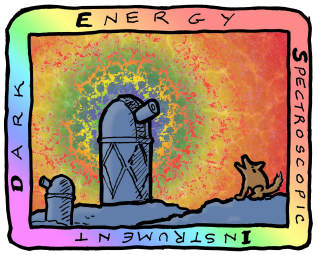
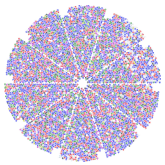
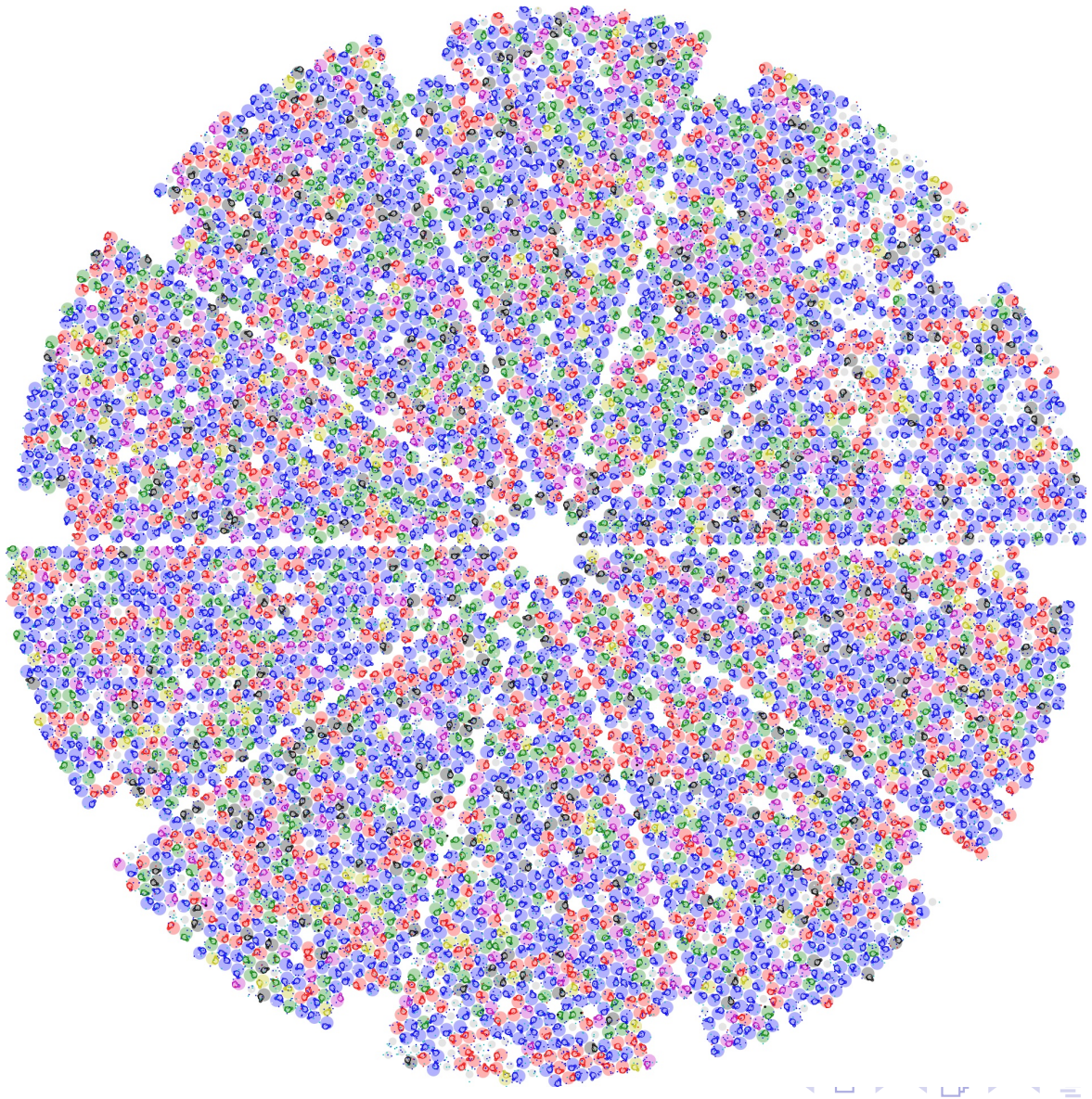
Fiber assignment for DESI

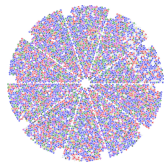
Bob Cahn and Louis Garrigue

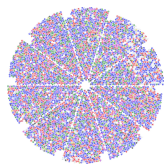
June 2015





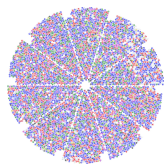






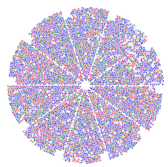
Code

- library of functions which can be easily adapted, with a "main" producing the assignment
- well "pipelined", and someone can use or modify it easily
- can be found on DESI git repository
- create the executable with a "make all" and run it on NERSC with "qsub run"



