Modelling the evolution of Dense Star Clusters

Gosala Narasimha Reddy,National Institute of Technology Roukela, India
under the guidance of
Prof.Alison Sills,McMaster University,Canada

The GAIA mission have unlatched a way to explore the kinematics of stars in a unprecedented way.It provides a way to shed some light on the birth and evolution of young star clusters. We used a sample of 17 young clusters which are analyzed in X-ray,IR and optical studies(MYStIX project).We used a crossmatching algorithm which compares the MYStIX sources with GAIA DR2 objects. We provided a detailed analysis of crossmatching algorithm in respect to GAIA DR2.We find that there is a regular trend in magnitudes of sources that are getting matched.We also find that the stars with high proper motions are getting rejected in crossmatching.We modified the previous AMUSE code to incorporate velocity information from GAIA DR2.

Index Terms

MYStIX, GAIA DR2, Crossmatching ,AMUSE

I. Introduction

Star clusters(SC) are a group of stars gravitationally bound to each other.This is as far as simple and a vague definition, but it tells us about the nature of star clusters. While some papers define the star cluster in terms of their properties like number of stars,velocity dispersion, mass density, crossing time. Since the star clusters dissipiate in time,they may be gravitationally bound or unbound.The unbound clusters are called associations and the bound clusters can be categorized as open and globular clusters.One of the classical statements states that all the stars form in clusters.For this to be true, the clusters have to be get dissolved in background. So the understanding of star clusters is crucial in the evolution of galaxies and in star formation. Young star clusters have very complex structure which makes it difficult for cluster identification.These protoclusters may or may not form a gravitationally bound systems.So these embedded clusters provides information about the role of gas and dust in the evolution of star clusters(section 1 in [1]).

Clusters are born in clumps in Giant Molecular clouds(GMC). These GMC obey several relations given in [2],such as the 1D velocity dispersion is to be supersonic and varies with length, the KE is nearly same as PE (which states they are nearly bound) and also the column density is approximately constant within a galaxy. The clumps in GMC are formed by supersonic turbulence and self gravity.The more massive clumps which in turn massive GMCs are expected in the gas rich regions like the formation of galaxies. The collapse of clouds results in formation of stars. The galaxy interactions and the collision between clouds causes the pressure difference which results in formation of stars. The star formation efficiency is very low.One of the best explanations states that the star forming gas is turbulent that is sufficient to unbound the most subcluster regions even though the cluster is bound in largest scales(section 3.1 in[1]).

The evolution of star clusters depends on star formation rate which is turn depends on stellar feedback mechanism which ejects gas from clusters. Protostellar outflows is one of basic and important mechanisms for feedback which helps in breaking dense regions and thus making the star formation rate low. The radiation from stars heats gas in clouds which causes them to move at sound speeds. If that gas is free,then it flows outside the cluster and thereby decreasing its mass. The zero age population stars mostly emit in UV. The ISM is highly opaque to UV and thus it applies pressure on the matter making it to move outside the cluster.Since the emitted light acts directly on matter,this phenomena is referred as Direct Radiation Pressure(DRP). Sometimes we find the regions where the interstellar dust absorbs UV radiation,then remits at infrared wavelengths. This time IR radiation gets absorbed and transfer the momentum.This processes happens for a long time.Since the force is acting indirectly here, it is called Indirect Radiation Pressure(IRP). Stars with temperatures $> 25000K$ and metallicity $Z = 0.05Z_{\Theta}$ drives winds with speeds of several thousand km/s producing temperatures $> 10^7K$ when the wind spreads into ISM. The radiative cooling sufficiently takes time which leads to formation of hot bubbles that pushes the colder gas and resulting in ejecting out of cloud. Winds are one of dominant feedback mechanisms which depends on hydrodynamic leakage of hot gas through low-density channels and reduction of thermal energy of hot gas by turbulent mixing. Supernovae(SNe) also results in ejection of material with total amount of energy that is begin generated in its life time.But this phenomena become effective in clusters of age $>$ 3Myr ,which is the time taken for a supernova to occur. For the case of smallest clusters($\leq 100M_{\odot}$),only protostellar outflows are present. For the case of medium clusters (present at the disks of galaxies), SNe, photoionization and DRP become significant.As we go to higher masses,these three phenomena become ineffective due to presence of dense matter which helps IRP to become dominant(section 3.3 in [1]).

Even after disruption of cluster by gas removal soon after the formation of cluster, there are some processes which causes disturbances in cluster either their origin is internal or external. At early ages (3-30Myr),the mass loss is mainly due

to supernova and after that ,it is due to shedding of envelopes by AGB (Asymptotic Giant Stars) stars.The latter will remain for a long time.These mass loss effects are greatly amplified by tidal fields.Over the long time scales,the relaxation becomes dominant internal process for mass loss. Stars with too much energy escape the cluster. Since the stars exchange energy by weak encounters,there may be chance of attaining sufficient energy to escape from cluster through Lagrangian points. Most of the young clusters orbit in galactic plane(not in circular orbit),they experience a change in tidal field due to close encounters with massive objects. The main source of these tidal perturbations is encounters with GMC which results in a destructive or diffusive ones based on crossing time and encounter time. The clusters that cross disk or bulge experience more tidal forces.Dynamical friction causes the clusters to move towards galactic centres by lowering their angular momentum. Massive clusters gets decelerate rapidly gets disrupted more as compared to low mass clusters(section 4 in [1]).

