gain knowledge into their star formation histories.

The presence of at least a certain degree of substructure is evident even to visual inspection in the youngest part of Sco-Cen, USCO. This rather compact ($98 \times 24 \times 18$ pc^3; Galli, Joncour & Moraux 2018) region of Sco-Cen, home to the bright Antares (Ohnaka et al. 2013), received a great attention on its own due to the interplay of kinematic and age peculiarities. Notably, a consistent age determination for USCO has long been elusive. While the first photometric studies argued for an age of $\sim$5 Myr with no significant spatial and temporal spread (Preibisch & Zinnecker 1999; Preibisch et al. 2002), recent work has been increasingly prone to an older age ($t \sim 11$ Myr) with a significant spread ($\Delta t \sim 7$ Myr; Pecaut et al. 2012); the debate on the dependence of the latter on position (Pecaut & Mamajek 2016), spectral class (Rizzuto et al. 2016), systematic artefacts due to stellar models (Feiden 2016), or an extended star formation history (Fang, Herczeg & Rizzuto 2017) has been vivid in recent years.

In this context, insight from kinematic studies are pivotal to shed light on the problem. While the first studies, limited to few bright members, could not but aim at assessing a single common expansion age (Blaauw 1978) – a solid lower limit of $\sim$10 Myr, in this regard, was put by Pecaut et al. (2012)–, nowadays we do have the means to investigate the whole kinematic substructure of USCO.

This paper is organized as follows: After defining the selection criteria for our sample of USCO stars, together with the astrometric, kinematic and photometric data employed throughout this work (Section 2), we introduce in Section 3 our tool, MADYS, and apply it to the region to recover and characterize the dual kinematic substructure found within the association. Section 4 is dedicated to the age determination of clustered and diffuse populations, conducted in a threefold way. In Section 5, we discuss our results within the framework of previous studies of the region, with particular emphasis on its star formation history. Finally, in Section 6, we provide a brief summary of the results of this work. Appendices A and B explore in greater detail two quantities introduced to correctly handle data coming from *Gaia*: namely, a quality cut defined to exclude unreliable G_{BP} and G_{RP} photometric measurements and a set of corrections to remove the fraction of individual proper motions due to the reflection of the relative motion of USCO with respect to the Sun.

2 DATA

2.1 Sample selection

Motivated by the idea of exploiting the full potential of the latest *Gaia* release (EDR3; Gaia Collaboration et al. 2021), we decided to construct a novel sample of USCO sources, independent from the DR2-based samples already present in the literature (e.g. Luhman & Esplin 2020). A preliminary deep query was done in a region virtually encompassing the whole USCO, employing just minimal cuts on astrometry (α, δ, π) and kinematics ($\mu_\alpha^* = \mu_\alpha \times \cos(\delta)$, μ_δ) to exclude stars either from the field or belonging to the nearby UCL subgroup (Table 1). No attempt has been done to remove, as many previous studies, sources belonging to the nearby Rho Ophiuchi region (from this moment on, ρ Ophiuchi) since we intend to explore

Table 1. Criteria for the selection of the 2D sample.

Initial query	
Query position (α_0, δ_0)	$(245°, -25°)$
Query radius (°)	20
Parallax (mas)	$5 < \pi < 11$
Proper motion along α (mas yr^{-1})	$-50 < \mu_\alpha^* < 0$
Proper motion along δ (mas yr^{-1})	$-53 < \mu_\delta^* < 0$
No. of sources	408 465
Final criteria	
Right Ascension (°)	$236 < \alpha < 251$
Declination (°)	$-29 < \delta < -16$
Parallax (mas)	$5.7 < \pi < 8$
Parallax error	$\sigma_\pi/\pi < 0.1$
Velocity along α (km s^{-1})	$v_\alpha > -12.8$
Velocity along δ (km s^{-1})	$-20.4 < v_\delta < -12.8$
Apparent G magnitude (mag)	$G < 20$
No. of sources	2745

Note. Coordinates and proper motions are referred – as usual for *Gaia* EDR3 – to the ICRS at epoch J2016.0.

in detail its relation with the bulk of USCO. We will simply refer to our sample as *USCO*.

Membership to USCO has been defined operationally, by inspecting the 5D phase space $(\alpha, \delta, \pi, v_\alpha, v_\delta)$, with

$$v_\alpha[\text{km s}^{-1}] = A \times \mu_\alpha^*/\pi, \qquad (1)$$

$$v_\delta[\text{km s}^{-1}] = A \times \mu_\delta/\pi, \qquad (2)$$

with $A = 4.74$ km yr s^{-1} being the conversion factor between AU yr^{-1} and km yr s^{-1}. The line-of-sight velocities v_α and v_δ are more suitable than proper motion components in a region of non-negligible radial depth $(\Delta r \sim 50$ pc) and with parallax uncertainties no more as limiting as it was in the pre-*Gaia* era.

A clear concentration of sources emerges, distinguishing USCO from the field (Fig. 1). A comparison with an independent sample, the DR2-based catalogue of Sco-Cen members by Damiani et al. (2019),[1] yielded excellent agreement: out of their 2330 stars, 2129 (~ 91 per cent) were recovered; the fraction would have risen to 2298/2330 (~ 98 per cent), if we employed their same cut on minimum distance ($\pi < 10$ mas). However, we opted for a more conservative $\pi < 8$ mas not to detrimentally affect field contamination.

We consider those stars – the only additional caveats being $G < 20$ and $\sigma_\pi/\pi < 0.1$ – as our final sample (which we will call *2D sample*). The complete set of defining criteria are summarized in Table 1, while the sky distribution of the sample, comprising 2745 stars, is shown in Fig. 2.

2.2 Radial velocities

An unbiased analysis of an extended region on the sky cannot be achieved, due to projection effects, without a full knowledge of the 6D phase space of its members: RVs are crucial not only to identify interlopers but also, more importantly, to correctly analyse stellar motions (see Appendix B for details).

A complete analysis in the 6D phase space has been performed on the subsample possessing reliable RV measurements. In addition

[1] Actually, with the subsample defined by the same cuts on α and δ as our sample in order to exclude UCL members. Only their *bonafide* members were considered.

to *Gaia* EDR3, we collected data from APOGEE DR16 (Ahumada et al. 2020) and GALAH DR3 (Buder et al. 2021); whenever multiple measurements were present, the datum with the smallest error bar was chosen. After selecting sources with a relative error on RV < 0.1 or an absolute error <1 km s^{-1}, we defined the Cartesian frame (x, y, z):

$$\begin{cases} x = r\cos(\delta - \delta_P)\sin(\alpha - \alpha_P) \\ y = r\cos(\delta - \delta_P)\cos(\alpha - \alpha_P) \\ z = r\sin(\delta - \delta_P) \end{cases}, \qquad (3)$$

where the pole $(\alpha_P, \delta_P) = (243°09, -23°03)$ points, for convenience, towards the mean equatorial coordinates of the sample. Finally, we restricted only to sources with propagated errors on v_x, v_y, and v_z simultaneously satisfying the three conditions[2]:

(i) $|\sigma_{v_x}/v_x| < 0.1$ OR $\sigma_{v_x} < 0.1$ pc Myr^{-1},
(ii) $|\sigma_{v_y}/v_y| < 0.1$ OR $\sigma_{v_y} < 0.1$ pc Myr^{-1},
(iii) $|\sigma_{v_z}/v_z| < 0.1$ OR $\sigma_{v_z} < 0.1$ pc Myr^{-1}.

The final RV sample (*3D sample*) comprises 771 stars, ~ 28 per cent of the 2D sample (Table 2).

Although we decided not to appoint the 3D sample as our main focus, because it only imperfectly reproduces the real distribution of sources,[3] we will employ it in a twofold way: On the one hand, it will provide us with a way to quantify the effect of the association's mean motion with respect to the Sun (which we will call, from this moment on, *bulk motion*); on the other hand, it will offer a constant comparison with the 2D sample to check the validity of our results: by comparing, whenever possible, features observed in 2D with their 3D counterparts, we are able to rule out the possibility of a random alignment of sources lying at different distances, i.e. a perspective effect.

As regards the former aspect, a simple geometrical argument proves that, taken a group of stars, the knowledge of their proper motions alone is not sufficient to disentangle between a real expansion and a non-zero mean radial motion with respect to the line of sight through its centre, i.e. a *virtual expansion* (see discussion in de Zeeuw et al. 1999). Transverse motions, too, are split up into velocity components depending on (α, δ, π). Thus, any non-zero bulk motion will manifest itself as a bias in the kinematic reconstruction.

Once having estimated that the centre of the 3D sample approximately lies at $(\alpha_c, \delta_c, r_c) = (244°55, -23°79, 143.3$ pc), we computed the mean velocity components in a Cartesian frame centred on it. Expressed in the standard right-handed Cartesian Galactic frame, our 3D sample has median velocity components $(U, V, W) = (-4.788 \pm 0.019, -16.378 \pm 0.015, -6.849 \pm 0.016)$ km s^{-1}, with a precision gain of almost one order of magnitude relative to previous estimates (Luhman & Mamajek 2012; Galli et al. 2018). Finally, we determined the projections of this bulk motion on the proper motions of each star of the 2D sample, and subtracted them (see Appendix B for details). We verified that, due to the angular extent of USCO, this bias does not significantly affect the shape of the substructures nor the timing of their maximum spatial concentration.

Even correcting for bulk motion, a similar (although smaller) projection effect keeps affecting individual stellar velocities, due to the rotation of the $(v_\alpha, v_\delta, v_r)$ plane with (α, δ). Again, the
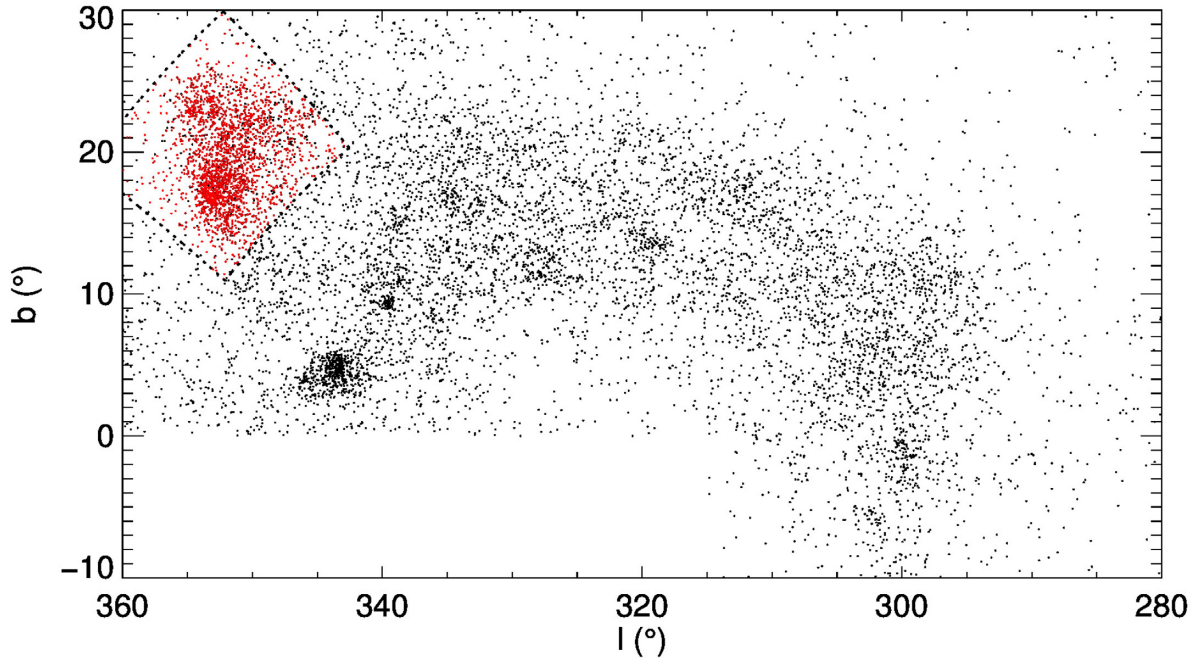
[2] 1 km s^{-1} = 1.02 pc Myr^{-1}.
[3] The distribution of sources possessing RV does not appear as a random pick of the *Gaia* sample, but rather – as expected – as the union of distinct surveyed regions.
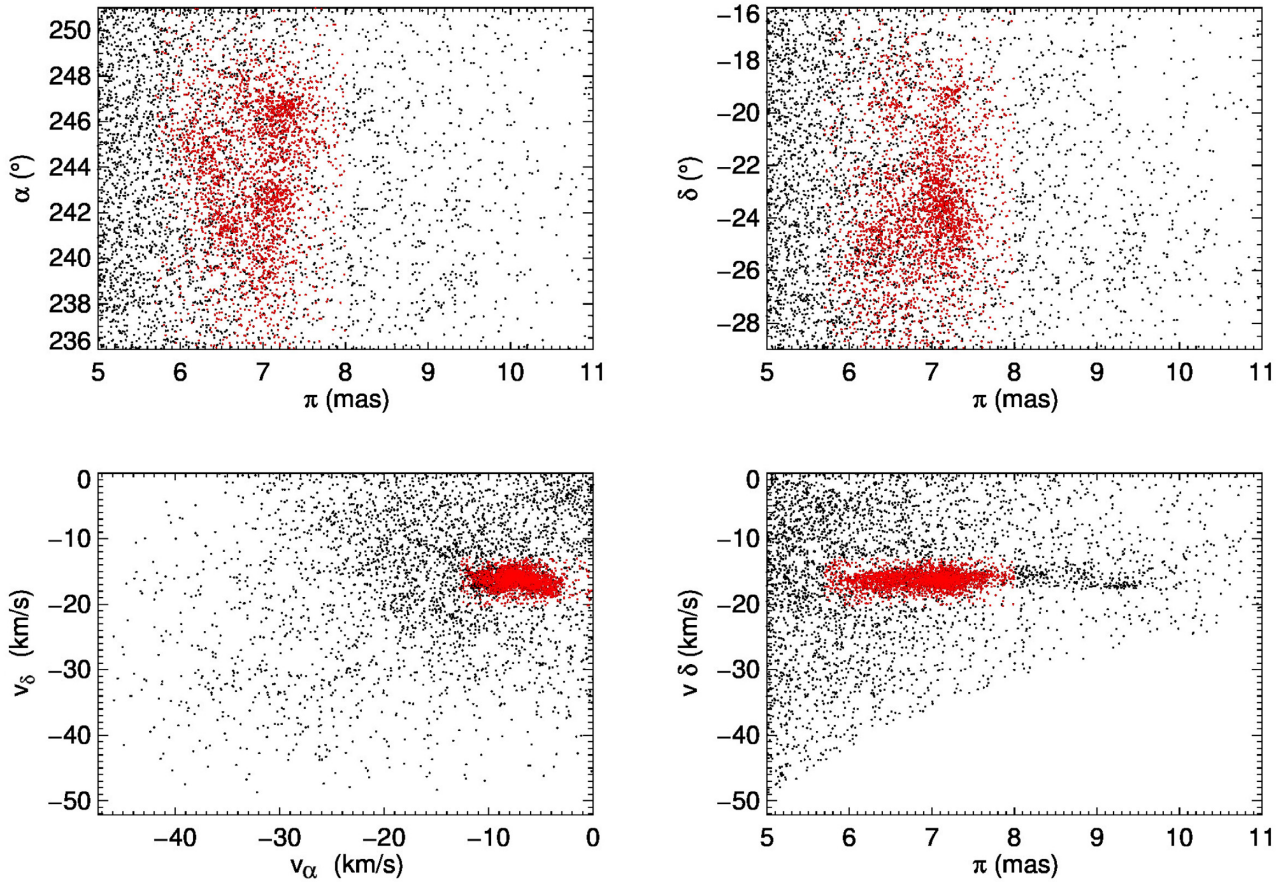
Figure 1. Detection of USCO (red) within the 5D phase space. Only field stars (black) with $236° < \alpha < 251°$, $-29° < \delta < -16°$, $G < 20$ mag, and $\sigma_\pi/\pi < 0.1$ are shown for the sake of clarity.