Input-Output

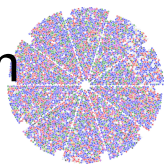
- Parameters : all parameters, including addresses of other input files
- Target DB : information, before the study, on all possible targets
- Obs DB : database constructed from the ongoing DESI observations after the data has been processed by the Spectroscopic pipeline. This DB has not been designed yet
- Survey tiles : file containing the positions of all the tiles to be observed
- Fiber positions : locations of the positioners in the focal plane



Output

FITS files

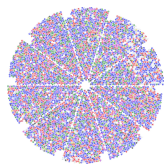
- fiber
- positioner
- number of available objects
- ID of available objects
- objtype : ELG, LRG, QSO, SKY, STDSTAR, GAL, OTHER
- targetid
- desi target : targeting info
- ra, dec
- xfocal, yfocal : mm from center in positioner coordinate system



Source files

Source files are file.h and file.cpp and are in this increasing dependency order :

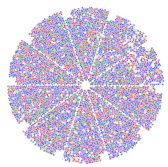
- misc : manipulate lists, tables, pairs, plots, ...
- collision : collision checking of positioners
- features : carries all parameters
- structs : objects, plates, positioners, assignment information
- global : main high-level functions and algorithms
- main : neat and quickly understandable code that sums up all the steps of the assignment



Objects

Kind	Id	Priority	Nobs	Density ($obj \cdot deg^{-2}$)
QSO Ly-α	0	1	5	50
QSO Tracer	1	1	1	120
LRG	2	3	2	300
ELG	3	4	1	2400
Fake QSO	4	1	1	90
Fake LRG	5	3	1	50
Standard Star	6	5	1	140
Sky Fiber	7	6	1	1400

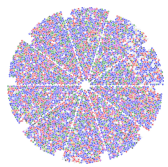
Table : Characteristics of objects samples



Parameters

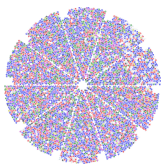
- input and output directories
- Output = true, whether you want to release the output
- Randomize = false randomize order of plates in making plans
- Pacman = false selects only spectrometers 0, 1, 2, 7, 8, 9 of the pacman
- Npass = 5 number of passes
- MaxSS = 10 ; MaxSF = 40 number of fibers assigned to SS and SF on a petal
- PlateRadius = 1.65° radius of the plate
- InterPlate = 0 minimal number of plates between two observations of the same galaxy
- Analysis = 0, tile distance for getting the information from previously observed tiles
- InfDens = false, simulate infinite density of SS and SF
- TotalArea = 15789.0 deg^2total area of the sky considered
- invFibArea = 700 inverse of area in deg^2accessible to a fiber (fiber density for a deg^2)
- moduloGal = 1, if 2 for instance, reads only one object over two in galFile
- moduloFiber = 1, same for fiber
- PatrolRad = 6.0 mm maximum distance between a positioner to a galaxy
- Collision = true when we want to allow collisions
- Exact = true whether we want exact geometry collision checking (otherwise just circles)
- AvCollide = 3.2 mm in case of no exact collision checking
- Collide = 1.98 mm, NoCollide = 7 mm, used to optimize collision checking
- NeighborRad = 14.0 mm maximum distance to consider that two fibers are neighbors
- PlotObsTime, etc, whether we want to plot some information into output files

- Verif = false, whether we verificate that the assignment is sane



Rules

- a tile-fiber is only assigned once
- no collision between fiber positioners
- two observations of a QSO or an LRG can't be separated by less than `InterPlate` plates
- in last pass, only ELG, SS and SF are considered
- there are 10 and 40 fibers assigned to SS and SF per petal
- try to assign a fiber to the highest prioritized object, then to the least observed
- between an already observed Ly-α and an unknown QSO, one chooses the already observed Ly-α
- the same SS or SF can be observed several times
- SS has priority over SF
- our policy is : while improving the assignment, never unassign an object without reassigning it right then