By above theory, we understand the processes that are meddling with evolution of clusters. In order to understand these evolution of clusters, we need to understand how the gas and stars simultaneously evolve with time by interacting with each other and also among themselves.Very few simulations allow the both the components to influence the motion of each other.Many codes have been developed to understand these processes. F.I.Pelupessy and Portegies Zwart[3] used the AMUSE code to model the spherical clusters suggests the effect of gas removal in the evolution of star clusters.In order to understand better,we need to evolve both gas and stars in simulations and also by comparing with observations.

In this report, we present the simulations of embedded protoclusters by taking into account of simulatenous evolution of gas and stars.More specially we selected the subclustered data from MYStIX. Then crossmatching the MYStIX data with GAIA DR2 data to get proper motions which are used to compute transverse velocities. We used the MYStIX positional data with kinematics of GAIA data to provide the initial conditions for simulations.

In the next section,we describe the significance of MYStIX and GAIA data in terms of sample data. We discuss the crossmatching method between MYStIX and GAIA in section 3. We discuss the theoretical aspects of simulations used in the AMUSE related to this work in section 4. We discuss the results in section 5 and future work in section 6.

II. Sample Data

A. MYStIX data

MYStIX aims to characterize 20 OB-dominated young clusters and their environments at distances $d < 4$ kpc . MYStIX project is mainly based on analysis of X-ray data from NASA's Chandra X-ray Observatory,near-infrared from UK Infrared Telescope(UKIRT) and the Two Micron All Sky Survey (2MASS), and mid-infrared observations from NASA's Spitzer Space Telescope[4].The main goal of MYStIX project is to compare the properties of different star-forming regions and young clusters in multiwavelengths. The regions that are included in MPCM(MYStIX Probable Complex Members) study are nearby high star formation sites that are included in entire MYStIX sample. A sample of 52.3 $\%$ out of all MYStIX sources were used in the analysis of spatial distributions.The finite fixture model is used for characterizing the subclusters for MSFRs. The studies of these subclustered structures are well studied in [5].The positional data(RA,DEC) we used in this report is found in [5].For the purpose of positional errors,we used the data from the paper[6]. .

B. GAIA data

GAIA is a space observatory of European Space Agency(ESA) launched in 2013. GAIA allows us to observe the positions with high accuracy precisions to about 1.7 billion sources and parallaxes and proper motions to about more than 1 billion sources reaching magnitude $G \approx 20.7$. The summary of astrometric,photometric and survey properties for GAIA DR2 is given in [7].As compared to GAIA DR1,DR2 provides 48 $\%$ additional sources,parallaxes and proper motions with an accuracy for 77 $\%$ of all sources.

In this report, we took the positional data from the MYStIX paper and then crossmatched with GAIA data .We consider only the GAIA objects which are having all four parameters(ra(α),dec(δ),pmra(μ_α),pmdec(μ_δ)).The proper motion of an object is an apparent motion of object across the sky.It is a measure of observed changes in the apparent places of stars and other objects,as seen from centre of mass of Solar system with respect to the background of more distant stars.The components of proper motion in equatorial coordinate system are given in the direction of right ascension(μ_α) and declination(μ_δ).Their total value is given by proper motion(μ).Parallax is difference in the apparent position of an object viewed from two different line of sights.In astronomy,parallax is defined as semi-angle of inclination between two sight-lines to the star, as observed when Earth is on opposite sides of the Sun in its orbit.The algorithm used for crossmatching is given in next section.

III. Crossmatching Algorithm

Crossmatching the sources in different catalogues is very useful in astronomical research.Since it provides a way to understand the same source in different wavelengths and also using different parameters.The crossmatching process is complex and challenging problem when dealing with large data sets and in different wavelength bands. The criteria for to use crossmatching is also sometimes depends on objectives of crossmatch.

The simplest matching method uses the stars' positions and use a nearest neighbour approach with a maximum cutoff distance when matching stars between two catalogues.Within the critical separation , two stars in two catalogues whose closest star in the other catalogue is each other will be assigned as a match.This nearest neighbour matching scheme didn't consider possibility that closest match may not be a counterpart.Also it may be possible that the source in question may have properties that place it outside the nearest distance.If one telescope has high resolution,it will detect more number

of sources in same region than other telescope.It results in placing many objects in the vicinity of actual source,thus affecting the crossmatching method. Since the observations of two telescopes are taken as different times,the objects had moved in the sky in the meantime. Sometimes we have to crossmatch the catalogues taken in different wavelengths.The objects which are bright in UV may not be bright in visible or IR,so sometimes it happen that the object which is present in UV image might not be present in visible image.These all problems are not considered in Nearest neighbour method.In order to reduce these problems,we use the following method.