Figure 2. Sco-Cen bonafide members from Damiani et al. (2019), shown in black. USCO can easily be distinguished in the upper left. The sample used throughout this work and defined by the cuts of Table 1 is displayed in red. The criteria on Right Ascension and Declination define the region bordered by the dashed lines.

Table 2. Median uncertainties on astrometry and kinematics for the 2D and 3D samples.

2D sample		3D sample	
No. of sources	2745	No. of sources	771
Median σ_π (mas)	0.06	Median σ_x (pc)	0.03
Median $\sigma_{\mu_\alpha^*}$ (mas yr^{-1})	0.07	Median σ_y (pc)	0.82
Median σ_{μ_δ} (mas yr^{-1})	0.05	Median σ_z (pc)	0.03
Median σ_{v_α} (km s^{-1})	0.08	Median σ_{v_x} (pc Myr^{-1})	0.05
Median σ_{v_δ} (km s^{-1})	0.14	Median σ_{v_y} (pc Myr^{-1})	0.04
		Median σ_{v_z} (pc Myr^{-1})	0.10

angular extent of the association is not too big to hinder the approach altogether.[4]

2.3 Photometry

Whereas astrometric and kinematic data were gathered from *Gaia* EDR3, the photometry comes from *Gaia* DR2 (Gaia Collaboration et al. 2018): The reason is that the set of isochrones that we used relies on *Gaia* DR2, and the filter response is not exactly the same between the two releases.[5] As G_{BP} and G_{RP} photometry can be severely contaminated by the background at faint magnitudes (Busso et al. 2021), a quality cut was needed to discriminate whether G_{BP} and G_{RP} magnitudes could be considered reliable or not. Following the line of reasoning by Riello et al. (2021), a colour-independent *BP-RP excess factor* C^* was defined, starting from the available bp_rp_excess_factor C:

$$C^* - C = \begin{cases} a_0 + a_1\Delta G + a_2\Delta G^2 + a_4 G & \Delta G < 0.5 \\ a_0 + a_1\Delta G + a_2\Delta G^2 \\ \quad + a_3\Delta G^3 + a_4 G & 0.5 \leq \Delta G < 3.5 \\ a_0 + a_1\Delta G + a_4 G & \Delta G \geq 3.5 \end{cases} \quad (4)$$

where $\Delta G = (G_{BP} - G_{RP})$; the distribution of C^* peaks at about 0 for well-behaved sources at all magnitudes but, when considering subsamples of stars with similar brightness, it tends to widen out for fainter G; a varying standard deviation $\sigma_{C^*}(G)$ can be defined as

$$\sigma_{C^*}(G) = k_1 + k_2 \times G^{k_3}. \quad (5)$$

Setting a rejection threshold at 3σ, we labelled 889/2745 sources (32 per cent) as having unreliable (G_{BP}, G_{RP}) magnitudes; the effect is larger, as expected, at fainter $(G \gtrsim 15)$ magnitudes. Details on the derivation of C^* and σ_{C^*}, as well as the numerical values of the constants of equations (4)–(5) can be found in Appendix A.

Infrared measurements from 2MASS (Skrutskie et al. 2006) were gathered and cross-matched with *Gaia* magnitudes, by inspecting the radial distance r between nearby *Gaia* and 2MASS sources. About 96 per cent of the source pairs have $r < 0.7$ arcsec, so this value was chosen to establish whether a pair actually referred to the *same* source. Only measurements labelled by the best quality flag ('A') were used.

In order to study the fraction of disc-bearing stars, additional photometric data were collected from *WISE* (Wright et al. 2010) and

---

[4]Choosing a fixed Cartesian $(\hat{x}, \hat{y}, \hat{z})$ frame around a star defined by $(\alpha_0, \delta_0, \pi_0, v_{\alpha,0}, v_{\delta,0}, v_{r,0})$ such that $\hat{x} \parallel v_{\alpha,0}$, $\hat{y} \parallel v_{r,0}$, and $\hat{z} \parallel v_{\delta,0}$, the mixing between velocity components for a second star having $(\alpha_1 = \alpha_0 + 10°, \delta_1 = \delta_0 + 10°, \pi_1, v_{\alpha,1}, v_{\delta,1}, v_{r,1})$ is such that $v_x = 0.985v_{\alpha,1} - 0.030v_{\delta,1} + 0.171v_{r,1}$, $v_y = -0.174v_{\alpha,1} - 0.171v_{\delta,1} + 0.970v_{r,1}$, $v_z = 0.985v_{\delta,1} + 0.174v_{r,1}$.

[5]We verified that, for our initial sample of 408 465 sources, G_{BP} magnitudes are on average 0.3 mag dimmer in EDR3.
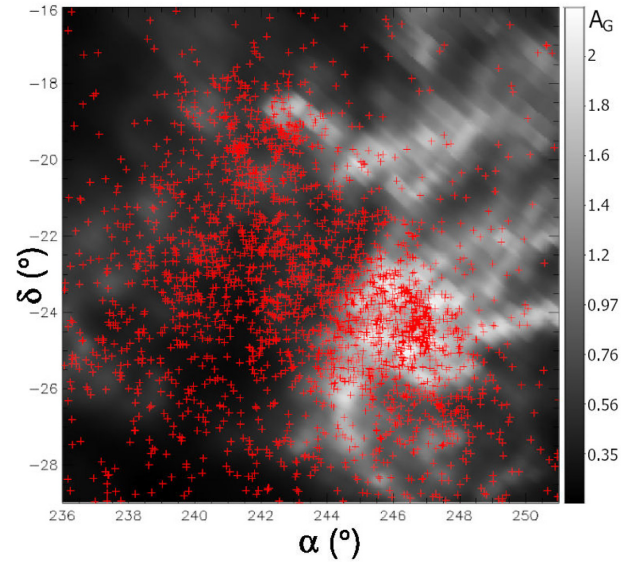
Figure 3. Line-of-sight *G*-band extinction (mag) in the direction of USCO, plotted with the help of a colour bar. The 2D sample is overplotted in red. The values, derived from the map by Leike et al. (2020), are computed for simplicity at a fixed $d = 160$ pc, encompassing ~ 90 per cent of the sample.

ALLWISE (Cutri et al. 2021). If simultaneously present, ALLWISE magnitudes were preferred over *WISE* data. As with 2MASS data, only measurements with the best quality flag ('0') were employed; measurements with $W1 < 8.1$ mag, $W2 < 6.7$ mag, $W3 < 3.8$ mag, and $W4 < -0.4$ mag were not considered, due to saturation problems at bright magnitudes (Cutri et al. 2012). The effect of flux contamination from nearby sources, increasingly affecting surveys with longer wavelengths, will be assessed in Section 4.3.

2.4 Extinction

Despite the proximity of the region, interstellar extinction is not to be overlooked: The youngest region here studied, the ρ Ophiuchi cloud, can reach $A_V \sim 40$–50 mag in its core (Wilking & Lada 1983), preventing detection of its embedded protostars (Grasser et al. 2021) in optical surveys, while the corners of the association hardly reach $A_V \approx 0.2$ mag. In order to take such intrinsic spatial variability into account, the absolute photometry of each star was corrected via the interstellar extinction map by Leike, Glatzle & Enßlin (2020). The resolution of the map is 1 pc, which at $d \sim 140$ pc typically translates into an angular resolution $\Delta\theta = 0°.4$, comparable to the extent of the largest molecular clouds within ρ Ophiuchi (Rigliaco et al. 2016). The provided *G*-band extinction was converted, case by case, into the appropriate band using a total-to-selective absorption ratio $R = 3.16$ and extinction coefficients A_λ taken from Wang & Chen (2019). A sketch of the integrated *G*-band extinction at a constant distance of $d = 160$ pc is shown in Fig. 3.

3 ANALYSIS

We developed a flexible PYTHON tool, MADYS, which, given a list of stars, automatically retrieves and cross-matches photometry from different catalogues, corrects for interstellar extinction, and employs photometric data judged reliable to get ages and masses of individual stars by comparison with pre-MS isochrones. Additionally, MADYS can look for evidence of primordial discs and perform a group kinematic analysis. The tool, to be shortly made publicly available,
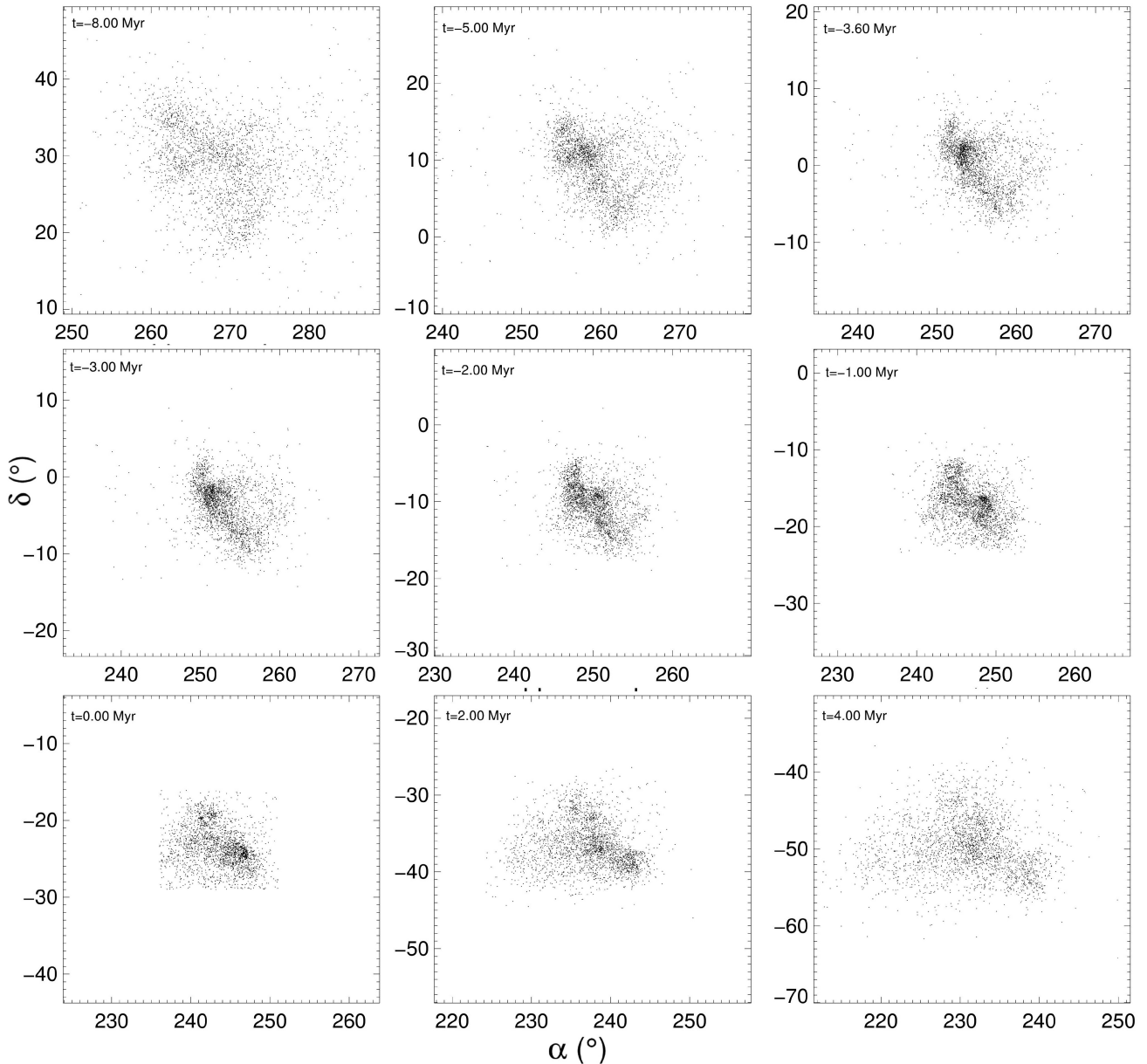
Figure 4. A few frames of the time evolution of USCO. Some clear overdensities of sources emerge at different times.

constitutes a development of the age determination tool already employed in Janson et al. (2021). We describe in this section the results of the kinematic analysis of our samples.

3.1 Kinematic subgroups

As hinted in the previous sections, our method works by tracing back the celestial coordinates of the 2D sample on the basis of their present μ_α and μ_δ, under the assumption that dynamical interactions among members are negligible. The inspection of the time evolution of the sample[6] in the (α, δ) plane suggests that a wealth of information is still encapsulated beneath the present structure of USCO: The sight of different parts of the association, clustering at different times in

the past, was what encouraged us to tentatively distinguish subgroups based on a purely kinematic way (Fig. 4).

Our 2D sample was inspected via a semi-automated approach based on iterative k-means clustering (Everitt et al. 2011). The analysis takes place in n 4D planes $(\alpha(t), \delta(t), v_\alpha, v_\delta)$:

$$\alpha(t) = \alpha_0 + \frac{\mu_\alpha^*}{\xi \cos(\delta)}, \tag{6}$$

$$\delta(t) = \delta_0 + \frac{\mu_\delta}{\xi}, \tag{7}$$

where (α_0, δ_0) are the present coordinates, $\xi = 3°\!.6\,\mathrm{Myr}^{-1}/\mathrm{mas}\,\mathrm{yr}^{-1}$ is a factor needed to express angular velocities in units of $[°\,\mathrm{Myr}^{-1}]$, and n represents the number of time-steps of an evenly spaced temporal grid ($t \in [-15, 15]$ Myr, $\Delta t = 0.2$ Myr). Every time a coherent group was visually identified at a certain t, it was subtracted from the sample. The procedure, which we will refer to as *2D*

---

[6]Available in the supplementary material as a .mp4 movie.
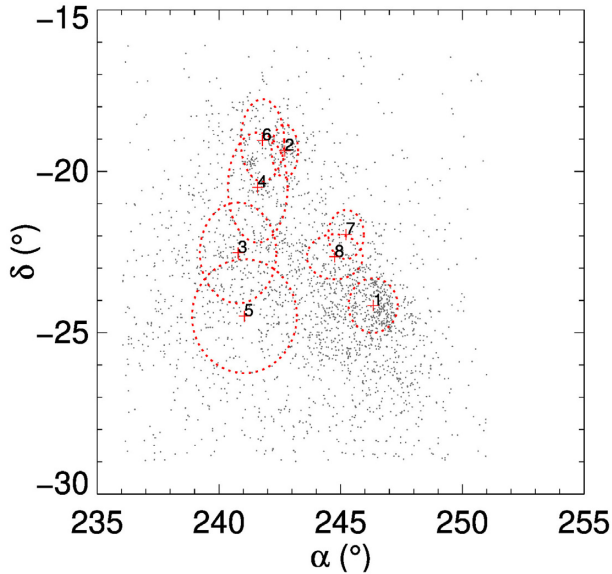
Figure 5. Mean positions of the groups at present time, with 1σ ellipses shown as dashed red curves.

analysis, was iterated, until no additional groups could be found with confidence.