Choose by density

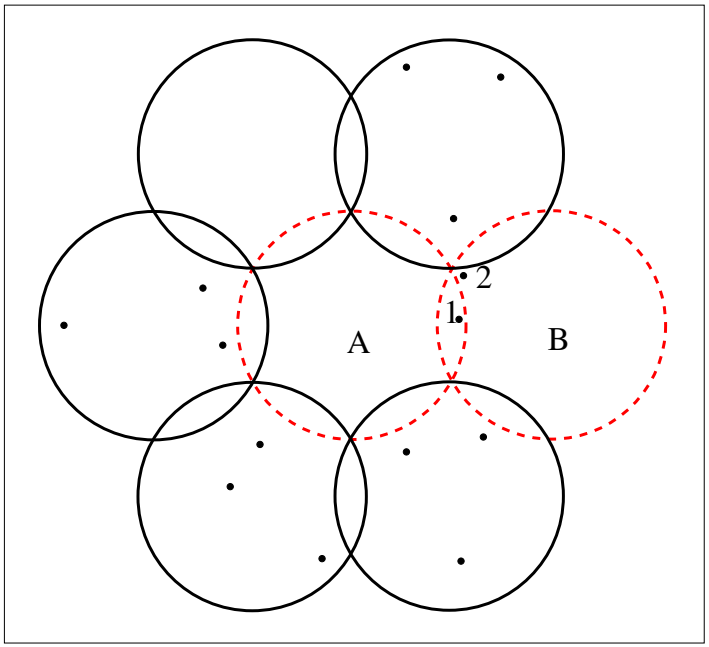
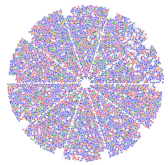


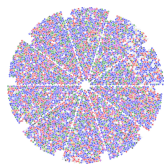
Figure : Choice by density : A is chosen to observe 1



Doesn't improve the results

Assignment plan

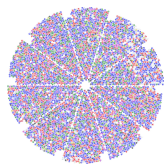
- make a plan before begining the study, without SS nor SF
- begining the real 5-years observation, assigning SS and SF just before observing a tile
- update this plan during observations, from information from analysis of exact kinds of objects



First assignment plan

For all plates (tiles) :