The purpose of this crossmatching algorithm is to minimize the number of mismatches and to achieve the most probable counter part. The algorithm we defined is mainly based on GAIA Dr1 [8],GAIA DR2[9] and Pineau et al (2011)[10] with a slight modification.We consider the MYStIX catalogue as leading catalogue and GAIA catalogue as secondary catalogue. A good neighbour for a MYStIX source is a nearby object in GAIA catalogue whose position is in agreement with source. If only one good neighbour is found, then we say it is a counterpart. When more than one neighbour is found, the best neighbour is determined among the good neighbours based on Figure of merit(FOM). Due to presence of significant angular separations between GAIA and MYStIX objects,there is huge difference of nearly six orders in positional errors.As a result of these errors, the FOM of some objects become extemely small. For these cases,we used the magnitudes to determine the best neighbour. Due to high resolution of GAIA,we find many of stars in GAIA have the same counterpart in MYStIX. It is a positional algorithm which uses mainly source positions, positional errors,their correlation and proper motions.For this work,we used only GAIA proper motions.The FOM is mainly based on the probability of finding optical counterpart which uses the angular distance between the MYStIX source and GAIA object and on their positional errors.

The use of magnitudes and colors requires transformations between photometric systems.Using magnitudes for GAIA and MYStIX is kind of difficult since they both operate in different wavelengths.So in this algorithm we didn't use the magnitude as primary criteria but uses it in defining FOM.The surveys are taken at different epochs and sources may move in meantime,so it is necessary to take proper motion into account.We moved the MYStIX objects to move along the possible matches using individual epochs using the concepts in Hipparcos and Tycho catalogue(ESA 1997).We used the simplified treatment present in section1.5.4 in [11].

$$\alpha = \alpha_0 + (T - T_0)\mu_{\alpha*0} \sec \delta_0$$

$$\delta = \delta_0 + (T - T_0)\mu_{\delta 0}$$

where α and α_0 are right ascensions of MYStIX and GAIA objects,δ and δ_0 are declinations of MYStIX and GAIA objects,$(T - T_0)$ is the epoch difference,$\mu_{\delta 0}$ is the proper motion along declination and $\mu_{\alpha*0}$ is the proper motion along the right ascension.

We are using GAIA data having all the four parameters ra,dec,pmra and pmdec.Since MYStIX data is taken from four different telescopes described in section 2.1,the epoch differences between GAIA and each telescope may be different.Since the MYStIX data is taken from four telescopes(Chandra,UKIRT,Spitzer and 2MASS),it is not easy to find a particular epoch difference.We know the Epoch difference for 2MASS and GAIA is 17.7 years.Chandra is launched in 1999,so the epoch may not be greater than 20. For the sake of simplicity,we used a epoch difference of 18 years. Crossmatching is not only one to one matching but also depends on number density which comes into picture when defining FOM.

A. Details

The crossmatching algorithm uses plane-sweeping technique which requires the second catalogue or GAIA catalogue to be sorted by declination.This process speeds up the computation analysis.We used the double filter technique. A first filter is defined by a large radius centered on MYStIX source which is different for each source .This is used to select neighbours and calculate local number density for each MYStIX source.A second filter of less radius is used to select good neighbours among neighbours.And finally best neighbour or neighbours is selected based on FOM.If there are more than one best neighbour,then we select the most probable one based on magnitudes.The angular distances between MYStIX and its counterpart are calculated using Haversine formula.

1) Initial Search Radius: The initial search radius depends on position of MYStIX(m) source in GAIA(g) catalogue.It is defined around each MYStIX object as

$$R_m = H_\gamma PosErr_{m,max} + \left(\frac{PM * EpochDiff}{1000}\right)$$

where $H_\gamma = 3$ corresponds to a confidence level γ of 0.9973002038;$PosErr_{m,max}$ is the combined position error for each Mystix source with maximum positional error in Gaia catalogue which is given by

$$PosErr_{m,max} = max\Big[RAerr_m, DECerr_m\Big] +$$

$$max\Big[max\big(RAerr_g\big), max\big(DECerr_g\big)\Big]$$

where $RAerr$ and $DECerr$ are uncertanities in right ascension and declination.$EpochDiff$ is the maximum reference epoch difference between MYStIX source and Gaia object and is given by

$$EpochDiff_{m,max} = max\Big[|max\big(refEpoch_m\big) - min\big(refEpoch_g\big)|,$$

$$|min\big(refEpoch_m\big) - max\big(refEpoch_g\big)|\Big]$$

And finally PM is the maximum of the Gaia catalogue proper motions.Here R_m is in arcsec, $PosErr_{m,max}$ in arcsec,PM in mas/yr and $refEpoch$ is in years.