A total of eight subgroups was identified, each with a peculiar mark in the phase space (Figs 5–6); their main properties are summarized in Table 3, and include an estimate of the moment of maximum coherence t_K, which we will define in Section 4.1.

To get an estimate of the false alarm probability for these overdensities, i.e. to rule out that similar features could be produced by chance, we set up, for each subgroup, the following test: we built reshuffled 2D samples by randomly assigning to each star a quintuplet $(\alpha, \delta, \pi, \mu_\alpha, \mu_\delta)$, every parameter being drawn independently from its natal distribution; then, we traced back the positions at the appropriate t_K, and excluded the most distant star – with respect to the 5D normal distribution – in an iterative way, until as many stars as in the original subgroup were left. Then, we estimated the false alarm probability (2DP) as the ratio between the number of simulations ending up – at any time – with a more clustered group than the observed one and the total number of performed simulations:

$$2DP \approx f = \frac{N_{\text{good}}}{N_{\text{good}} + N_{\text{bad}}}. \tag{8}$$

We considered as a *success* a simulation yielding a median angular distance from its centre and a 1σ velocity ellipse $\sigma_{v_\alpha} \cdot \sigma_{v_\delta}$ smaller than the observed ones. Similarly, a 3D false alarm probability (3DP) was computed by assigning to each star a sextuplet (x, y, z, v_x, v_y, v_z) and considering as *success* a simulation yielding a smaller median distance from the centre than the real one and a 1σ velocity ellipse[7] $\sigma_{v_x} \cdot \sigma_{v_y} \cdot \sigma_{v_z}$ smaller than the observed ones.

As already mentioned in Section 2.2, the results is not influenced significantly by removal of projection effects due to USCO's bulk motion: The distinction among subgroups was carried out on the original 2D sample, using the corrected velocities as a tool to estimate their t_K. Also, errors on v_α and v_δ, increasing linearly with time, cannot be the source of the pattern observed, as 95 per cent of the sources has error $<0°.5$ at $t = -5$ Myr.

We find strong evidence for the physical nature of Groups 1, 2, 3, 4, 5 and the tiny Group 7. While Group 1 corresponds to the ρ Ophiuchi region, as confirmed by a cross-match with the catalogue by Cánovas et al. (2019),[8] Groups 2 and 3 visually resemble Groups E and H of a recent work by Kerr et al. (2021). The latter, an elongated structure spanning $\sim 1° \times 3°$ with a north–south orientation, is remarkable in many aspects: at a mean distance of 141 ± 4 pc, its shorter side, somewhat broadened by proper motion uncertainties (the 1σ angular error at $t = -3.5$ Myr is $\sim 0°.15$), spans just ~ 2pc.

The randomness of Groups 6 and 8 – for which the 3D sample is quite small – cannot be ruled out. For the moment being, we will keep on retaining the division in subgroups, which will be further discussed in Section 5.

3.2 Clustered and diffuse populations

An even more interesting result is obtained if we put together the stars belonging to the subgroups, and compare the resulting population with the remaining stars. We might call them the *clustered population* (1442 stars) and the *diffuse population* (1303 stars), respectively. The difference between the two populations is profound: while the former appears to clump, both in 2D and in 3D, at different times in the past, the latter does not.

Since the clustered population has been created by assembling subgroups detected individually during the 2D analysis, the result is somewhat surprising. The 3D sample strengthens the idea of a kinematic duality within the association, which we tried to quantify by taking the median distance from the mean 3D position for the whole sample and for the two subpopulations:

$$d_M(t) = \text{median}\left(\sqrt{[x(t) - \bar{x}(t)]^2 + [y(t) - \bar{y}(t)]^2 + [z(t) - \bar{z}(t)]^2}\right). \tag{9}$$

Looking at Fig. 7, it is clear that $\tilde{d}_M := \min(d_M(t))$ is smaller for the clustered population ($\tilde{d}_M = 7.4 \pm 0.1$ pc) than for the diffuse population ($\tilde{d}_M = 12.4 \pm 0.2$ pc), and reaches its minimum value earlier for the former ($t = -3.2$ Myr) than for the latter ($t = 0$ Myr). Errors are computed via a Monte Carlo simulation for uncertainty propagation, i.e. by repeating the procedure while randomly varying, in a normal fashion, input velocities according to their uncertainties.

While the moment minimizing $d_M(t)$ cannot be assumed as an age estimate but rather as a lower limit (see Section 5), its minimum value is indeed a measure of the degree of concentration of the population at that time. To shed light on the significance of its difference between the clustered and diffuse population, we ran the same set of 3D simulations described in Section 3.1, using $t_K = 3.2$ Myr. Out of 100 000 simulations, none behaved better than the original sample; the best-fitting Gaussian distribution of their $\tilde{d}_M(t)$, defined by $(\mu, \sigma) = (9.68\,\text{pc}, 0.16\,\text{pc})$, places a confidence level on the observed clustering at $\sim 14\sigma$.

4 AGE DETERMINATION

The presence of a kinematic duality within USCO poses new questions: Is there any difference in the age distribution of the clustered and diffuse population? Is the distribution of t_K reflected into an age spread between the groups? To provide tentative answers,

---

[7]Computed, in the 3D case, from the samples restricted to the [16,84] percentiles to minimize the impact of possible outliers.

[8]Out of 517 stars that are both in their sample and in our 2D sample, 342 (66 per cent) have been assigned to Group 1. Most of the remaining stars have not been assigned to any group.
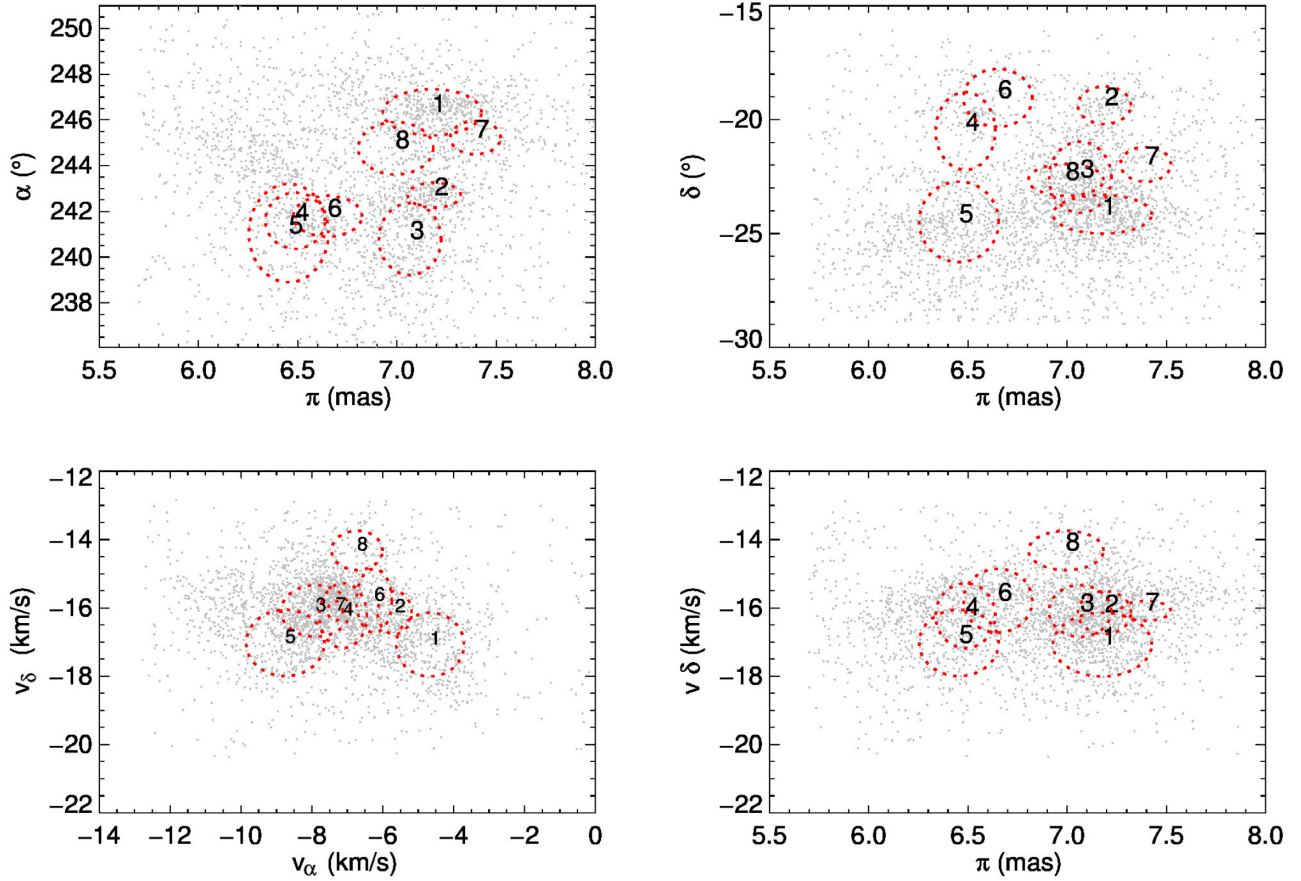
Figure 6. Distribution of the groups in the 5D phase space, with 1σ ellipses shown as dashed red curves.

Table 3. Number of stars (n) and mean positions in the phase space of the groups.

Group	2D n	3D n	2DP (per cent)	3DP (per cent)	α (°)	δ (°)	π (mas)	v_α (km s^{-1})	v_δ (km s^{-1})	MRS
1	467	121	<0.1	<0.1	246.3 ± 1.0	-24.2 ± 0.8	7.2 ± 0.2	-4.7 ± 1.0	-17.1 ± 0.9	i Sco
2	114	64	<0.1	<0.1	242.7 ± 0.5	-19.4 ± 0.8	7.2 ± 0.1	-5.7 ± 0.5	-16.1 ± 0.6	ν Sco B
3	396	196	<0.1	<0.1	240.8 ± 1.6	-22.5 ± 1.6	7.1 ± 0.2	-7.9 ± 1.0	-16.1 ± 0.8	b Sco
4	156	67	<0.1	<0.1	241.6 ± 1.2	-20.5 ± 1.7	6.5 ± 0.2	-7.1 ± 0.7	-16.2 ± 1.0	HD 144273
5	166	44	12.2	0.2	241.0 ± 2.2	-24.5 ± 1.8	6.5 ± 0.2	-8.7 ± 1.1	-17.0 ± 1.0	HIP 77900
6	58	13	13.5	18.4	241.8 ± 0.9	-19.0 ± 1.3	6.7 ± 0.2	-6.2 ± 0.5	-15.8 ± 0.9	HIP 78968
7	45	23	1.0	<0.1	245.2 ± 0.7	-22.0 ± 0.8	7.4 ± 0.1	-7.3 ± 0.2	-16.1 ± 0.3	HIP 79910
8	40	11	40.0	<0.1	244.8 ± 1.1	-22.6 ± 0.7	7.0 ± 0.2	-6.7 ± 0.7	-14.3 ± 0.6	HD 146457

Note. Errors should be read as sample standard deviations, equivalent to the semimajor axes of the ellipses in Fig. 6. The brightest star of each group (MRS) is also indicated.

we tried to compute the age of both the groups and the diffuse population in a threefold way: via usual isochrone fitting, through the fraction of disc-bearing stars and in a purely kinematic way.

The results of our analysis are summarized in Table 4.

4.1 Kinematic age

We have seen in Fig. 7 that at least part of the association was more clustered in the past, as visually evident in Fig. 4. We might think to exploit this observation to constrain the age of the groups, since their very detection encapsulates, by definition, an age estimate: The moment of maximum spatial coherence might reflect that of the common birth of the members.

For each group, a quantitative estimate can be obtained by defining a coherence function $K(t)$ as the length of the minimum spanning tree connecting the points at time t. To minimize the impact of possible outliers, we excluded the 10 per cent longer branches for 2D estimates, and the 32 per cent for 3D estimates.

The minimum of $K(t)$ provides us with an age estimate that we will call *kinematic* age t_K. For 2D estimates, we employed both corrected and uncorrected proper motion components, with minimal influence on the results: from this moment on, we will always think as t_K as referring to the latter case. The trend of $K(t)$ for all the groups is shown in Fig. 8, while the appearance of individual groups at $t = t_K$ is shown in Fig. 9. The error on t_K, again, was computed via $N = 1000$ Monte Carlo simulations per group.
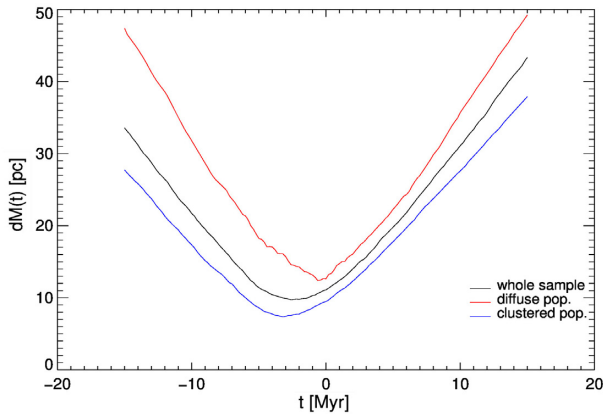
Figure 7. The trend of $d_M(t)$, defined by equation (9), for the clustered and the diffuse population. The clustered population reaches a minimum $d_M \approx$ 7.4 pc at $t = -3.2$ Myr, while the diffuse population does not appear not to have been more clustered in the past.

Table 4. Age estimates obtained through photometry (t_P), discs (t_D), and kinematics (t_K).