- assign QSO
- assign LRG
- assign ELG

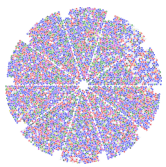


Improvement functions ideas

Figure : Improve

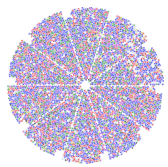
Figure : Redistribute

Figure : Improve from kind

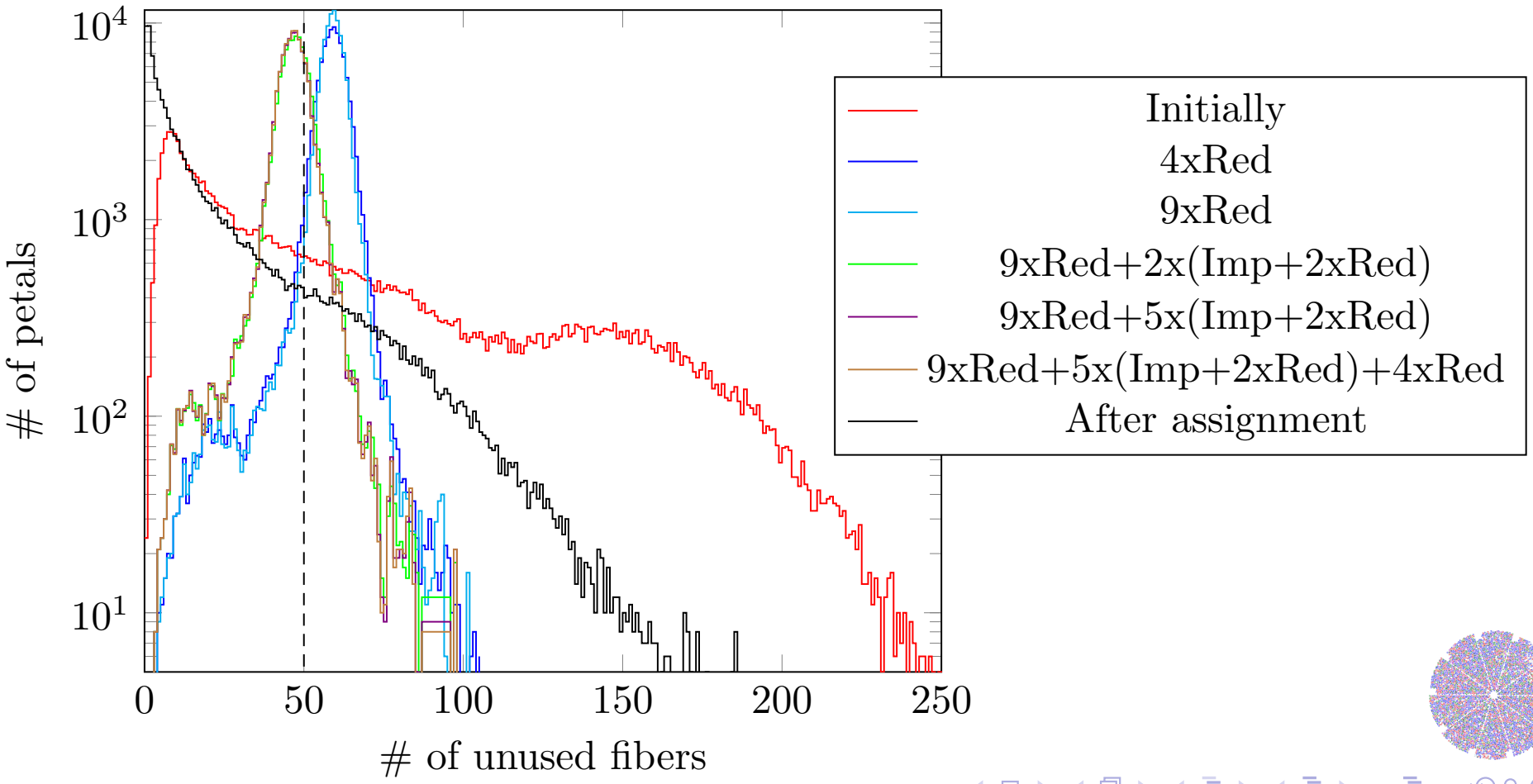


Improving the plan

- Run, globally, on the list plates from 0 to last :
- Simply assign fibers
- Redistribute (several times)
- Improve + 2x Redistribute (several times)
- Redistribute (several times)



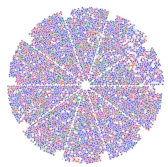
Improving the plan



Applying the plan

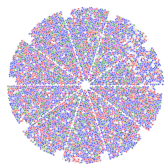
Tile after tile, as we observe :

- Assign SS and SF, replacing some ELG
- Assign further observed objects to unused fibers if possible
- Observe the tile
- Update the plan left with information from analysis of previously observed galaxies
- Improve the plan left (several redistribute, improve, redistribute)



Execution

```
# Read 71,998,144 galaxies from objects_ss_sf0.rdzipn
# Read 10,666 plate centers from desi-tiles.par and 5000 fibers from fiberpos.txt
# Start building HTM tree at 13.8 s ... took : 25.5 s
# Begin collecting available galaxies ... took : 31.6 s
# Begin computing available tilefibers ... took : 1 mn 12.5 s
# Start assignment at : 2 mn 27 s
# Begin new assignment :
50,518,743 assignments on all left next plates
# ... took : 18 mn 38.3 s
# Begin improve :
565,801 more assignments (1.120 % improvement)
# ... took : 36.5 s
# Begin redistribute TF :
1,760,465 redistributions of couples of TF
# ... took : 23.3 s
# Begin real time assignment at 23 mn 24.5 s
- Plate 0 : 550 not as - 3852 unas & 2772 replaced
- Plate 1 : 98 not as - 3784 unas & 2773 replaced
# Begin redistribute TF :
1,160,465 redistributions of couples of TF
# ... took : 23.3 s
# Begin improve :
106,948 more assignments (0.205 % improvement)
# ... took : 30.2 s
- Plate 2 : 48 not as - 3128 unas & 2475 replaced
- Plate 3 : 96 not as - 3407 unas & 2485 replaced
```

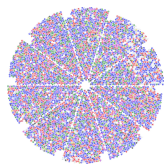


Results

	Times observed									
	0	1	2	3	4	5	Total	Fibused	obs %	weighted %
QSOLy-a	0	1	5	12	19	10	49	180	99.151	72.206
QSOTracer	1	118	0	0	0	0	119	118	99.141	99.141
LRG	13	42	243	0	0	0	298	528	95.505	88.434
ELG	480	1,930	0	0	0	0	2,411	1,930	80.054	80.054
FakeQSO	0	89	0	0	0	0	90	89	99.139	99.139
FakeLRG	2	47	0	0	0	0	50	47	95.792	95.792

Table : Densities (objects/deg^2) as as a function of # of observations (with total), and % observed, once and weighted

51,044,452 assignments in total (95.7143 % of all fibers)