2) Good neighbour selection[10]: We consider selecting counterparts taking into account of arbitrary error ellipses on the source's spherical coordinates.Lets consider a MYStIX source with α_m, δ_m the coordinates of MYStIX source and $\sigma_{\alpha_m}, \sigma_{\delta m}$ and $\rho_{m_{\alpha,\delta}}$ are errors on $\alpha_m \cos\delta_m$ and on δ_m and correlation between their errors, and a GAIA counterpart with α_g, δ_g the coordinates of GAIA source and $\sigma_{\alpha_g}, \sigma_{\delta g}$ and $\rho_{g_{\alpha,\delta}}$ are errors on $\alpha_m \cos\delta_m$ and on δ_m and correlation between their errors,respectively.

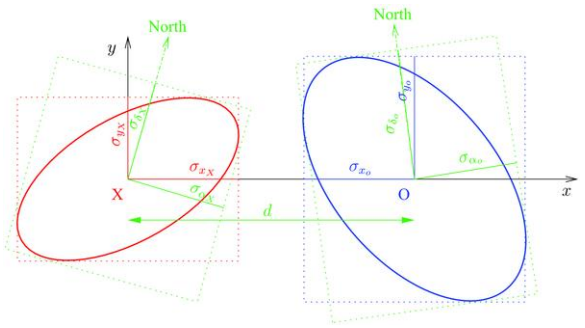

Fig. 1: Chosen projection plane: the xy frame is centred on the MYStIX source position X; the x-axis is the direction towards the Gaia candidate, located at point O. d is the angular distance between the two sources.

We convert the spherical problem into a plane one and assuming the positional errors follows 2D gaussian.We chose the projection on a 2D plane with a frame centring on MYStIX source and placing the GAIA counterpart on x-axis at a distance d. Then the errors on positions become gaussian distributions:

$$N_m\Big(x, y; \sigma^2_{x_m}, \sigma^2_{y_m}, \rho_m \sigma_{x_m} \sigma_{y_m}\Big)$$

$$N_g\Big(x, y; \sigma^2_{x_g}, \sigma^2_{y_g}, \rho_g \sigma_{x_g} \sigma_{y_g}\Big)$$

where d is the angular separation between the MYStIX source and GAIA counterpart.The density of probability that two sources are at same location,so the same object is given by convolution of two distributions

$$N_c\Big(x, y; \sigma^2_{x_c}, \sigma^2_{y_c}, \rho_c \sigma_{x_c} \sigma_{y_c}\Big) = f(x, y) =$$

$$\frac{1}{2\pi \sigma_{x_c} \sigma_{y_c} \sqrt{1-\sigma_c^2}}$$

$$exp\left[-\frac{1}{2(1-\sigma_c^2)} \left(\frac{x^2}{\sigma^2_{x_c}} + \frac{y^2}{\sigma^2_{x_c}} - \frac{2\rho_c xy}{\sigma_{x_c}\sigma_{y_c}} \right) \right]$$

where

$$\sigma^2_{x_c} = \sigma^2_{x_m} + \sigma^2_{x_g}$$

$$\sigma^2_{y_c} = \sigma^2_{y_m} + \sigma^2_{y_g}$$

$$\rho_c \sigma_{x_c} \sigma_{y_c} = \rho_m \sigma_{x_m} \sigma_{y_m} + \rho_g \sigma_{x_g} \sigma_{y_g}$$

The variance -covariance matrix of $N_c$is is given by

$$\begin{pmatrix} \sigma^2_{x_c} & \rho_c \sigma_{x_c} \sigma_{y_c} \\ \rho_c \sigma_{x_c} \sigma_{y_c} & \sigma^2_{y_c} \end{pmatrix}$$

By using the eigen decomposition of the variance-covariance matrix of N_c, we define σ_M and σ_m as the semi-major and semi-minor axis in the eigenvector frame (x_1, y_1).

$$\sigma_{M,m} = \frac{1}{2}\left(\sigma^2_{x_c} + \sigma^2_{y_c} \pm \sqrt{\left(\sigma^2_{x_c} - \sigma^2_{y_c}\right)^2 + 4\left(\rho \sigma_{x_c} \sigma_{y_c}\right)^2} \right)$$

Then we change into polar coordinates(r,θ) and integrating over θ, the density of probability can be written as a Rayleigh distribution

$$f(r) = r exp\Big(-\frac{1}{2} r^2 \Big)$$

where r is a dimensionless parameter defined as

$$r = \sqrt{\frac{x_1^2}{\sigma_M^2} + \frac{y_1^2}{\sigma_m^2}}$$

The probability density function depends on x,y only through exponent component

$$-\frac{1}{2(1-\sigma_c^2)} \left(\frac{x^2}{\sigma^2_{x_c}} + \frac{y^2}{\sigma^2_{x_c}} - \frac{2\rho_c xy}{\sigma_{x_c}\sigma_{y_c}} \right) = K_\gamma^2$$

or

$$\begin{pmatrix} x \\ y \end{pmatrix}^t \quad \begin{pmatrix} \sigma^2_{x_c} & \rho_c \sigma_{x_c} \sigma_{y_c} \\ \rho_c \sigma_{x_c} \sigma_{y_c} & \sigma^2_{y_c} \end{pmatrix}^{-1} \quad \begin{pmatrix} x \\ y \end{pmatrix} = K_\gamma^2$$