Group	n	t_P (Myr)	s	f_D (per cent)	t_D (Myr)	t_K (Myr)
1	467	2.6 ± 0.1	1.5	31 ± 3	3.0 ± 0.3	0.0 ± 0.1
2	114	4.4 ± 0.1	1.0	27 ± 6	$3.2^{+0.6}_{-0.5}$	0.1 ± 0.2
3	396	4.8 ± 0.1	1.1	25 ± 3	3.5 ± 0.3	3.7 ± 0.3
4	156	5.2 ± 0.1	1.4	17 ± 4	$4.5^{+0.7}_{-0.5}$	3.4 ± 0.6
5	166	6.1 ± 0.2	1.6	14 ± 4	$4.8^{+0.8}_{-0.6}$	3.8 ± 0.4
6	58	4.4 ± 0.2	1.2	26 ± 8	$3.4^{+0.9}_{-0.7}$	2.1 ± 0.6
7	45	6.2 ± 0.2	1.2	9 ± 4	$6.0^{+1.7}_{-1.0}$	0.3 ± 0.4
8	40	5.4 ± 0.3	1.7	12 ± 6	$5.3^{+1.7}_{-1.0}$	0.4 ± 0.8
clust.	1442	4.5 ± 0.1	1.5	24 ± 2	3.6 ± 0.2	–
diff.	1303	8.2 ± 0.1	3.7	10 ± 1	5.7 ± 0.3	–

Note. The number of stars in each group (n), the sample standard deviation of isochronal ages (s), and the fraction of discs (f_D) are also shown.

4.2 Isochronal ages

We derive a second age estimate for the groups by employing the most common method of age determination for pre-MS stars: isochrone fitting.

We decided to employ BT-Settl CIFIST2011_2015 isochrones (Baraffe et al. 2015) due to their large dynamical range ($m \in [0.01, 1.4]$ M$_\odot$, $a \in [1, 1000]$ Myr), spanning from young pre-MS objects in the brown dwarf regime to F-type stars. A constant solar metallicity, typical of most nearby star-forming regions, was assumed (D'Orazi, Biazzo & Randich 2011).

The best-fitting solution for individual stars was found by leveraging as more as possible the available *Gaia* and 2MASS photometry: We averaged three estimates per stars, coming from the *channels* $\mathcal{F} = \{G, J\}$, $\mathcal{F} = \{G, H\}$ and $\mathcal{F} = \{G_{BP}, G_{RP}\}$; in this way, it was possible to simultaneously employ the precision of *Gaia G* magnitude, the long colour baseline of $G - J$ and $G - H$, protecting against measurements errors, and a third channel that is simultaneously independent of G and 2MASS data.

Whenever a filter in one channel did not fulfil the conditions described in Section 2.3, the corresponding channel was not used; also, a channel was ignored every time its total photometric error was larger than 0.15 mag; stars lying at more than 3σ outside the region delimited by the isochrones were considered unfitted.

The results for the subgroups are shown in the second row of Table 4: Going from Groups 1 to 7, the median age grows by ∼4 Myr.

What really stands out is the comparison between the clustered and the diffuse population (Fig. 10): While the former has a median age of ∼4.5 Myr and two well-defined peaks at ∼1 and ∼5 Myr, the latter has a median age of ∼8.2 Myr, with a much flatter distribution.

4.3 Fraction of discs

Even though USCO has largely depleted its initial gas and dust reservoir (Mathews et al. 2012), hints of accretion might still be found for some stars (Dahm & Carpenter 2009). Indeed, the disappearance of primordial discs, signalled by the opening of holes and gaps increasingly straining their initial SED (Espaillat et al. 2012), has been shown to occur with an exponential fashion (Haisch, Lada & Lada 2001): the fraction of disc-bearing stars within young populations ($t \lesssim 10$ Myr) can therefore be used as an independent age estimator.

The long wavelengths of *WISE W*1, *W*2, *W*3, *W*4 filters are particularly suitable for detecting young discs, with longer wavelengths probing larger distances from the star due to their intrinsic thermal structure (Lada & Wilking 1984). While IR excesses in *W*2 and *W*3 are suggestive of inner discs, *W*4 is associated with a colder outer disc (Kuruwita et al. 2018). The effect that we wish to look for is enormous: the luminosity of these discs in *W*3 and *W*4 can outshine that of the star itself (Luhman & Mamajek 2012). Previous works like Luhman & Mamajek (2012), Pecaut & Mamajek (2016) have greatly used *W*4 magnitudes – often in combination with *W*3 – for this scope, due to its capability of probing outer disc zones.

To construct a reliable *W*3 (*W*4) sample, we selected only data with the best photometric quality flag ('0'), photometric error <0.2, and apparent magnitudes *W*3 > 3.8 (*W*4 > −0.4).[9] An insidious problem is that of *flux contamination*: since the angular resolution θ_i increases for redder bands,[10] the measured flux becomes more at risk of including a non-negligible contribution of nearby, unresolved sources. To quantify this effect, all the sources within an angular distance θ_i were looked for in *Gaia* EDR3.[11] Then, the G, G_{BP}, G_{RP} fluxes were converted into *W*3 and *W*4 fluxes. A semi-empirical relation between input and output fluxes was used, combining the 8-Gyr BT-Settl isochrone for $0.01 < M/M_\odot < 1.4$ and the empirical tables by Pecaut & Mamajek (2013) for $1.4 < M/M_\odot \lesssim 20$.

The fluxes from field sources were derived averaging four estimates: from G and parallax, if $\sigma_\pi/\pi < 0.5$, from G and G_{BP} if $\sigma_G + \sigma_{G_{BP}} < 0.4$, from G and G_{RP} if $\sigma_G + \sigma_{G_{RP}} < 0.4$, from G_{BP} and G_{RP} if $\sigma_{G_{RP}} + \sigma_{G_{BP}} < 0.4$. Correction for extinction was applied whenever possible. If no estimate was available for even a single neighbour of a certain star, we employed G fluxes, for consistency, for all the neighbourhood too.

The so-called contamination fraction f_c was defined as $f_c = F_{cont}/F_{tot}$, where F_{tot} is the flux from the *M*-band of interest and F_{cont} the flux from all the field sources within θ_i. Its effect consists in a magnitude decrease $\delta M = 2.5 \times \log_{10}(1 - f_c)$. We subtracted it to measured magnitude: this correction goes *against* disc detection, as it makes magnitudes fainter. To be protected against the uncertainties in field star distances, ages, and on the correction itself, we decided to employ in this analysis only stars with $f_c < 0.2$. This limits the

---

[9]The last criterion, due to saturation problems leading to a flux overestimation (Cutri et al. 2012), was virtually unnecessary, given the distance of USCO.

[10]2 arcsec for *J, H,* and *K*, 6.1 arcsec for *W*1, 6.4 arcsec for *W*2, 6.5 arcsec for *W*3, and 12.0 arcsec for *W*4.

[11]*Gaia* EDR3 is essentially complete at $\theta > 1.5$ arcsec, and can be relied upon until $\theta = 0.7$ arcsec for equal-mass sources (Fabricius et al. 2021).
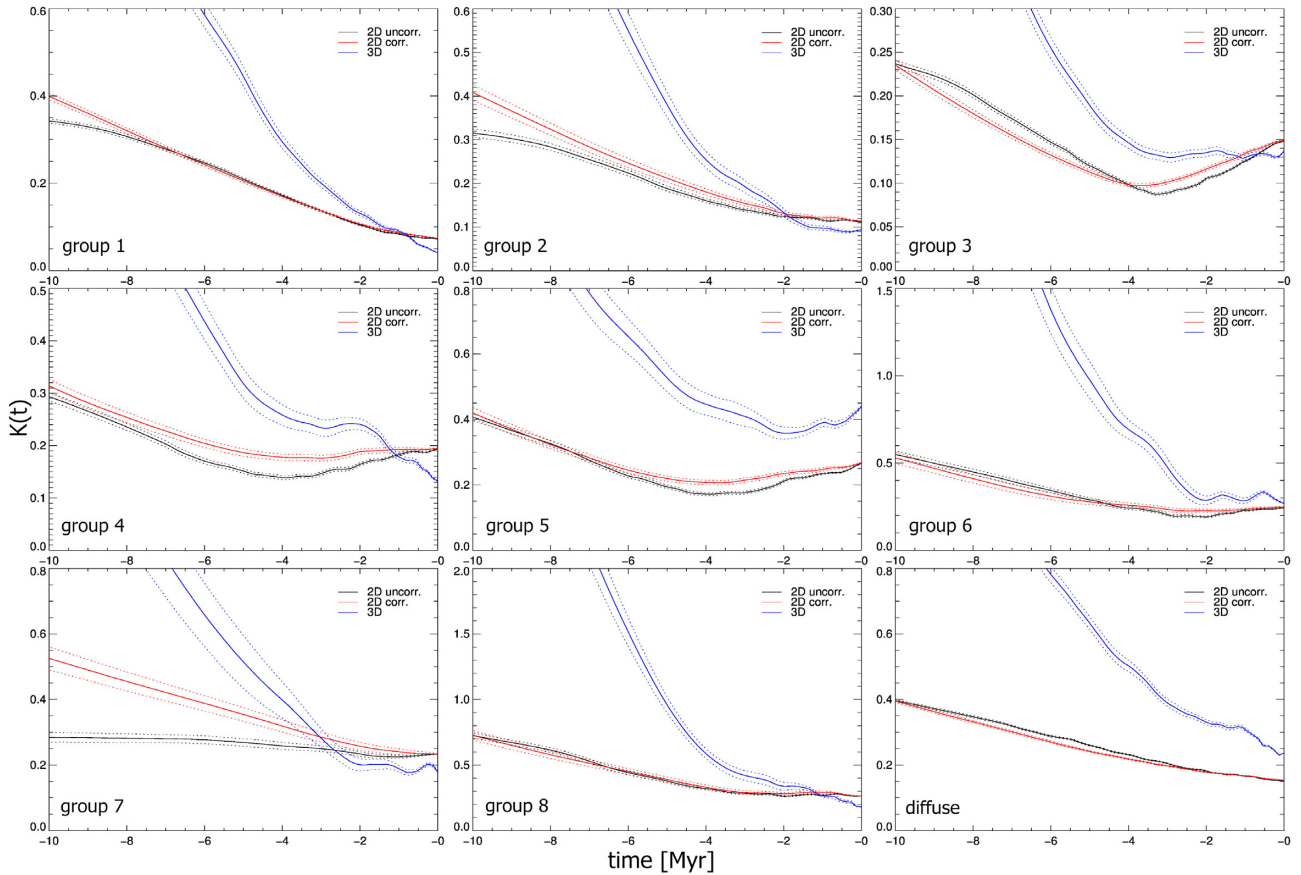
Figure 8. Coherence function $K(t)$ for the groups, computed in the 2D case – with uncorrected proper motions (black) and corrected proper motions (red) – and in the 3D case (blue), and shown in arbitrary units. 1σ errors are shown as dashed lines. The minimum of $K(t)$ pinpoints a kinematic age. The diffuse population, as expected, does not cluster in the past.

corrections to $-\delta M \approx 0.25$ mag, far less than the expected effect for full discs (several mag).

The combination of quality and contamination cuts, though, greatly reduces the availability of $W4$ data, which additionally show a tendency for being *always* redder than expected, perhaps indicating debris discs (Cody & Hillenbrand 2018) that are known to be common in the association (Carpenter et al. 2009). For these reason, we decided at the end to employ only $W3$ magnitudes.

Abundant groups like 1 and 3 show a neat bimodality in their $(G − W3, G)$ CMD, with a second sequence of stars running parallel to discless stars, but with an offset of $\sim$2 mag. Hence, we chose to use the same criterion as Pecaut & Mamajek (2016), i.e. we compute the excess $E(K − W3)$ relative to the expected colour, and identify full discs as those with $E(K − W3) > 1.5$ mag. The choice of K is based on the fact that its λ ($\approx$2.2 μm) is simultaneously too short to carry along a significant non-photospheric contribution and long enough to be protected against uncertainties in the extinction. We compute the expected $(K − W3)$ as that of the isochrone corresponding to the group isochronal age, computed in Section 4.2. We perform the same computation[12] with $(G − W3)$, both as a consistency check and as a way to recover those few stars excluded by quality cuts in 2MASS photometry.

The fraction of primordial discs f_D spans from ~ 30 to ~ 10 per cent, going from Group 1 to the diffuse population. Under the assumption of an exponential decay of f_D (see, e.g. Mamajek 2009), we derived the *disc age*,

$$t = -\tau \ln f_D, \qquad (10)$$

where $\tau = 2.5$ Myr (Mamajek 2009).

The fraction of discs we found within USCO is $f_d = 0.19 \pm 0.01$, comparable to that found by Esplin et al. (2018) and Luhman & Esplin (2020).

The clustered population appears younger than the diffuse population, whose disc fraction has been computed using as expected colours those produced by individual ages and masses. Again, setting a fiducial line in this way produces a bias *against* disc detection and hence against an age spread.[13]

A comparison of our sample of disc-bearing stars with that by Luhman & Esplin (2020) shows that, among 161 full discs having a cross-match, 145 (148) are there identified as full discs, the remaining 16 (13) being labelled as debris/evolved transitional discs, for the criterion employing $G − W3$ $(K − W3)$. It is significant that the

---

[12]Adopting the same threshold of 1.5 mag, as G-band emission is purely photospheric.