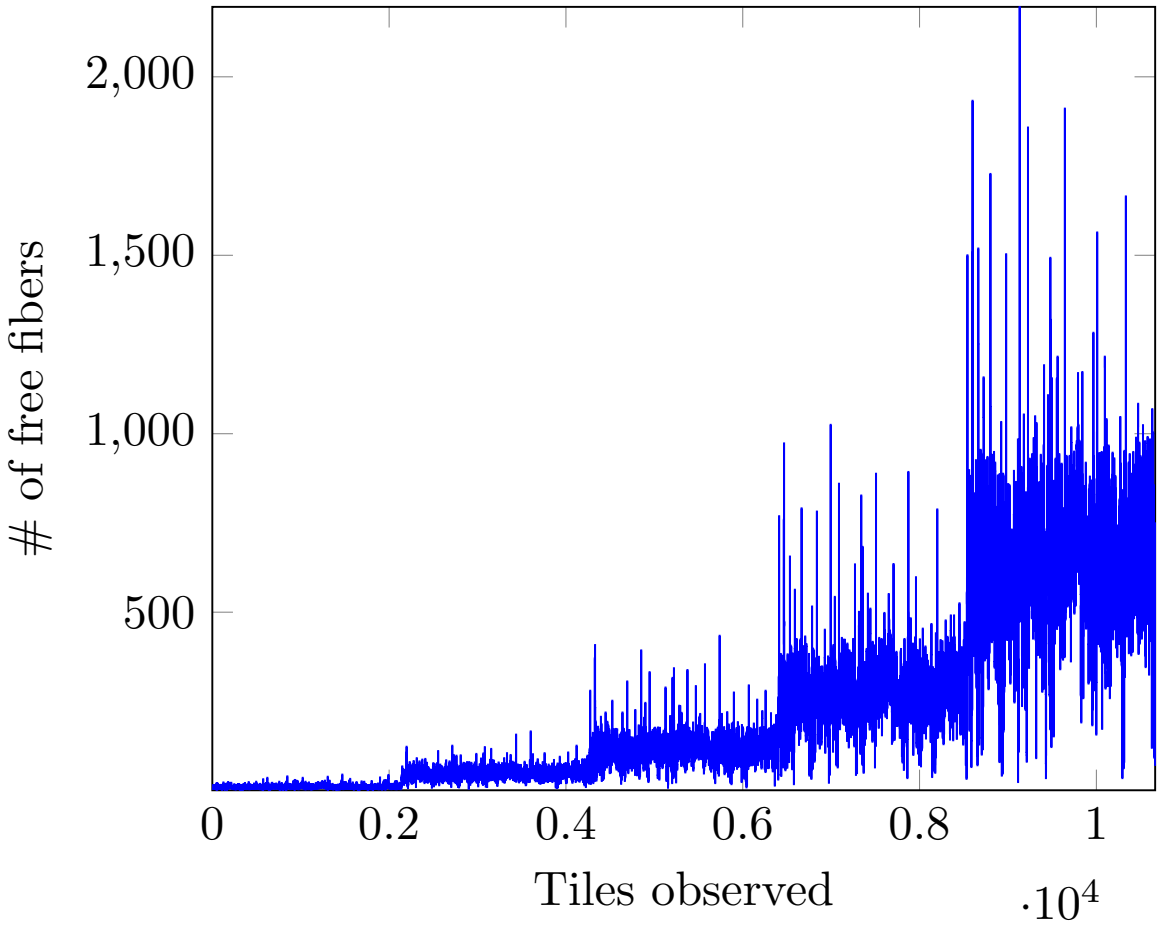
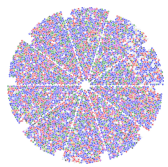


Figure : # free fibers as a function of time (plates)



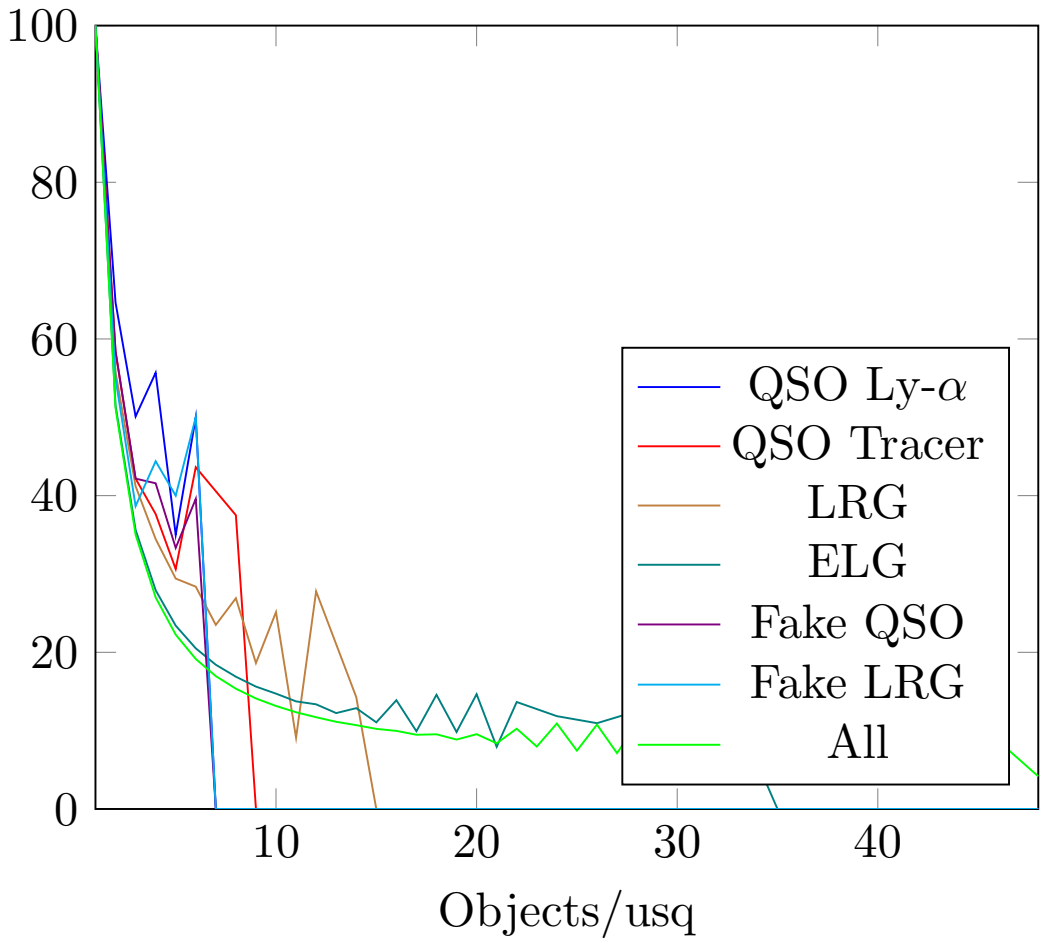
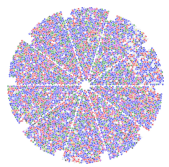