These are the equations of ellipses which define confidence regions,which are the lines of constant probability density; where K_γ^2 has a χ^2 distribution with two degrees of freedom.If K_γ is equal to a critical value of the χ^2 distribution,the probability that (x,y) will fall within the ellipse is equal to confidence level γ

$$P\left[-\frac{1}{2(1-\sigma_c^2)} \left(\frac{x^2}{\sigma^2_{x_c}} + \frac{y^2}{\sigma^2_{x_c}} - \frac{2\rho_c xy}{\sigma_{x_c}\sigma_{y_c}} \right) \le \chi^2_{2,\alpha} \right] = 1 - \alpha = \gamma$$

where α is the probability that (x,y) will fall outside the ellipse. Good neighbours are defined as neighbours that fall within the ellipse defined by confidence level γ as

$$-\frac{1}{2(1-\sigma_c^2)} \left(\frac{x^2}{\sigma^2_{x_c}} + \frac{y^2}{\sigma^2_{x_c}} - \frac{2\rho_c xy}{\sigma_{x_c}\sigma_{y_c}} \right) \le K_\gamma^2$$

From the figure of both catalogues,we have the coordinates of Gaia catalogue as x=d and y=0,then the above equation becomes

$$\frac{d}{\sigma_{x_c}\sqrt{1-\sigma_c^2}} \le K_\gamma$$

where d is the angular separation given by

$$d = 2 \arcsin\left(\sin^2 \frac{\delta_g - \delta_m}{2} + \sin^2 \frac{\alpha_g - \alpha_m}{2} \cos\delta_g \cos\delta_m \right)^{1/2}$$

Also σ_{x_c} is the convolution ellipse error in the direction from MYStIX object to the possible counterpart,ρ_c is the correlation between σ_{x_c} and σ_{y_c} and K_γ depends on the confidence level γ:

if $\gamma = 0.99730002038, K_\gamma = \sqrt{11.8290}$;
if $\gamma = 0.9999366575, K_\gamma = \sqrt{19.3448}$;
if $\gamma = 0.9999994267, K_\gamma = \sqrt{27.6310}$.

The completeness we have chosen is a 3σ criterion. In the MYStIX data,the positional errors are circular and the errors are given by

$\sigma_x = \sigma_y = PosErr/\sqrt{2}$

$\rho\sigma_x\sigma_y = 0$

In the Gaia data,the positional errors are elliptical and the errors are given by

$\sigma_x = raError$

$\sigma_y = decError$

$\rho\sigma_x\sigma_y = 0$

3) Best Neighbour Selection:Figure of Merit: The definition of figure of merit is taken from Pineau et al(2011)[10].We did not consider any prior knowledge like magnitude,number of possible candidates,but we use magnitude when FOM is unable to find a counterpart.The FOM is defined as ratio between the possibility of a counterpart to a random object that is just found by chance.It depends on angular seperation,position errors,epoch difference and on local surface density.

Let the probability of finding a counterpart at a distance between r and r+dr is given by

$$dp(r|cp) = rexp\Big(-\frac{1}{2}r^2\Big)dr$$

Let the probability of finding a spurious object is $dp(r|spur)$.It is evaluated using poisson distribution.The Poisson probability of finding one or more objects by chance in an infinitesimal annulus area is

$$dp(r|spur) = \sum_{k=1}^{\infty}\frac{s^k}{k!}exp(-s) =$$

$$\sum_{k=0}^{\infty}\frac{s^k}{k!}exp(-s) - exp(-s) = 1 - exp(-s) \approx s$$

where $s = \rho\sigma_M\sigma_m dA \approx \rho' 2\pi r dr$ is the number of sources within an area dA.The local density ρ is defined by counting number of objects within the initial search radius.It is better to take local density rather than global density since it will give you a better precision.

The figure of merit(FOM) is thus defined as

$$FOM\big(r\big) = \frac{dp(r|cp)}{dp(r|spur)}$$

By using the equations,we write FOM as

$$FOM(r) = \frac{1}{2\pi\rho}exp\Big(-\frac{1}{2}r^2\Big)$$

We have used all the information related to position like positions,their errors and correllations between them.In order to access the FOM values in reasonable ranges, we apply asinh to FOM(r) defined above:

$$score = asinh(FOM(r)) = asinh\Big[\frac{1}{2\pi\rho}exp\big(-\frac{1}{2}r^2\big)\Big]$$

The best neighbour is defined as the good neighbour with the highest score value.Even after this we find some of sources may have more than one counterpart.For these cases we used the J-magnitude of MYStIX sources and r-magnitude of Gaia sources.For the cases where score is same,we chose the best neighbour which is having the smallest difference between J and r magnitudes.