[13]If the excess is computed starting from individual positions rather than from a fiducial line, it will be harder for the youngest stars to have an excess beyond 3σ; the opposite applies for stars older than the fiducial line, but they are usually not expected to show an IR excess at all. The overall effect goes against disc detection.
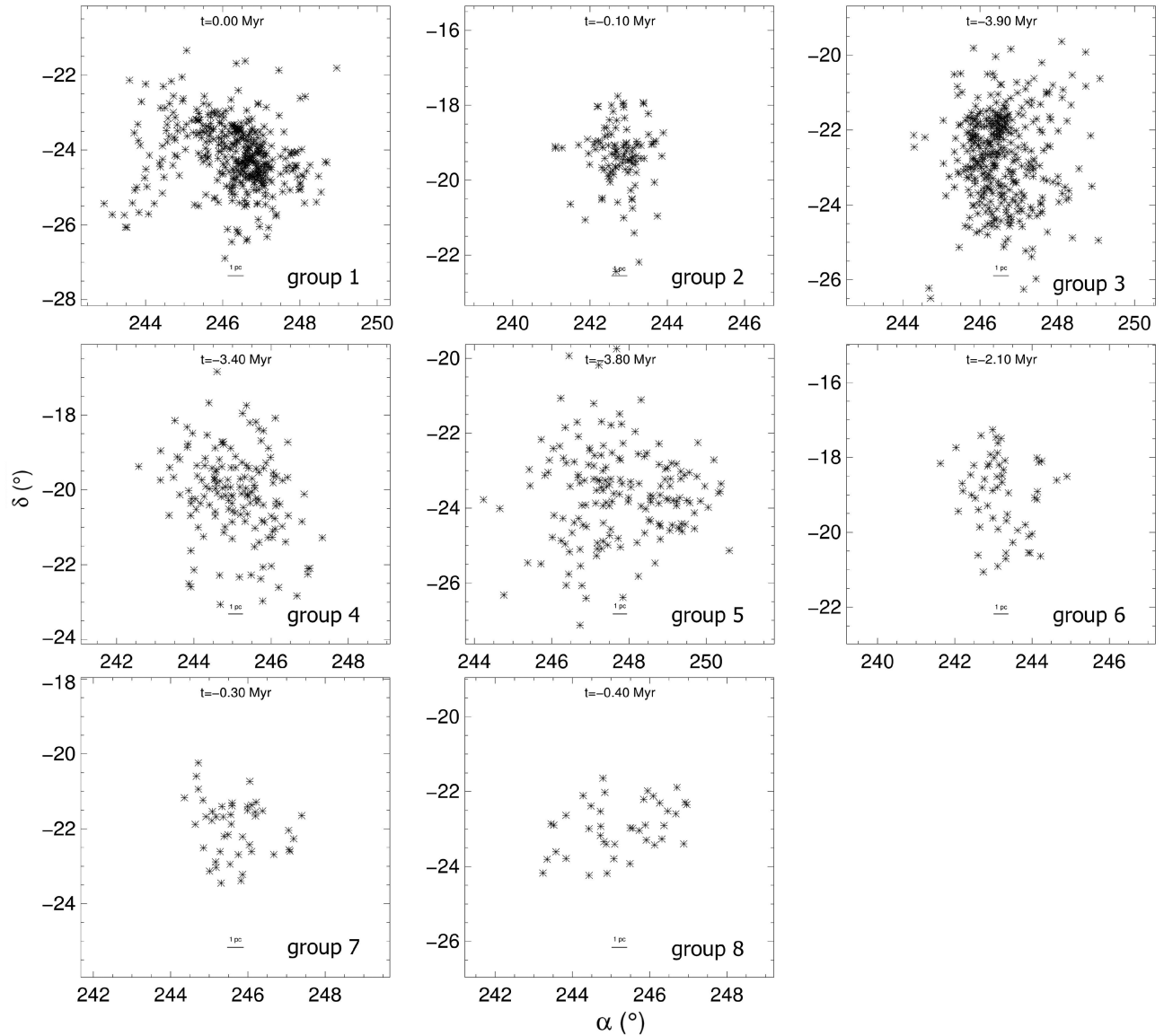
Figure 9. All groups, at their maximum coherence, span just a few degrees in sky coordinates, corresponding to projected separations of a few pc.
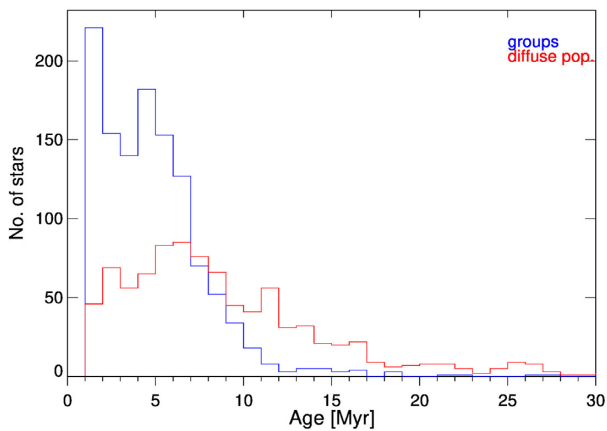

**Figure 10.** Age distribution for the clustered (blue) and diffuse (red) population.

number of false negatives, i.e. sources not labelled here as full discs but that are identified there as such, is comparable: 84/1261 (81/1261). We can reasonably assert that our disc fraction estimates are not biased in one direction or another.

The assumption of a single τ hides the different time-scales of disc decay with stellar mass: it has been shown that the lifetime of a disc steadily decreases with stellar mass (Ribas, Bouy & Merín 2015). The dependence of the disc fraction on stellar mass is shown in Fig. 11: The fractions for the first bin and the last two are in complete agreement with the dedicated IR survey performed by Luhman & Mamajek (2012) and Pecaut & Mamajek (2016).

We show in Figs 12–14 the comparison between the three estimates for the eight groups. While disc and isochronal age show a remarkable correlation (0.84), kinematic ages appears always underestimated, hinting at the presence of factors not taken into account so far. We will discuss the findings of the age analysis in Section 5.
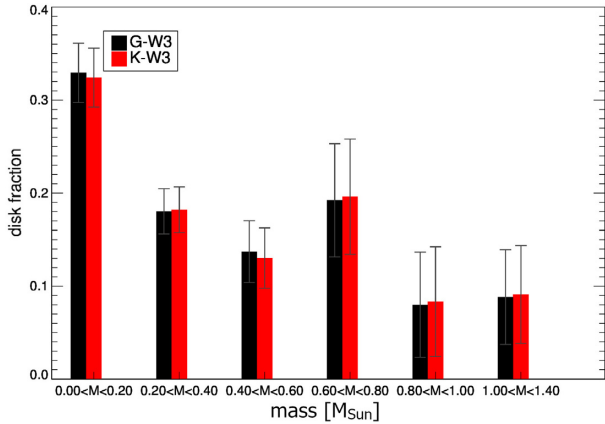
Figure 11. Disc fractions for different bins of stellar mass. Black bars give the fraction of stars where discs were detected from the $G - W3$ colour; red bars those with discs detected from the $K - W3$ colour. Going towards bins of increasing stellar mass, a clear decrease in the disc fraction is seen.
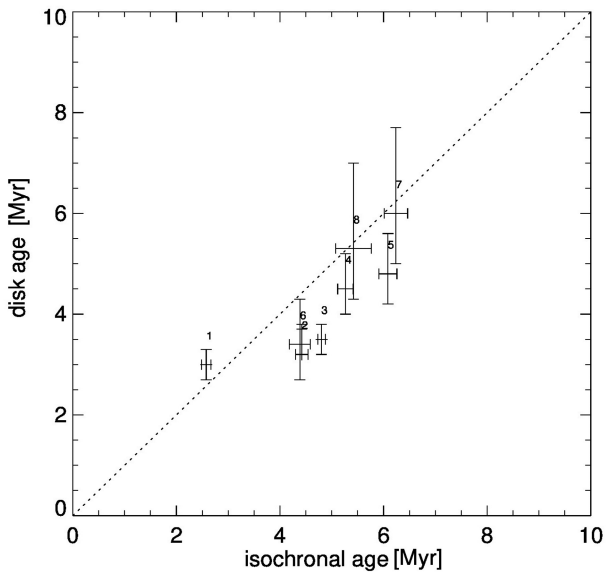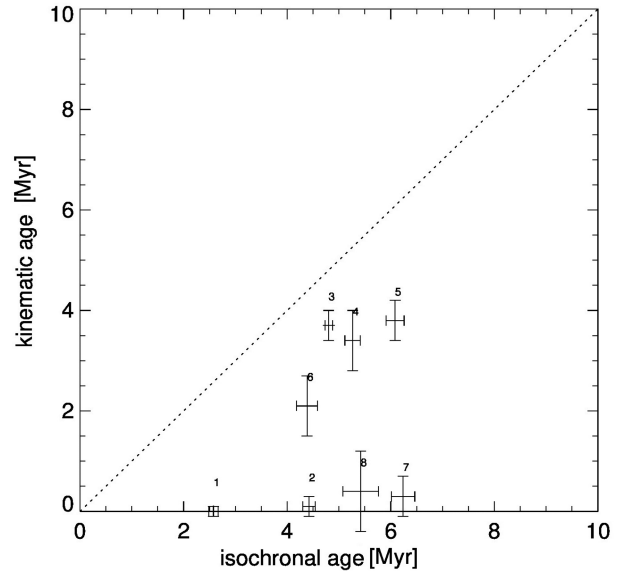

Figure 12. Isochronal versus disc ages for USCO groups. These estimates show a remarkably high correlation (0.84).

5 DISCUSSION

5.1 Kinematic analysis

As many recent studies have shown (e.g. Roccatagliata et al. 2020), the formation of associations cannot be reduced to the simple scenario of a monolithic burst. The idea of a coherent, uniform expansion has given way to that of a complex star formation history, prolonged over million of years and with a large spatial variability. While this might seem an argument against kinematic reconstruction (e.g. Pecaut & Mamajek 2016), the substructures themselves dissolve over time (Damiani et al. 2019), so they might in principle be targeted by the same approach.

The presence of a high degree of substructure within USCO has been recently quantified by González et al. (2021), confirming a long-held suspicion (e.g. Wright & Mamajek 2018). However, Damiani et al. (2019) concluded that this complex structure seen at present time cannot be unfolded and brought out into a global pattern of

Figure 13. Isochronal versus kinematic ages for USCO groups. Kinematic ages always appear to be underestimated.
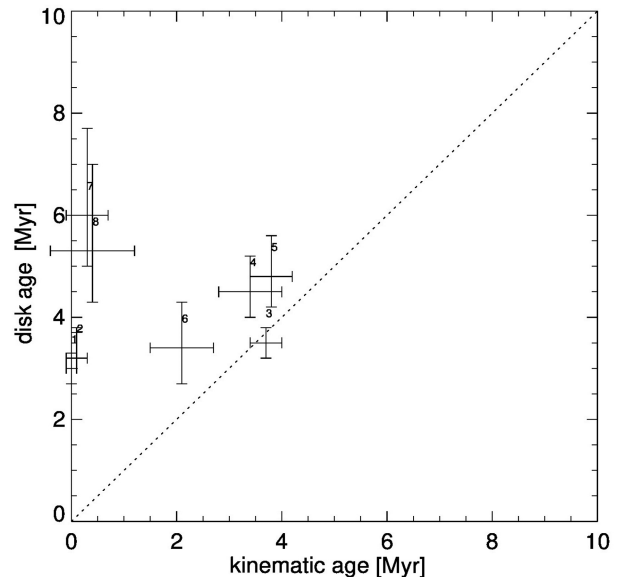
Figure 14. Kinematic versus disc ages for USCO groups. The same considerations as in Fig. 13 apply.

motion: Regions of higher and lower density do not bear distinct kinematic imprints, and close proximity on the sky is not equivalent to proximity in velocity, as if the different populations are mixed up. This is similar to what is observed in young clusters (e.g. Jeffries et al. 2014; Sacco et al. 2015; Tobin et al. 2015; Damiani et al. 2017). The general trend that denser regions appear younger than the sparser ones hints at the presence of multiple populations or, at least, of an extended star formation history.

The initial substructure of a group of stars can be irremediably altered not only by the violent relaxation following the removal of the initial gas reservoir (Lynden-Bell 1967), but also by subsequent dynamical interactions (Scally & Clarke 2002), leading to its complete disappearance, if the dynamical time-scale of the region is

shorter than its age.[14] If, like in associations, dynamical interactions may be neglected, a memory of the initial velocity pattern can be preserved for a longer time (Goodwin & Whitworth 2004), since the main factor leading to the erosion of the original velocity structure now becomes the galactic tidal field, acting on time-scales of $\sim 10^7$ yr (Wright & Mamajek 2018).

Galli et al. (2018), employing *Gaia* DR1 data of 1322 stars, found that the shape of USCO is approximately ellipsoidal; they conclude that the association is not dynamically relaxed, meaning that its shape is an imprint of its star formation history. Based on these considerations, we decided to employ a kinematic approach to see how much the initial structure of the association is still perceptible beneath its present velocity structure.

The method employed was the classical linear trace-back, in which individual star motions are traced back in time. The approach differed from the classical studies of USCO both in depth, number of stars and purposes: Contrary to earlier studies, with typical sample sizes of a few hundreds, our sample comprises about 3000 stars and did not aim at finding a single kinematic age for the whole USCO – something already excluded by Pecaut et al. (2012) – but rather at investigating its degree of substructure.

As can be imagined, the feasibility of a trace-back analysis is ultimately limited by the precision of the available velocities (Goldman et al. 2018). The need for precise RVs ($\sigma_v \ll 1$ km s^{-1}; Donaldson et al. 2017) was here satisfied by combining RV data from APOGEE and GALAH to those provided by *Gaia*: We have built a catalogue of 771 USCO sources (our '3D sample') with median velocity uncertainty $\sigma_v = 0.12$ km s^{-1}. Being the 3D sample size only $\sim$ 30 per cent of the 2D sample, the former was employed to verify and complement the results of the main analysis, carried in 2D.

Our cluster analysis found eight groups, which are seen to clump at different times at the past. The real nature of Groups 1, 2, 3, 4, 5, and 7 is extremely likely, comparing it with random clustering of sources in 2D or in 3D, while Groups 6 and 8 would need additional RV measurements to confirm their physical nature. None the less, it is worth mentioning that the disc fraction of Group 6 is higher than that of the field with a confidence level of $\sim 2\sigma$.

The results for Group 1 that should be interpreted as ρ Ophiuchi's bark, the region that is not too extinct to be undetectable in *Gaia*'s optical bands, are consistent with the study by Erickson et al. (2011), that derived a disc fraction of 27 ± 5 per cent and an isochronal age of about 3.1 Myr. When cross-matching our USCO sample with the ρ Ophiuchi sample by Cánovas et al. (2019), we find that 5 per cent of its sources belong to our Group 7 and another 5 per cent to our Group 8. The subdivision of sources between Groups 0 and 8 closely resembles that of Grasser et al. (2021), with compatible (α, δ, v_a, v_d, π) between our Group 1 and their 'Pop 1' and our Group 8 and their 'Pop 2'. The concentration of sources in the northern region of USCO ($352° < l < 355°$, $22° < b < 25°$), coming with a complex velocity structure, was already noted by previous studies (Damiani et al. 2019; González et al. 2021); here we suggest the division of those stars in three different groups (2, 4, and 6) as a possible solution to this conundrum.

A particular fruitful comparison can be done with the recent work by Kerr et al. (2021), which employs a density-based hierarchical clustering algorithm to identify substructures within young nearby associations. Contrary to their approach, based on a thorough inspection of the present phase space structure, our simpler, semi-automated method tried to incorporate in the classification scheme the time variable too. This is what led us to identify, for instance, the rather spread and populous Group 3, which was very concentrated in the past. A comparison of their table 6 with our Table 3 likely leads to associate their Group E, F, G, H, and I with our Groups 2, 7, 6, 3, and 1, respectively.

The most interesting result came when putting together the groups, and comparing them with the stars that could not be put in a group. The clustered population appears to have been more compact in the past (with a peak at $t \sim 3.5$ Myr ago, dominated by Group 3), with a confidence level on this result of 14σ. This is not true for the diffuse population, which instead shows its tightest configuration at present time. These results came as a surprise, as the age for USCO members should be 5–11 Myr, and led us to investigate whether the retrieved substructures were correlated to age gradients within the region: to this aim, we employed both isochrones for pre-MS and discs.

5.2 The age conundrum

The first robust result of the age analysis is that the clustered population appears younger than the diffuse population, in a similar way as Damiani et al. (2019) noticed with their distinction based on the present projected source density. The natural explanation for this fact is that star formation in USCO appears to have happened in small groups that disperse after a few Myr, dissolving in the field of the older population, but retaining for some time memory of their original velocity structure.

The clustered population itself shows an internal age gradient, as expected from previous works on the region (Pecaut & Mamajek 2016). The youngest group is Group 1 (ρ Ophiuchi), with a $t \sim 3$ Myr consistent with the literature (Erickson et al. 2011); Group 2, 3 and 6 are approximately coeval, while the small comoving Group 7 appears the remnant of an older formation event.