Figure : % of observed galaxies as a function of objects density



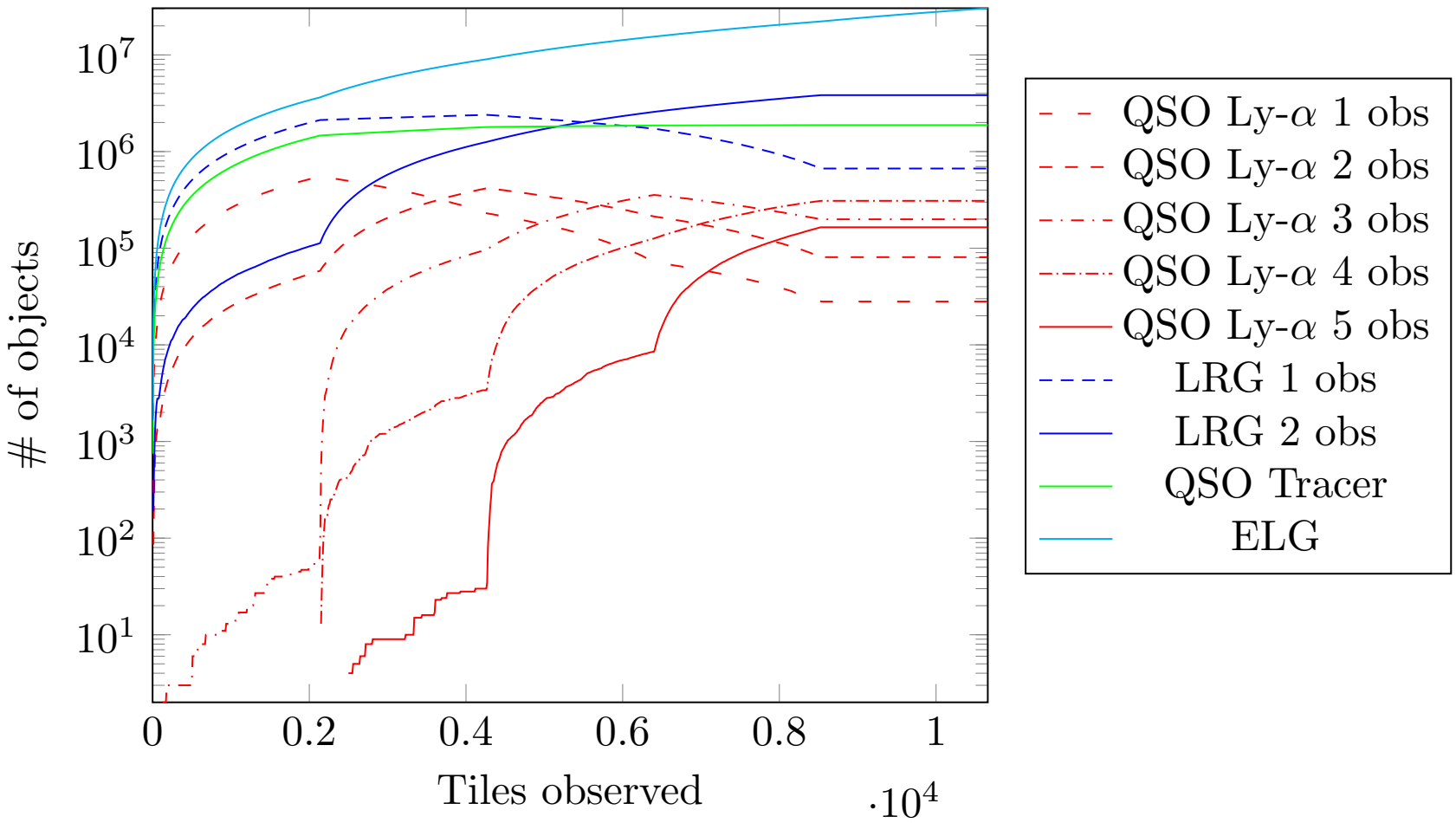
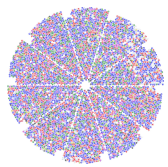
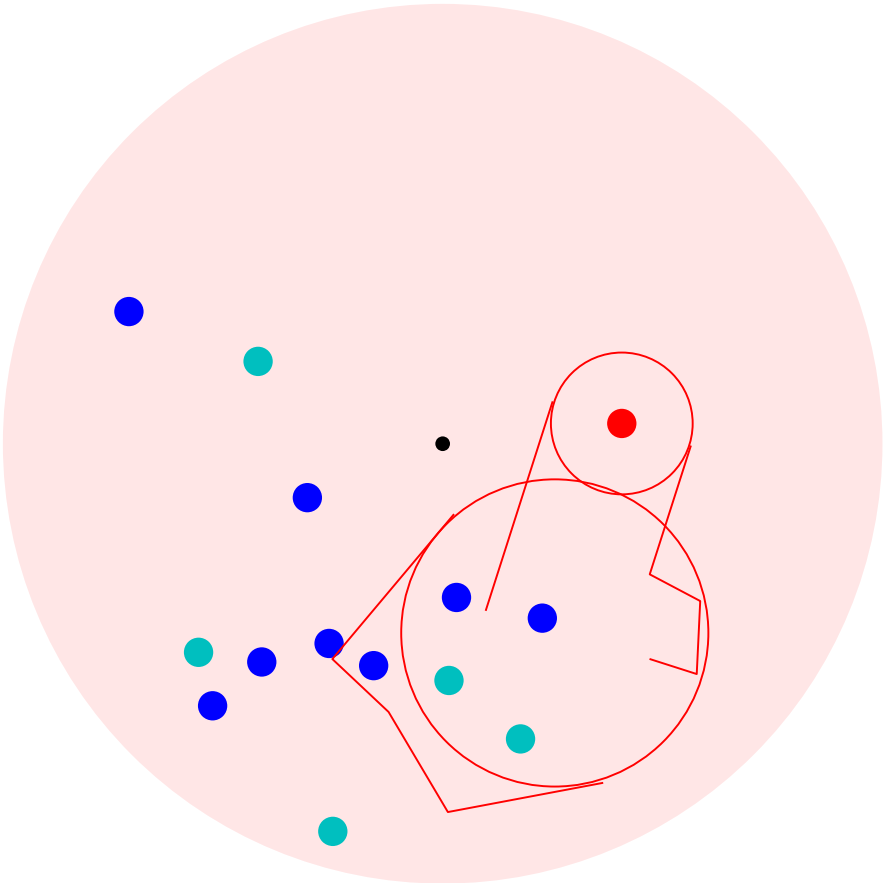