Cluster (1)	N_M (2)	N_G (3)	N_{G_p} (4)	N_T (5)	N_1 (6)
Orion Nebula	1524	2315	1766	670	392
Flame Nebula	484	537	391	151	93
W 40	426	538	429	109	64
RCW36	308	625	567	94	63
NGC 2264	1173	8433	6939	653	560
Rosette	1730	50713	42350	1285	1176
Lagoon	2056	41006	24362	1600	1342
NGC 2362	491	20827	18638	471	454
DR 21	979	3844	3381	612	400
RCW38	886	8926	7736	371	237
NGC 6334	1663	17342	10918	737	530
NGC 6357	2235	28131	18101	1175	818
Eagle nebula	2574	23299	16953	1918	1469
M 17	1977	13626	9569	719	322
Carina nebula	7334	179845	150615	5488	4613
Trifid Nebula	532	30807	22145	448	398
NGC 1893	1275	8077	6123	949	726

TABLE I: Crossmatching results between MYStIX star forming regions and GAIA.Column 1:Target region name.Column 2:Number of MYStIX stars (Kuhn 2014[5]).Column 3:Number of all Gaia stars present in region determined by MYStIX cluster.Column 4:Number of Gaia stars containing ra,dec,pmra and pmdec.Column 5:Number of best neighbours in Gaia catalogue.Column 6:Number of good neighbours in Gaia catalogue with exactly one best neighbour.

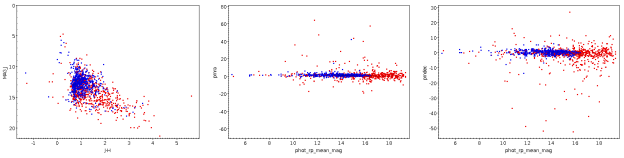

Fig. 2: Left:J vs J-H color-magnitude diagram for sources in Orion Nebula.All the MYStIX sources are shown in red and matched GAIA sources are shown in blue.Center:Proper motion measurement in right ascension direction vs r-band magnitude.Right:Proper motion measurement in declination direction vs r-band magnitude

The number of stars matched per cluster varies from 25 % (W 40) to 95.9 % (NGC 2362).This variation of matching indicates that the sources which are present in MYStIX aren't present in GAIA.It means that the clusters contains more sources that are absent in visible band.In the case of W40,dust from the molecular cloud obscures the nebula,making it very difficult to observe in visible region[12].The X-ray,infrared and radio observations easily pass through these regions so we find these regions in MYStIX data easily rather than in GAIA data.In the case of NGC2362,the diffuse cloud gets dispersed by the stellar winds from the massive young stars like Tau Canis Majoris. We also find that number of stars within the search region of cluster is more in GAIA than in MYStIX.One of the reasons is due to high angular resolution of GAIA as compared to MYStIX (four telescopes).

In figure 2,the left panel shows a near-infrared color–magnitude diagram for both the full catalog of

MYStIX in Orion Nebula and the matched stars with the Gaia. We find there are only few sources that got matched with GAIA at the right down of CMD.The center and right panel shows the proper motion measurements as a function of r-magnitude. We find that the sources which are present within certain range of magnitudes got matched with GAIA sources.Also we find that the sources with high proper motions are also getting rejected in crossmatching.

IV. Simulations

Astrophysical simulations have become a vital tool to understand the formation and evolution of astrophysical systems.Problems like planet formation, stellar evolution, formation of galaxies, structure formation in the universe have benefited from development of codes that can simulate physics involved.Most of the codes that are developed for a particular problem like on stellar evolution,gravitational dynamics,hydrodynamics,but the astrophysical problems most often require the codes to work for a wide variety of regions.For example,understanding the ejection of heated gas from an embedded star cluster born during a galaxy merger requires simultaneous treatments of collisional stellar dynamics(protostellar cluster), collisionless stellar dynamics(galaxy merger), hydrodynamics(interstellar gas),stellar evolution(stars) and radiative processes(feedback between stars and interstellar gas).This problem is not only in astrophysics but also in other fields of science. The processes that happen in astrophysics also spans a wide range of temporal and spatial scales.Instead of developing new algorithms to solve all problems simultaneously,there is an another method which is that each of the physical domains solved separately, then combined their solutions with certain time intervals. Astrophysical Multipurpose Software Enivironment(AMUSE) is an astrophysical software that presents an interface and simplified access to astrophysical codes allowing for multiphysics simulations.AMUSE allows user to bind existing numerical codes into a easy understandable language(chapter 1 in[13],[14]).

For the simulations of the embedded dense star clusters for this report,we combine the gravitational dynamics code with hydrodynamics code.In the next section,we will give the description of basic principles of hydrodynamics and gravitational dynamics that are necessary for simulations.

A. Theoretical aspects and methods

1) Hydrodynamics: Hydrodynamics deals with the motion of fluids driven by internal and external forces due to pressure gradient and gravity. Hydrodynamics plays an important role in many astrophysical processes like for understanding of star and planet formation, accretion discs, etc. AMUSE considers only the continuum description of the gas ignoring the microscopic behaviour of its constituent particles and defining of thermodynamic properties like density,temperature on macroscopic scales.Some astrophysical processes like jets,supernova,stellar winds have symmetries that allow us to decrease to a one dimensional problem.A astrophysical fluid is defined as group of small elements which has position, velocity and thermodynamic properties such as density, composition, temperature,entropy(chapter 5 in [13]).