An independent age estimate was obtained through the disc fraction f_D, defined as the fraction of stars within the sample still bearing marks of a primordial disc. We decided to follow the criterion for distinction of disc classes used by other studies (Luhman & Mamajek 2012; Pecaut & Mamajek 2016), but the enormous restriction of $W4$ data, due to a combination of quality cuts and contamination from the field, made us lean towards the use of the sole criterion on $W3$. We employed as colour benchmark both $K - W3$ and $G - W3$, finding consistent results, and validated our proxy with a cross-match with Luhman & Esplin (2020). The price to pay was the impossibility to reliably identify looser disc evolutionary stages (evolved, transitional), which none the less are not used when inferring the age of a population.

We find an average full disc fraction of about 0.19 ± 0.01, consistent with previous work (Carpenter et al. 2006; Luhman & Mamajek 2012; Cody & Hillenbrand 2018; Luhman & Esplin 2020), and a well-defined trend with stellar mass consistent with Luhman & Mamajek (2012). The disc fraction is again significantly higher for the clustered (0.24 ± 0.02) than for the diffuse population (0.10 ± 0.01). Since disc statistics has been shown not to be influenced by binarity, as circumbinary discs in USCO decay with the same pace as circumstellar discs (Kuruwita et al. 2018), the robust correlation of f_D (0.84) with isochronal estimates reinforce the idea of a real age spread between the groups.

It should be underlined that, whilst recognizing full disc can be considered safe and little model-dependent, the e-folding time of

---

[14]The picture is even more complex, as the dynamical time-scale itself can change over time as a result of its density variations, with the counter-intuitive result that the presence of high degree of structure today rules out the existence of highly compact states in the past (Parker 2014).

disc decay is usually derived from comparison with isochronal ages, generating a circular argument when confronting t_K and t_P. Estimates of τ include 2.3 Myr (Fedele et al. 2010), 2.5 Myr (Mamajek 2009), 3 Myr (Ribas et al. 2014) and even 5 Myr (Richert et al. 2018) for models including magnetic-driven radius inflation; additionally, τ has been shown to depend on environmental conditions like the mass and the rotation rate of the parent molecular cloud core, and on the stellar mass itself (e.g. Lada et al. 2006; Mamajek 2009; Ribas et al. 2015). Therefore, our disc ages must be considered relative, rather than absolute.

The issue with isochronal ages is similar and, if anything, even more severe, as the history of age determinations in USCO bears witness. The first systematic studies of the region argued for a uniform age of ~5 Myr (Preibisch & Zinnecker 1999; Preibisch et al. 2002), but then Hillenbrand, Bauermeister & White (2008) showed that the ages of young stars depend on the spectral class, with low-mass star ages undestimated by 30–100 per cent and high-mass star ages overestimated by 20–100 per cent; a twofold variation of the inferred ages for intermediate and low-mass stars emerged for other clusters too (Mayne & Naylor 2008; Naylor 2009; Bell et al. 2013). Significant insights on the problem are given by binary systems for which dynamical mass estimates are available: discrepancies between ages inferred for presumably coeval components have been confirmed by Rizzuto et al. (2016) and Asensio-Torres et al. (2019). A total reassessment of the problem of age determination in USCO was carried out by Pecaut et al. (2012), who used F-type stars as a benchmark to establish a revised age of 11 ± 2 Myr. The picture was confirmed in a subsequent study (Pecaut & Mamajek 2016), alongside with a strong dependence of age estimates on spectral class, with younger ages for both K- (5 ± 2 Myr) and B-type stars (7 ± 2 Myr) than for F- (10 ± 1 Myr) and G-type (13 ± 1 Myr) stars. Interestingly enough, the derived age spread for the association was as large as 7 Myr.

While the debate on the age of USCO is still ongoing (e.g. David et al. 2019), a parallel discussion involves the above-mentioned age spread, either attributed to an extended star formation or to systematic effects inherent to models. An interesting solution has been put forward by Feiden (2016): highlighting the difficulties in modelling convection for pre-MS stars, he noticed that the effect of magnetic fields on a protostar – stronger for less massive stars – is to slow down its radial contraction along the Hayashi line; the resulting luminosity at a fixed age is higher than predicted, leading to incorrectly infer younger ages if the effect is not taken into account. By means of Dartmouth magnetic models (Feiden & Chaboyer 2012) equipped with the maximum allowed surface magnetic field strength, he found that a consistent 10-Myr isochrone could fit the observed Hertzsprung–Russell diagram (HRD) of USCO across all spectral types. The explanation by Feiden (2016), though, can create problems with moving groups with well-defined age (Tucana–Horologium association, β Pictoris; see Bell, Mamajek & Naylor 2015).

A completely different approach is that of Fang et al. (2017), who emphasize the importance of incorporating the star formation history (SFH) into the study of associations. Dividing stars by spectral type is risky because, at fixed mass, a younger star is of a later type than an older one. While low-mass stars ($M < 0.4 M_\odot$) enter the MS via the Hayashi line, intermediate stars develop a radiative core that makes them move along the horizontal Henyen track; given that the ascension toward earlier spectral types acts on a time-scale of 1–10 Myr, a population forming through a burst of comparable duration will generate spurious age differences if divided according to spectral class.

We will discuss how our results can be reconciled with this framework in Section 5.4, shifting for the moment the focus to the discussion of some biases impacting on both kinematic and photometric considerations.

5.3 Assessment of biases

A particularly tricky bias, affecting both the kinematics and isochronal ages, is that resulting from unresolved binaries. Assessing the fraction of binaries f_b in stellar populations has been the aim of several studies in the last years. A general trend is the increase of f_b with the primary mass: $f_b = 27 \pm 1$ per cent for $M = 0.075\text{--}0.6\,M_\odot$ (Winters et al. 2019), $f_b = 41 \pm 3$ per cent for $M = 0.78\text{--}1.02\,M_\odot$ (Raghavan et al. 2010), and $f_b = 54 \pm 4$ per cent for $M = 1.02\text{--}1.25\,M_\odot$ (Raghavan et al. 2010); however, environmental effects can play an important role, as shown by USCO itself that, contrary to the field, shows a fairly constant $f_b = 35$ per cent for early M and G–K stars (Tokovinin & Briceño 2020) and a value as high as ~ 70 per cent for B-A-F stars (Kouwenhoven et al. 2005, 2007).

We envisage the effect of unresolved binary stars on our sample to be twofold. As regards the kinematics, we expect that the wobbling of the photocentre is reflected in a perturbation of the system's true proper motion with the result that, when performing kinematic trace-back, binaries will be preferentially directed towards the diffuse population rather than to the clustered population. As regards the photometry, the presence of a cooler secondary star can shift the whole system towards lower temperatures and higher luminosities in the HRD. The combination of the two effects makes unresolved binaries appear younger than they are. It has recently been shown (Sullivan & Kraus 2021) that this factor is able to create a large apparent spread in a coeval population, especially when accompanied by high parallax uncertainties. None the less, the correlation that we found between the isochronal age and the disc age is reassuring, as the latter is expected to be independent of multiplicity (Kuruwita et al. 2018).

To quantify the impact of these considerations on our results, we started from the sample of 614 USCO stars constructed by Tokovinin & Briceño (2020) to assess the multiplicity in the region. Their sample, covering the mass range $[0.7, 1.5]\,M_\odot$, is invaluable for two reasons: on the one hand, it was accurately vetted not to be biased against or towards binaries; on the other hand, it was extensively studied to look for companions by means of speckle interferometry, pushing the detection sensitivity well below *Gaia*'s.[15]

A tentative diagnostic of possible binarity is the renormalized unit weight error (RUWE), a parameter associated with each source of *Gaia* EDR3 and quantifying how much the assumption of *isolated point source* is suited to the astrometric solution; values of RUWE > 1.4 are usually used as a threshold to flag a potential non-single star nature (Fabricius et al. 2021). After removing 70 companions of their list that are resolved by *Gaia* (having angular separations $\theta \gtrsim 1$ arcsec), we are left with a sample of 180 companions. We have two samples of primary stars, that we may dub 'single' (S) and 'binary' (B). We retrieved the RUWE parameter for all these stars and verified that the fraction of stars with RUWE > 1.4 (f_R) is significantly higher for sample B (76 per cent) than for sample S (19 per cent). On the ground of the expected multiplicity of USCO, we would

---

[15]The detection limits of their survey, outside $\theta \approx 0.1$ arcsec (corresponding to a projected physical distance $d \approx 15$ AU), are such that the survey is complete at a contrast $\Delta I = 2$ mag.
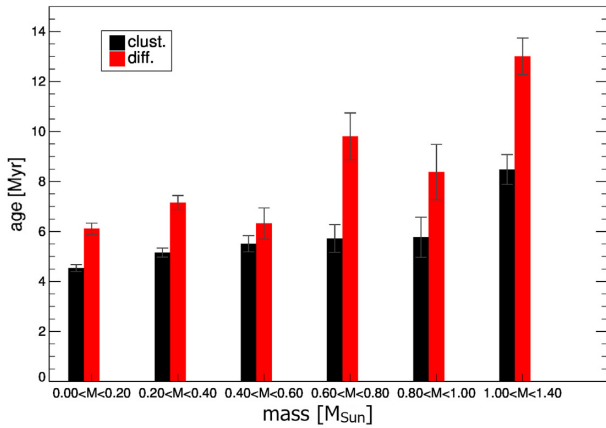
Figure 15. Average isochronal ages for different bins of stellar mass. Black bars give the average age for members of the clustered population; red bars for those of the diffuse population.

expect for our 2D sample an $f_R \sim 30$–35 per cent. Instead, we find that $f_R = 16 \pm 1$ per cent, equal for the clustered (15 ± 1 per cent) and the diffuse (17 ± 1 per cent) populations. Exceeding the natural bias against strict binaries proper to *Gaia*-based samples (Tokovinin & Briceño 2020), our selection criteria appear to have preferentially excluded binaries from the 2D sample.

To verify it explicitly, we applied the same selection criteria of Table 1. Whereas we recover 87 per cent of stars from sample S, we only recover 67 per cent from sample B. Although the median parallax uncertainty is larger (0.089 versus 0.023 mas), the effect is mostly due to velocity cuts. This basically means that proper motions of unresolved binaries can be so large to slide them out of the selection window 25 per cent of the time.

The high-RUWE stars of our clustered (diffuse) sample are significantly younger than their parent population, having a median $t = 2.6$ Myr ($t = 4.4$ Myr). To verify if this result is consistent with a population of unresolved binaries, we set up a simulation of 10 000 binary systems, with initial mass function (IMF) and mass ratio distribution (CMRD) as in Reggiani & Meyer (2011). After randomly generating primary masses according to the IMF and secondary masses according to the CMRD and the IMF, we assigned each star a set of magnitudes (J, H, G, G_{BP}, G_{RP}), coming from the same set of models used in the isochronal analysis, and a fixed age. For an age of $t = 5$ Myr (comparable to the BT-Settl result for the clustered population), the derived median age shifts to $t = 3.0$ Myr; for $t = 8$ Myr (similar to the diffuse population), it shifts to $t = 4.8$ Myr. The similar relative magnitude of the age deviation, coupled with the comparable fraction of high-RUWE stars, does not constitute an argument against the age spread between the two kinematic populations of USCO, but naturally explains the young tail seen in the diffuse one (Fig. 10).

A second issue worth considering is the already mentioned difficulty in assessing ages of low-mass stars. If we divide our sample in bins of fitted mass, we recover 119 stars with $M > 1\,M_\odot$[16]: 80 stars in the diffuse, 39 in the clustered population, with median ages 13.7 and 8.1 Myr, respectively (Fig. 15). The unequal division of mass might point to different properties of the two populations; Galli

---

[16]The smallness of the sample is due not only to the IMF but also to selection effects: Bright stars at the typical distance of USCO tend to have poorer astrometric solutions than low-mass stars.

et al. (2018) presented evidence for a somewhat different spatial distribution of the brightest and the faintest stars in USCO, hint of a different relaxation state that would imply an earlier formation of massive stars. We don't have the means to answer this question, but even a general increase of our fitted ages would not be detrimental to our main arguments.

As concerns field contamination, expected to some extent due to the rough definition of the 2D sample, it tends to create an age overestimate. Starting from the 3D sample, we applied an iterative exclusion of the stars that are more than 3σ out of the velocity distribution, up to convergence: 82/924 stars (~ 9 per cent) are excluded in this way. The median of their ages is twice as large (11 Myr) as that of the full sample. A similar fraction of interlopers is found by applying the same algorithm to the age distribution of the 2D sample: out of 2183 stars with an age estimate, those excluded are 210 (~ 10 per cent), mostly (91 per cent) belonging to the diffuse group.

5.4 The star formation history of USCO

Turning back the attention to our main point, we might notice that the fraction of stars found in the clustered population is ~ 50 per cent. This fraction, likely underestimated because of field contamination, is much higher than the 14.5 per cent found by Damiani et al. (2019) across the whole Sco-Cen, suggesting an evolutionary scenario of kinematic structures, consistent with the younger age of USCO compared to UCL and LCC. In the former, even though the present space distribution is quite well mixed, the velocity distribution is not, and still allows a distinction of subgroups; but velocity patterns are beginning to fade away, too, as confirmed by the diffuse population that has lost, or is losing, its initial kinematic imprint.

The classical picture of star formation in Scorpius-Centaurus, first put forward by Preibisch & Zinnecker (1999), envisions the first outburst of star formation (not before 30 Myr; see Damiani et al. 2019) in LCC, that triggered shortly after activity in the adjacent UCL, and finally events like the explosion of supernovae led to star formation in USCO. But this should be considered no more than a zeroth-order approximation of a more complex sequence of events. Indeed, a thorough study by Pecaut & Mamajek (2016) proved the existence of an intricate substructure in its age distribution within the subgroups themselves, reflected in the kinematic structure of the association: Wright & Mamajek (2018) concluded that there is no evidence to say that the three subgroups of Sco-Cen are the result of a global expansion of three small, independent clusters; rather, they argue for the existence of substructures, that they could not resolve due to the small sample size; such subgroups should not be the outcome of single, individual bursts of star formation, but rather of several minor subgroups, born independently of each other.

All the elements of our analysis converge towards a single solution: namely, that star formation in USCO must have lasted more than 10 Myr. After starting producing stars almost simultaneously with LCC (~ 15 Myr ago), as confirmed by the presence of stars coeval to the latter (Damiani et al. 2019), the formation continued: the existence of B-type stars with clear indication of youth like τ Sco (Rizzuto, Ireland & Kraus 2015) or HD 142184 (Tetzlaff, Neuhäuser & Hohle 2011) argues for some star formation occurred between the first outburst and the age of ρ Ophiuchi. Whether this star formation was continuous or rather experienced a late burst is a different question: In this regard, the existence of a kinematic bimodality similar, albeit less pronounced, to that of Taurus (Kraus et al. 2017), is significant, as it is intriguingly consistent with the late-burst scenario put forward by Fang et al. (2017).