Figure : Observed galaxy kind as a function of time (plates seen)



Geometry



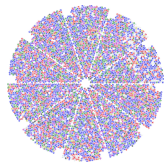


Figure : Geometry of a positioner, with fiber holder and central body

Collision rate

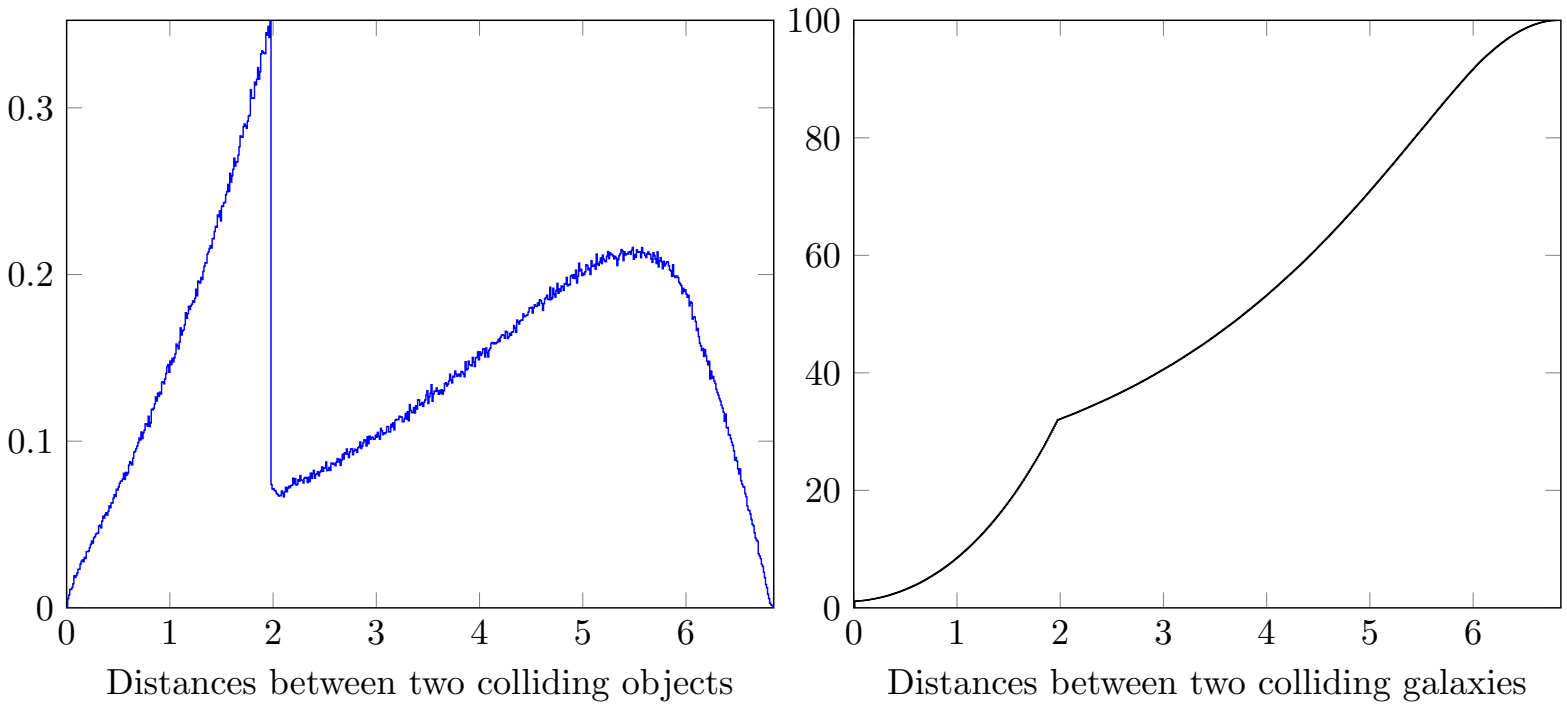
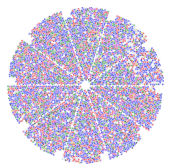