The dynamics of fluid can be described by three independent partial differential equations which are conservation of mass,momentum and energy[15]. The conservation of mass leads to continuity equation which is given by

$$\frac{\partial \rho}{\partial t} + \nabla.(\rho v) = 0$$

By applying Newton's II law to a fluid element,we get

$$\rho \frac{dv}{dt} = F$$

By considering the main forces in astrophysics which are pressure ($F = -\nabla P$) and gravity ($F = \rho g$),we get euler equation

$$\frac{dv}{dt} = -\frac{\nabla P}{\rho} + g$$

Finally If the gas heats at a rate H and cools at a rate C ,then rate of thermal energy per unit mass is

$$\frac{de}{dt} = H - C - \frac{P}{\rho}\nabla.v$$

where ρ is the density of fluid,v is the velocity of fluid and e is the internal energy per unit mass.

Hydrodynamics codes in AMUSE can be divided into two main categories:Lagrangian methods and Eulerian methods.In Lagrangian description of flow,the fluid is described as particles.We follow the motion of fluid as described by euler equation.In eulerian method,the equations of motion are solved on a fixed grid.Particle and grid based methods are combined by moving mesh methods that solve the fluid dynamical equations on unstructured grids that evolve with fluid(chapter 5 in [13]).

We used the Smoothed-Particle Hydrodynamics(SPH) method, a type of Lagrangian method for the hydrodynamical simulation.In this approach, the fluid is described as particles. Each particle represents a finite mass of fluid;density and other thermodynamic quantities are computed by kernel estimation from nearby particles.Because each particle may be thought of smoothed out a region ,these methods are referred as Smoothed Particle hydrodynamics(SPH).

2) Gravitational Dynamics: The equation of motion for a self-gravitating system[13] is given by

$$a_i \equiv \ddot{r}_i = G\sum_{j\neq i}^{N} m_j \frac{r_j - r_i}{|r_j - r_i|^3}$$

A simple N-body code is works as follows:We initialize the model which determines the initial positions and velocities of particles in space.We calculate the forces between the particles. Determine the number of steps based the given time step or interval.Then update the new velocities based on forces,time step and old velocities.Then update the positions based on old positions,new velocity,force and time step.The above procedure is as far as simple one.

The N-body codes are characterized into three types:Pure N-body codes, Direct N-body and approximate methods.Pure N-body codes solve the Newton's equations of motion with

no free physical parameters,but includes the close encounters, binary dynamics.Ph4 and Hermite codes comes under this scheme.Direct N-body codes solve the Newton's equations of motion with additional parameters sometimes used to speed up the code.Nbody6 code comes under this method. Approximate methods relaxes the complexity for computing force.These methods are fast.BH-Tree,MI6,Bonsai comes under this method(chapter 2 in [13]).

We used Hermite scheme method to solve gravitational dynamics.

3) Bridge technique: A straight forward way to couple codes is by directly propagating information from one code to another code.A more general procedure for coupling two or more simultaneously running codes is by using Bridge technique.When two dynamical systems interact,one may be a microscopic system (smaller time scale) and other may be a macroscopic system (larger time scale).In many cases, the spatial and temporal scales of the microscopic and macroscopic systems are well separated.So we can split the underlying Hamiltonian for the combined system into rapidly and slowly changing parts. Microscopic and macroscopic evolution can then be followed simultaneously by adopting a kick–drift–kick algorithm, allowing both systems to be evolved separately,then the results are combined into a consistent solution.We can give different methods for both the systems.The technique is called Bridge method(chapter 4,7 in [13]).

We used this technique in coupling N-body code and Hydrodynamics code.

B. Initial conditions

The MYStIX observations gives the right ascension and declination for each star observed in the region.They also give the number of subclusters, semimajor axis, semiminor axis and position angle of ellipse on the sky.There are some stars which are present as a background are called Unclustered stars.Also there are some stars which belongs to neither clustered ones nor background ones are classified as Unknown stars.

1) Positions of stars and gas particles: For the positional data for stars, we used the same procedure as described in the paper[[16]].For all the observed stars(including subcluster stars,unclustered and unknown stars),we used MYStIX's positional data (RA and DEC). We fixed the z-axis by a randomly distributed number from a uniform distribution of reasonable distances(0-1 pc). MYStIX observes stars which are more massive than $0.83\ M_{\odot}$.So we added the stars of masses (0.2-$0.83 M_{\odot}$).In Getman 2015[17],they estimated the low-mass stars by populating IMF normalized by number of observed stars.For this we created a plummer sphere with a core radius which is given by harmonic mean of the observed major and minor axes of the observed ellipses.We then multiply these plummer sphere positions with some factor to fit in the ellipses and using PA of the ellipses to get correct orientation.We also added the low-mass stars in unclustered regions doing the same way as for low mass stars in subclusters.We assigned the z range for these low mass stars to be in comparable with those of massive stars.