The intricacy of star formation in USCO lies in the fact that it happens in small formation episodes (a few tens to hundreds stars each), as already pointed out by (Pecaut & Mamajek 2016), and that the older events have at present time lost their kinematic imprint. This is seen also in the compact populations in Lupus within LCC, where the stars are younger than the diffuse population not only of the surrounding LCC, but also of USCO (Damiani et al. 2019). The most compact population of USCO, ρ Ophiuchi, has a population of protostars still embedded in the molecular clouds – hence invisible to *Gaia* – only 0.3–1 Myr old (Luhman & Rieke 1999; Wilking et al. 2008), with external, less extinct regions spanning between 2 and 5 Myr, as we have shown. The interesting fact is that the group is under strong influence from the rest of USCO (Ladjelate et al. 2020), with positive feedback on the star formation rate in the L1688 cloud (Nutter, Ward-Thompson & André 2006), whose stars have not dispersed yet (Ducourant et al. 2017); moreover, a possible echo of its formation related to a shock wave by a supernova might still be seen as an RV gradient of $\sim$1.0 km s^{-1} pc^{-1} (Rigliaco et al. 2016). An even more significant observation is that the bubble around the eponymous star ρ Ophiuchi is inflating with a velocity $\sim$1.3 km s^{-1}, so that the time needed to inflate the bubble ($d \approx 1.36$ pc) is $\sim$1.2 Myr (Pillitteri et al. 2016). The isochronal ages of the members of this bubble is about $\sim$5–10 Myr, much higher than the kinematic age of the bubble. Pillitteri et al. (2016) speculate that the assumption of a constant expansion is not valid in the first phases, when the dense material surrounding the young stars actively acts to delay the expansion.

This conclusion is directly reflected into what we have found in our work: kinematic ages of the subgroups of USCO are always smaller than isochronal ages, even neglecting issues inherent to age determination of low-mass stars that would further raise the discrepancy. The idea that stars in USCO formed well before their group started dispersing can be related to a initial bound state of the gas-rich structures, disrupted after the dispersal of gas by stellar feedback (Krumholz 2014). Therefore, the time-scale of gas removal – quantified as 2–7 Myr for some nearby galaxies (Kim et al. 2021) – could be intriguingly estimated as the difference between isochronal and kinematic ages.

6 CONCLUSIONS

This work has shown that the star-forming region USCO can be divided into two populations, carrying a distinct kinematic imprint. The existence of a clustered population cannot be attributable to a random concentration of sources, as shown with both a 2D and a 3D kinematic analysis. While the diffuse population does not appear to have been more concentrated in the past, the clustered population shows a clear tendency towards a more compact past configuration. We have further divided the clustered population in eight groups, the most evident clustering at $\sim$4 Myr ago.

This duality within USCO is clearly reflected in the age determinations obtained through isochrone fitting: The diffuse population is characterized by a flatter age distribution than the clustered population, whose relative youth is consistent with a late burst of star formation. The fraction of stars bearing mark of primordial discs is significantly higher for the latter ($f_D = 0.24 \pm 0.02$) than for the former ($f_D = 0.10 \pm 0.01$).

Even if the absolute ages t_P and t_D provided in this work should be taken with caution due to known uncertainties in theoretical models, the relative ages are significant, and argue for a strong difference between the populations and, on a lesser extent, among the groups

themselves. The star formation history in USCO appears to have been heavily substructured, with several small events comprising at most a few hundreds of stars, and spread over $\sim$10 Myr. The kinematic structure of the association is still visible inside the youngest part of the association, but has already been erased in the oldest. The systematic differences between kinematic and isochronal ages are likely due to the time-scale of gas dispersal, intriguingly building a bridge from the early phases of star formation in molecular clouds to the final stages of star dispersal into the galactic field.

paper includes data that have been provided by AAO Data Central (datacentral.aao.gov.au). This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This publication makes use of data products from the *Wide-field Infrared Survey Explorer*, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, and NEO-WISE, which is a project of the Jet Propulsion Laboratory/California Institute of Technology. *WISE* and NEOWISE are funded by the National Aeronautics and Space Administration. This research has made use of the SIMBAD data base and Vizier services, operated at CDS, Strasbourg, France, and of the Washington Double Star Catalogue maintained at the US Naval Observatory. Finally, we would like to thank NASA's Astrophysics Data System. RG and DM acknowledge funding by the PRIN-INAF 2019 'Planetary systems at young ages' (PLATEA) and the ASI-INAF agreement no. 2018-16-HH.0. MB acknowledges funding by the UK Science and Technology Facilities Council (STFC) grant no. ST/M001229/1.

DATA AVAILABILITY

MADYS is soon to be made publicly available for use.

The data underlying this paper can be freely accessible on Vizier at http://vizier.u-strasbg.fr/viz-bin/VizieR. Queries on the *Gaia* archive can be done using the dedicated web page: https://gea.esac.esa.int/archive/ for *Gaia*.

The trace-back animations of our samples, described throughout this work, are provided as supplementary material to this paper.

REFERENCES

Ahumada R. et al., 2020, ApJS, 249, 3
Ambartsumian V. A., 1954, Liege International Astrophysical Colloquia, p. 293
Arzoumanian D. et al., 2019, A&A, 621, A42
Asensio-Torres R. et al., 2019, A&A, 622, A42
Baraffe I., Homeier D., Allard F., Chabrier G., 2015, A&A, 577, A42
Barnes A. T., Longmore S. N., Dale J. E., Krumholz M. R., Kruijssen J. M. D., Bigiel F., 2020, MNRAS, 498, 4906
Bell C. P. M., Naylor T., Mayne N. J., Jeffries R. D., Littlefair S. P., 2013, MNRAS, 434, 806
Bell C. P. M., Mamajek E. E., Naylor T., 2015, MNRAS, 454, 593
Blaauw A., 1946, Publ. Kapteyn Astron. Lab. Groningen, 52, 1
Blaauw A., 1964, ARA&A, 2, 213
Blaauw A., 1978, in Mirzoyan L. V., ed., Problems of Physics and Evolution of the Universe. Publishing House of the Armenian Academy of Sciences, Yerevan, p. 101
Brown A. G. A., Dekker G., de Zeeuw P. T., 1997, MNRAS, 285, 479
Buder S. et al., 2021, MNRAS, 506, 150
Busso G. et al., 2021, Gaia EDR3 Documentation. European Space Agency (ESA) and Gaia Data Processing and Analysis Consortium (DPAC). Available at: https://gea.esac.esa.int/archive/documentation/GEDR3/index.html
Cánovas H. et al., 2019, A&A, 626, A80
Carpenter J. M., Mamajek E. E., Hillenbrand L. A., Meyer M. R., 2006, ApJ, 651, L49
Carpenter J. M., Mamajek E. E., Hillenbrand L. A., Meyer M. R., 2009, ApJ, 705, 1646
Chen C. H., Pecaut M., Mamajek E. E., Su K. Y. L., Bitner M., 2012, ApJ, 756, 133
Churchwell E. et al., 2009, PASP, 121, 213
Cody A. M., Hillenbrand L. A., 2018, AJ, 156, 71
Cutri R. M. et al., 2012, Explanatory Supplement to the WISE All-Sky Data Release Products. WISE Science Team. Available at: https://wise2.ipac.caltech.edu/docs/release/allsky/expsup/index.html
Cutri R. M. et al., 2021, VizieR Online Data Catalog, II/328
D'Orazi V., Biazzo K., Randich S., 2011, A&A, 526, A103
Dahm S. E., Carpenter J. M., 2009, AJ, 137, 4024
Damiani F., Prisinzano L., Jeffries R. D., Sacco G. G., Randich S., Micela G., 2017, A&A, 602, L1
Damiani F., Prisinzano L., Pillitteri I., Micela G., Sciortino S., 2019, A&A, 623, A112
David T. J., Hillenbrand L. A., Gillen E., Cody A. M., Howell S. B., Isaacson H. T., Livingston J. H., 2019, ApJ, 872, 161
de Zeeuw P. T., Hoogerwerf R., de Bruijne J. H. J., Brown A. G. A., Blaauw A., 1999, AJ, 117, 354
Donaldson J. K., Weinberger A. J., Gagné J., Boss A. P., Keiser S. A., 2017, ApJ, 850, 11
Ducourant C., Teixeira R., Galli P. A. B., Le Campion J. F., Krone-Martins A., Zuckerman B., Chauvin G., Song I., 2014, A&A, 563, A121
Ducourant C. et al., 2017, A&A, 597, A90
Elmegreen B. G., Elmegreen D. M., 2001, AJ, 121, 1507
Erickson K. L., Wilking B. A., Meyer M. R., Robinson J. G., Stephenson L. N., 2011, AJ, 142, 140
Espaillat C. et al., 2012, ApJ, 747, 103
Esplin T. L., Luhman K. L., Miller E. B., Mamajek E. E., 2018, AJ, 156, 75
Evans N. J., II, 1999, ARA&A, 37, 311
Evans D. W. et al., 2018, A&A, 616, A4
Everitt B., Landau S., Leese M., Stahl D., 2011, Cluster Analysis, 5th edn. John Wiley & Sons Ltd, West Sussex
Fabricius C. et al., 2021, A&A, 649, A5
Falgarone E., Phillips T. G., Walker C. K., 1991, ApJ, 378, 186
Fang Q., Herczeg G. J., Rizzuto A., 2017, ApJ, 842, 123
Fedele D., van den Ancker M. E., Henning T., Jayawardhana R., Oliveira J. M., 2010, A&A, 510, A72
Feiden G. A., 2016, A&A, 593, A99
Feiden G. A., Chaboyer B., 2012, ApJ, 761, 30
Gagné J., Faherty J. K., 2018, ApJ, 862, 138
Gaia Collaboration et al., 2016, A&A, 595, A1
Gaia Collaboration et al., 2018, A&A, 616, A1
Gaia Collaboration et al., 2021, A&A, 649, A1
Galli P. A. B., Joncour I., Moraux E., 2018, MNRAS, 477, L50
Goldman B., Schilbach E., Röser S., Schöfer P., Derekas A., Moor A., Brandner W., Henning T., 2018, in Recio-Blanco A., de Laverny P., Brown A. G. A., Prusti T., eds, Proc. IAU Symp. 330, Astrometry and Astrophysics in the Gaia Sky. Kluwer, Dordrecht, p. 214
González M. et al., 2021, A&A, 647, A14
Goodwin S. P., Whitworth A. P., 2004, A&A, 413, 929
Grasser N. et al., 2021, preprint (arXiv:2101.12200)
Gutermuth R. A. et al., 2008, ApJ, 674, 336
Hacar A., Tafalla M., Kauffmann J., Kovács A., 2013, A&A, 554, A55
Hacar A., Tafalla M., Forbrich J., Alves J., Meingast S., Grossschedl J., Teixeira P. S., 2018, A&A, 610, A77
Haisch K. E. J., Lada E. A., Lada C. J., 2001, ApJ, 553, L153
Hillenbrand L. A., Bauermeister A., White R. J., 2008, in van Belle G., ed., ASP Conf. Ser. Vol. 384, 14th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun. Astron. Soc. Pac., San Francisco, p. 200
Hoemann E., Heigl S., Burkert A., 2021, MNRAS, in press
Janson M., Lafrenière D., Jayawardhana R., Bonavita M., Girard J. H., Brandeker A., Gizis J. E., 2013, ApJ, 773, 170
Janson M. et al., 2021, A&A, 646, A164
Jeffries R. D. et al., 2014, A&A, 563, A94
Kerr R., Rizzuto A. C., Kraus A. L., Offner S. S. R., 2021, preprint (arXiv:2105.09338)
Kim J. et al., 2021, MNRAS, 504, 487
Kouwenhoven M. B. N., Brown A. G. A., Zinnecker H., Kaper L., Portegies Zwart S. F., 2005, A&A, 430, 137
Kouwenhoven M. B. N., Brown A. G. A., Portegies Zwart S. F., Kaper L., 2007, A&A, 474, 77

Kraus A. L., Ireland M. J., Martinache F., Hillenbrand L. A., 2011, ApJ, 731, 8

Kraus A. L., Herczeg G. J., Rizzuto A. C., Mann A. W., Slesnick C. L., Carpenter J. M., Hillenbrand L. A., Mamajek E. E., 2017, ApJ, 838, 150

Kroupa P., Aarseth S., Hurley J., 2001, MNRAS, 321, 699

Krumholz M. R., 2014, Phys. Rep., 539, 49

Kuruwita R. L., Ireland M., Rizzuto A., Bento J., Federrath C., 2018, MNRAS, 480, 5099

Lada C. J., Lada E. A., 2003, ARA&A, 41, 57

Lada C. J., Wilking B. A., 1984, ApJ, 287, 610

Lada C. J. et al., 2006, AJ, 131, 1574

Ladjelate B. et al., 2020, A&A, 638, A74

Larson R. B., 1981, MNRAS, 194, 809

Larson R. B., 2002, in Grebel E. K., Brandner W., eds, ASP Conf. Ser. Vol. 285, Modes of Star Formation and the Origin of Field Populations. Astron. Soc. Pac., San Francisco, p. 442

Leike R. H., Glatzle M., Enßlin T. A., 2020, A&A, 639, A138

Lodieu N., 2013, MNRAS, 431, 3222

Luhman K. L., Esplin T. L., 2020, AJ, 160, 44

Luhman K. L., Mamajek E. E., 2012, ApJ, 758, 31

Luhman K. L., Rieke G. H., 1999, ApJ, 525, 440

Lynden-Bell D., 1967, MNRAS, 136, 101

Mamajek E. E., 2009, in Usuda T., Tamura M., Ishii M., eds, AIP Conf. Ser. Vol. 1158, Exoplanets and Discs: Their Formation and Diversity. Am. Inst. Phys., San Francisco, p. 3

Mamajek E. E., Meyer M. R., Liebert J. W., 2002, AAS Meeting Abstracts #200, p. 71.14

Mathews G. S., Williams J. P., Ménard F., Phillips N., Duchêne G., Pinte C., 2012, ApJ, 745, 23

Mayne N. J., Naylor T., 2008, MNRAS, 386, 261

Moraux E., 2016, in Moraux E., Lebreton Y., Charbonnel C., eds, EAS Publ. Ser. Vol. 80, Stellar Clusters: Benchmarks of Stellar Physics and Galactic Evolution – EES2015. EDP Sciences, Paris, p. 73

Naylor T., 2009, MNRAS, 399, 432

Nutter D., Ward-Thompson D., André P., 2006, MNRAS, 368, 1833

Ohnaka K., Hofmann K. H., Schertl D., Weigelt G., Baffa C., Chelli A., Petrov R., Robbe-Dubois S., 2013, A&A, 555, A24