Figure : Histogram of distances between galaxies in collision cases, and integral



Collision rate : 10.8%

Mathematical optimum

Have we reached the optimum ? Mathematical expressions will give the answer

Imagine a portion S of the sky of the size the reacheable area by a single fiber

Let i be the indice for different kinds (Ly-α , fake QSO, ... except SS or SF)

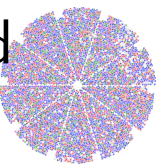
Let λ_i be the densities of the kinds of objects in S

Let N be the indice for the portions of sky with exactly N objects (when we take an area of size S)

Let f_k be the fraction of the sky which density corresponds to N

Let $poisson(x, n) = \frac{x^n e^{-x}}{n!}$ be the Poisson distribution

Then, for a given ELG in S, named e, the probability to be observed is :



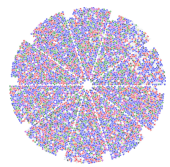
$p(\text{e}) = \sum_{1 \leq N \leq 10} f_k \sum_{0 \leq A \leq N-1 \text{and} 1 \leq B \leq \infty} p(\text{A objects having priority over an ELG}) \cdot$
$p(\text{B ELG and e is chosen among the N-A firsts})$
$= \sum_{1 \leq N \leq 10} f_k \sum_{0 \leq A \leq N-1 \text{and} 1 \leq B \leq \infty} p(A) \cdot p(\text{B ELG}) \cdot p(\text{e is chosen among the N-A firsts})$
$= \sum_{1 \leq N \leq 10} f_k \sum_{0 \leq A \leq N-1} \sum_{1 \leq B \leq \infty} p(A) \cdot \frac{\lambda_{ELG}^B e^{-\lambda_{ELG}}}{B!} \cdot \frac{N-A}{B}$
$= I \cdot \sum_{1 \leq N \leq 10} f_k \sum_{0 \leq A \leq N-1} p(A)(N-A)(1)$

where $I = e^{-\lambda_{ELG}} \int_0^{\lambda_{ELG}} \frac{e^{\lambda}-1}{\lambda} d\lambda$

Also we define the set L of objects with priority against ELG, which are Ly-α , fake QSO, QSO target, LRG and fake LRG. Then, for $i \in L$:

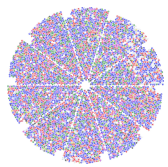
$p(A) = \sum_{n_i \, st \, \sum_i goal(i) n_i = A} \prod_i poisson(\lambda_i, n_i)$

Same method to compute theoretical percentage of observed QSO and LRG



On the code

- A lot of graphics displayed
- Hundreds of functions in this library, you can try yourself an other algorithm, or display other results easily
- Easy to adapt for other pipelines



To do

- Compute mathematical optimum
- Add SF and SS to the mock catalog (+0.5% of observed ELG when infinite density)
- Adapt the C FITS output function to C++
- There are still things to adapt to this code, for example the fact that the weather can be bad
- It can be used to optimize the tilling strategy
- Lado Samushia is computing the bias introduced by the fiber assignment on the correlation function
- Other ideas for improving ?