We add the gas to the simulations after the stellar positions are determined.We added gas by using a plummer gas model which is a stretched plummer sphere with same axis ratio and PA as stars for a subcluster.Then we added the background gas by a plummer model of highly extended axis ratio.

2) Velocities for stars and gas particles: We modified the code in the paper[[16]] to include the velocity information. We used the observed velocity information for the stars that are present in both MYStIX and GAIA catalogues(from now we call Matched stars).This includes all observed stars(inside the subclusters, unclustered and unknown stars) that got matched with GAIA .For all the unmatched stars, we assigned the velocity drawn from a normal distribution of mean and standard deviation for the velocities of matched stars.Since we are using observed velocities, we neglected the relative velocities between the sub clusters.For all the stars, we assigned the velocity along z-axis by a random number drawn from a uniform distribution of -1 to 1 km/s.

For all gas particles,we assigned the velocity drawn from a normal distribution of mean and standard deviation for the velocities of matched stars.Similarly,we assigned the velocity along z-axis by a random velocity drawn from a uniform distribution of -1 to 1 km/s.

V. Results

VI. Future work

We are going to use these initial conditions in the simulations of evolution of star clusters.

References

[1] Mark R. Krumholz, Christopher F. McKee, and Joss Bland - Hawthorn. Star Clusters Across Cosmic Time. *arXiv e-prints*, page arXiv:1812.01615, Dec 2018.

[2] R. B. Larson. Turbulence and star formation in molecular clouds. , 194:809–826, March 1981.

[3] F. I. Pelupessy and S. Portegies Zwart. The evolution of embedded star clusters. , 420(2):1503–1517, Feb 2012.

[4] Michael A. Kuhn, Matthew S. Povich, Kevin L. Luhman, Konstantin V. Getman, Heather A. Busk, and Eric D. Feigelson. The Massive Young Star-Forming Complex Study in Infrared and X-Ray: Mid-infrared Observations and Catalogs. , 209(2):29, Dec 2013.

[5] Michael A. Kuhn, Eric D. Feigelson, Konstantin V. Getman, Adrian J. Baddeley, Patrick S. Broos, Alison Sills, Matthew R. Bate, Matthew S. Povich, Kevin L. Luhman, Heather A. Busk, Tim Naylor, and Robert R. King. The Spatial Structure of Young Stellar Clusters. I. Subclusters. , 787(2):107, Jun 2014.

[6] Patrick S. Broos, Konstantin V. Getman, Matthew S. Povich, Eric D. Feigelson, Leisa K. Townsley, Tim Naylor, Michael A. Kuhn, Robert R. King, and Heather A. Busk. Identifying Young Stars in Massive Star-forming Regions for the MYStIX Project. , 209(2):32, Dec 2013.

[7] Gaia Collaboration. Gaia Data Release 2. Summary of the contents and survey properties. , 616:A1, Aug 2018.

[8] P. M. Marrese, S. Marinoni, M. Fabrizio, and G. Giuffrida. Gaia Data Release 1. Cross-match with external catalogues. Algorithm and results. , 607:A105, Nov 2017.

[9] P. M. Marrese, S. Marinoni, M. Fabrizio, and G. Altavilla. Gaia Data Release 2. Cross-match with external catalogues: algorithms and results. , 621:A144, Jan 2019.

[10] F. X. Pineau, C. Motch, F. Carrera, R. Della Ceca, S. Derrière, L. Michel, A. Schwope, and M. G. Watson. Cross-correlation of the 2XMMi catalogue with Data Release 7 of the Sloan Digital Sky Survey. , 527:A126, Mar 2011.

[11] ESA, editor. *The HIPPARCOS and TYCHO catalogues. Astrometric and photometric star catalogues derived from the ESA HIPPARCOS Space Astrometry Mission*, volume 1200 of *ESA Special Publication*, 1997.

[12] M. A. Kuhn, K. V. Getman, E. D. Feigelson, B. Reipurth, S. A. Rodney, and G. P. Garmire. A Chandra Observation of the Obscured Star-forming Complex W40. , 725:2485–2506, December 2010.

[13] S. Portegies Zwart and S. McMillan. *Astrophysical Recipes; The art of AMUSE*. December 2018.
[14] F. I. Pelupessy, A. van Elteren, N. de Vries, S. L. W. McMillan, N. Drost, and S. F. Portegies Zwart. The Astrophysical Multipurpose Software Environment. , 557:A84, Sep 2013.
[15] J. J. Monaghan. Smoothed particle hydrodynamics. , 30:543–574, 1992.
[16] Alison Sills, Steven Rieder, Jennifer Scora, Jessica McCloskey, and Sarah Jaffa. Dynamical evolution of stars and gas of young embedded stellar sub-clusters. , 477(2):1903–1912, Jun 2018.
[17] M. A. Kuhn, K. V. Getman, and E. D. Feigelson. The Spatial Structure of Young Stellar Clusters. II. Total Young Stellar Populations. , 802:60, March 2015.