Parker R. J., 2014, MNRAS, 445, 4037

Pecaut M. J., Mamajek E. E., 2013, ApJS, 208, 9

Pecaut M. J., Mamajek E. E., 2016, MNRAS, 461, 794

Pecaut M. J., Mamajek E. E., Bubar E. J., 2012, ApJ, 746, 154

Pillitteri I., Wolk S. J., Chen H. H., Goodman A., 2016, A&A, 592, A88

Preibisch T., Zinnecker H., 1999, AJ, 117, 2381

Preibisch T., Brown A. G. A., Bridges T., Guenther E., Zinnecker H., 2002, AJ, 124, 404

Proctor R. A., 1869, Proc. R. Soc. I, 18, 169

Raghavan D. et al., 2010, ApJS, 190, 1

Reggiani M. M., Meyer M. R., 2011, ApJ, 738, 60

Ribas Á., Merín B., Bouy H., Maud L. T., 2014, A&A, 561, A54

Ribas Á., Bouy H., Merín B., 2015, A&A, 576, A52

Richert A. J. W., Getman K. V., Feigelson E. D., Kuhn M. A., Broos P. S., Povich M. S., Bate M. R., Garmire G. P., 2018, MNRAS, 477, 5191

Riello M. et al., 2018, A&A, 616, A3

Riello M. et al., 2021, A&A, 649, A3

Rigliaco E. et al., 2016, A&A, 588, A123

Rizzuto A. C., Ireland M. J., Kraus A. L., 2015, MNRAS, 448, 2737

Rizzuto A. C., Ireland M. J., Dupuy T. J., Kraus A. L., 2016, ApJ, 817, 164

Roccatagliata V., Franciosini E., Sacco G. G., Randich S., Sicilia-Aguilar A., 2020, A&A, 638, A85

Sacco G. G. et al., 2015, A&A, 574, L7

Scally A., Clarke C., 2002, MNRAS, 334, 156

Skrutskie M. F. et al., 2006, AJ, 131, 1163

Stetson P. B., 2000, PASP, 112, 925

Sullivan K., Kraus A. L., 2021, ApJ, 912, 137

Tetzlaff N., Neuhäuser R., Hohle M. M., 2011, MNRAS, 410, 190

Tobin J. J., Hartmann L., Fűrész G., Hsu W.-H., Mateo M., 2015, AJ, 149, 119

Tokovinin A., Briceño C., 2020, AJ, 159, 15

Wang S., Chen X., 2019, ApJ, 877, 116

Wilking B. A., Lada C. J., 1983, ApJ, 274, 698

Wilking B. A., Gagné M., Allen L. E., 2008, in Reipurth B., ed., Handbook of Star Forming Regions. Astron. Soc. Pac., San Francisco, p. 351

Winters J. G. et al., 2019, AJ, 157, 216

Wright N. J., Mamajek E. E., 2018, MNRAS, 476, 381

Wright E. L. et al., 2010, AJ, 140, 1868

Wright N. J., Bouy H., Drew J. E., Sarro L. M., Bertin E., Cuillandre J.-C., Barrado D., 2016, MNRAS, 460, 2593

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

paper_videos.zip

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: GAIA DR2 CORRECTED BP_RP_EXCESS_FACTOR

A known problem of *Gaia*'s G_{BP} and $G_R P$ photometry is the insensitivity to variation in local background levels, affecting the derived photometry especially for faint magnitudes (Riello et al. 2018). For this reason, Evans et al. (2018) put forward the idea of a quality metric, named *BP-RP excess factor*, which is the ratio of the sum of the two fluxes and the *G*-band flux:

$$C = \frac{F_{BP} + F_{RP}}{F_G}, \tag{A1}$$

and derives its effectiveness from the behaviour of instrumental passbands and response; $C \approx 1$ for well-behaved sources. Riello et al. (2021) show that the actual distribution of C is more complex and colour-dependent, with larger expected values at redder colours, and introduce a *corrected BP-RP flux excess factor* C^* defined as

$$C^* = C - f(G_{BP} - G_{RP}), \tag{A2}$$

where $f(G_{BP} - G_{RP})$ is an appropriate piecewise polynomial function. C^* is defined in such a way that its expected value is zero for well-behaved sources, and its value can be used as a way to discriminate between sources with *good* and *bad* (G_{BP}, G_{RP}) photometry. As the standard deviation of C^* increases at fainter magnitudes, with standard deviation given by

$$\sigma_{C^*}(G) = c_0 C + c_1 G^m, \tag{A3}$$

they suggest to remove the sources such that $|C^*| > N\sigma(G)$.

As a thorough analysis of this kind is not known to the authors for *Gaia* DR2 photometry, it was chosen to follow the same line of reasoning of Riello et al. (2021), but applied to the photometry of *Gaia* DR2. To start with, we recovered the set of standard stars by Stetson (2000) that was used by Riello et al. (2021), comprising ∼200 000 stars. After reproducing their results for *Gaia* EDR3 photometry, we repeated the analysis for *Gaia* DR2. Reasoning in the same way as in their discussion, a piecewise polynomial was fitted to the data (Fig. A1). However, an additional dependence on magnitude was observed in the corrected excess, showing up as a
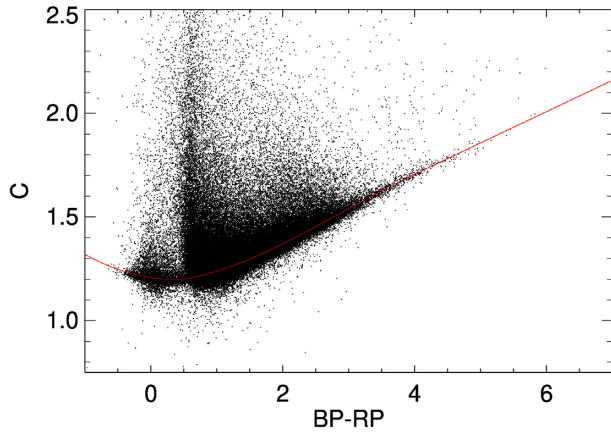
Figure A1. Dependence of the bp_rp_excess_factor on $G_{BP} - G_{RP}$ colour, using the set of standard stars by Stetson (2000). The best-fitting polynomial is overplotted in red.

Table A1. Best-fitting parameters for equation (A4).

|	$\Delta G < 0.5$	$0.5 \leq \Delta G < 3.5$	$\Delta G \geq 3.5$
a_0	$-1.121\,221$	$-1.124\,4509$	$-0.928\,8966$
a_1	$+0.050\,5276$	$+0.028\,8725$	$-0.168\,552$
a_2	$-0.120\,531$	$-0.068\,2774$	0
a_3	0	$0.007\,952\,58$	0
a_4	$-0.005\,552\,79$	$-0.005\,552\,79$	$-0.005\,552\,79$
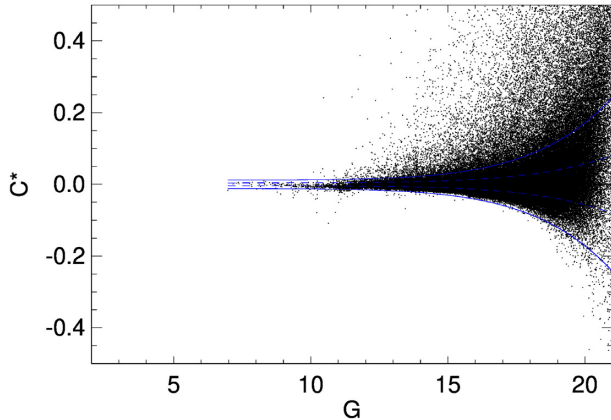
Figure A2. Trend of C^* with G for the set of standard stars. As expected, the distribution widens out at fainter magnitudes. In blue, the 3σ threshold that we applied to exclude sources with unreliable (G_{BP}, G_{RP}) photometry.

small counterclockwise rotation in the (G, C^*) plane, well fitted by a straight line.

The final equation defining the *corrected BP-RP excess factor* is given by

$$C^* = C + a_0 + a_1 \Delta G + a_2 \Delta G^2 + a_3 \Delta G^3 + a_4 G, \qquad (A4)$$

where $\Delta G = (G_{BP} - G_{RP})$; numerical values for the constants are provided in Table A1. The distribution of C^* peaks at about 0 for well-behaved sources at all magnitudes but, when considering subsamples of stars with similar brightness, it tends to widen out for fainter G (Fig. A2); a varying standard deviation $\sigma_{C^*}(G)$ can be
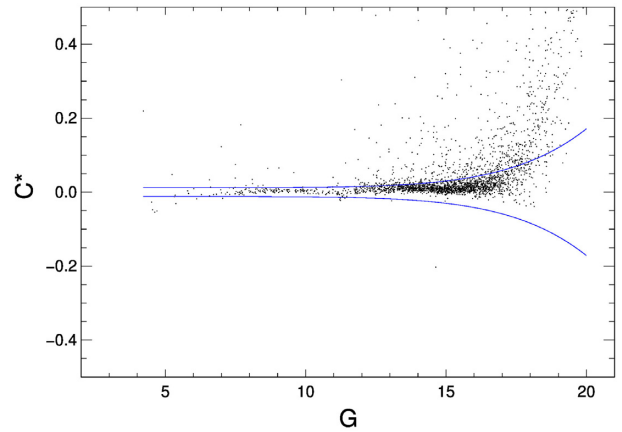
Figure A3. Same as Fig. A2, but using our 2D sample.

defined as

$$\sigma_{C^*}(G) = 0.004 + 8 \times 10^{-12} \times G^{7.55}. \qquad (A5)$$

We decided to use a cut at 3σ, as it effectively excludes stars that look visibly far from their expected positions in the $(G_{BP} - G_{RP}, G)$ diagram. The trend of C^* with G for our 2D sample is shown in Fig. A3: 889/2745 sources (32 per cent) were flagged for unreliable (G_{BP}, G_{RP}) photometry. Given that most of these stars behave well in the (G, J), (G, H) colours, this filter yields a significant improvement of the quality of isochronal age estimates.

APPENDIX B: CORRECTED PROPER MOTIONS

Any non-zero motion of an association with respect to the Sun will manifest itself as a bias in the kinematic reconstruction.

Let us suppose, for instance, that a group of stars is rigidly approaching the Sun with velocity $v_{r,c}$. If we conveniently define a Cartesian frame $(\hat{x}, \hat{y}, \hat{z})$ with one axis, let uss say $\hat{y}$, connecting the centre of the group to the Sun, we might say that all the stars have velocity $(0, v_{r,c}, 0)$. However, when we describe the position of individual stars on the sky plane, $v_{r,c}$ will be split in $(v_\alpha, v_\delta, v_r)$ depending on the relative (α, δ) with respect to the centre of the association. It is easy to verify that this results in a contribution to proper motions, that are seen to 'escape' from the centre. This is referred to as *virtual expansion*, and reflects the simple idea that the approaching association will span, as time passes, a greater extension on the sky. In a similar way, any tangential motion of the group along x and z will contribute differently to $(v_\alpha, v_\delta, v_r)$ according to (α, δ), producing a spurious velocity difference among group members.

This argument explains why, if interested in the analysis of peculiar motions within an association, one must estimate and subtract from observed proper motions the contribution coming from the bulk motion.

In order to take this factor into account, we began by looking for the centre of USCO; starting from our 3D sample, we estimated the centre at approximately $(\alpha_c, \delta_c, r_c) = (244°.55, -23°.79, 143.3 \text{ pc})$.

Then, we performed a coordinate transformation analogue to that of equation (3):

$$\begin{cases} x = r\cos(\delta - \delta_c)\sin(\alpha - \alpha_c) \\ y = r\cos(\delta - \delta_c)\cos(\alpha - \alpha_c) \\ z = r\sin(\delta - \delta_c) \end{cases} . \qquad (B1)$$

The relation among velocities is found by deriving

$$
\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} r\cos\Delta\alpha\cos\Delta\delta & -r\sin\Delta\delta\sin\Delta\alpha & \cos\Delta\delta\sin\Delta\alpha \\ -r\sin\Delta\alpha\cos\Delta\delta & -r\sin\Delta\delta\cos\Delta\alpha & \cos\Delta\delta\cos\Delta\alpha \\ 0 & r\cos\Delta\delta & \sin\Delta\delta \end{pmatrix}
$$

$$
\times \begin{pmatrix} \mu_\alpha \\ \mu_\delta \\ v_c \end{pmatrix}, \tag{B2}
$$

where we define $\Delta\alpha := \alpha - \alpha_c$ and $\Delta\delta := \delta - \delta_c$. We might write, in a compact form, $v_{\text{cart}} = A v_{\text{eq}}$. Let us consider the simplest case in which the peculiar motions are null, and the associations moves rigidly in the sky with constant (V_x, V_y, V_z). The reflection of the bulk motion in equatorial coordinates is given by $v_{\text{eq}} = A^{-1} v_{\text{cart}}$, i.e.

$$
\begin{pmatrix} \mu_\alpha \\ \mu_\delta \\ v_r \end{pmatrix} = \begin{pmatrix} \frac{\cos\Delta\alpha}{r\cos\Delta\delta} & \frac{\sin\Delta\alpha}{-r\cos\Delta\delta} & 0 \\ \frac{\sin\Delta\alpha\sin\Delta\delta}{-r} & \frac{\cos\Delta\alpha\sin\Delta\delta}{-r} & \frac{\cos\Delta\delta}{r} \\ \sin\Delta\alpha\cos\Delta\delta & \cos\Delta\alpha\cos\Delta\delta & \sin\Delta\delta \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}. \tag{B3}
$$

In order to get consistent $\mu_\alpha^* = \mu_\alpha \times \cos\delta$, μ_δ [mas yr^{-1}], and v_r [km s^{-1}], a conversion factor $p = 1000/4.74$ must be used in the equations for proper motion components.

We estimate bulk Cartesian velocities as the median velocity components of the 3D sample: $(V_X, V_Y, V_Z) = (-7.20 \pm 0.01,$ $-4.58 \pm 0.02,$ $-16.29 \pm 0.01)$ km s^{-1}, where the errors are computed with $N = 10\,000$ realizations of the same Monte Carlo approach employed throughout this work. The velocity dispersion, computed using the 16th and 84th percentiles of the distribution, corresponds to $(\sigma_{V_X}, \sigma_{V_Y}, \sigma_{V_Y}) = ([-1.98, +2.62], [-1.64, +1.61], [-1.25, +0.50])$ km s^{-1}, for a total 3D dispersion ~ 3 km s^{-1}.

To directly compare our results to those obtained by previous studies, we converted our velocities into the standard galactic right-handed frame UVW, where the origin lies in the Sun, X heads towards the Galactic Centre, Y follows the Galactic rotation, and Z is directed towards the Galactic North Pole. We find $(U, V, W) = (-4.788 \pm 0.019, -16.378 \pm 0.015, -6.849 \pm 0.016)$ km s^{-1}, comparable with both Galli et al. (2018) and Luhman & Mamajek (2012).

Subtracting the first two equations of equation (3) from proper motions from *Gaia* EDR3, we have shifted to a reference frame jointed to USCO, so that only peculiar motions are left. Projection effects of peculiar motions are not eliminated (see footnote 4), but it is the best we can achieve without possessing RVs for all the stars.

This paper has been typeset from a TEX/LATEX file prepared by the author.