HFD filter system design for Gaia

New proposals and general insights concerning MBP

Coryn A.L. Bailer-Jones

Max-Planck-Institut für Astronomie, Heidelberg

PWG Meeting, Copenhagen, 28-29 June 2004

see GAIA-CBJ-016 for details

Considerations

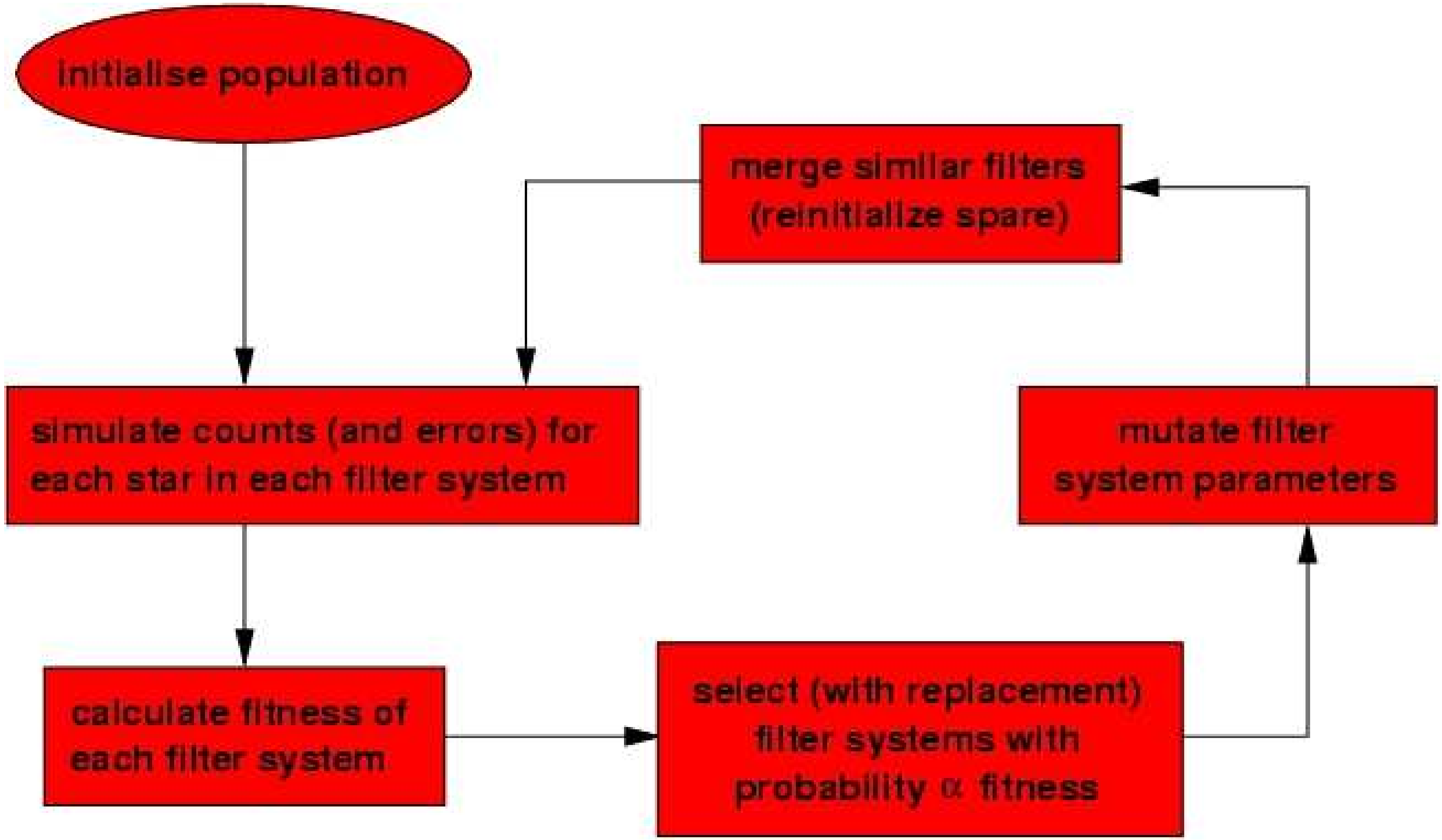
Goal: *design a filter system to determine multiple APs across a wide parameter space subject to instrumental constraints*

- conflicting demands on filter system
- manual design is complex; gives no idea of optimality
- cast as a mathematical optimization problem:
  - parametrize filter system
  - establish a figure-of-merit of filter system performance
  - maximise this as a function of the filter system parameters

Heuristic filter design (HFD) model

- figure-of-merit / fitness function:
  - measure of ability of filter system to maximally separate stars with different APs
  - stellar grid shows variance in APs of interest (T_{eff}, logg, [Fe/H], A_V)
  - use an instrument model to simulate photon counts and errors in a filter for each star in grid
- Evolutionary Algorithms: population-based optimization
  - 1 individual = 1 candidate filter system
  - genetic operators: search, selection
  - provides a stochastic (but not random) search
  - evolve population and find fittest filter system

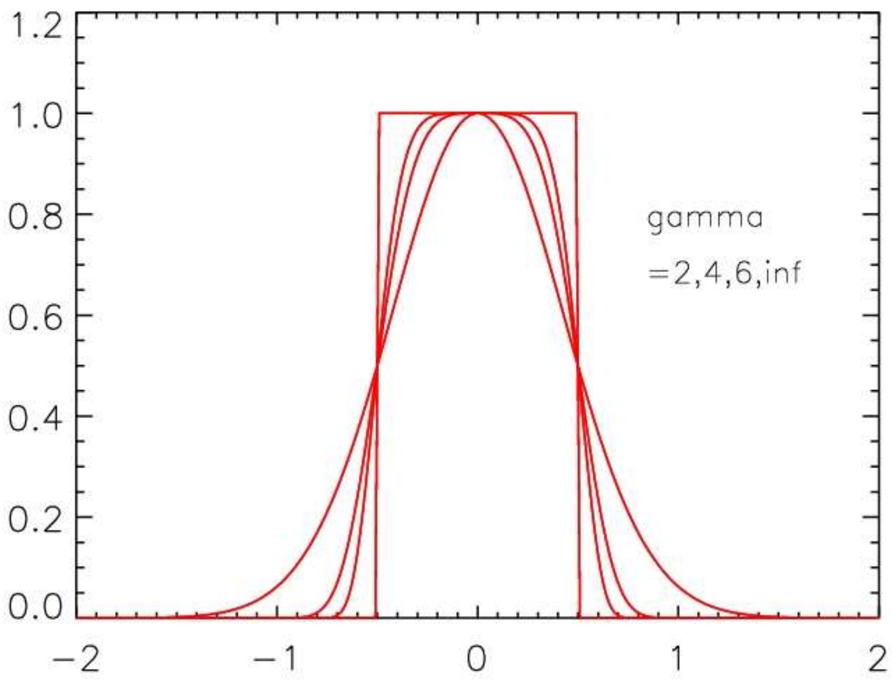
HFD model


initialise population
merge similar filters
(reinitialize spare)
simulate counts (and errors) for
each star in each filter system
mutate filter
system parameters
calculate fitness of
each filter system
select (with replacement)
filter systems with
probability α fitness

Filter system representation

Each filter system consists of *I* filters each with 3 parameters:

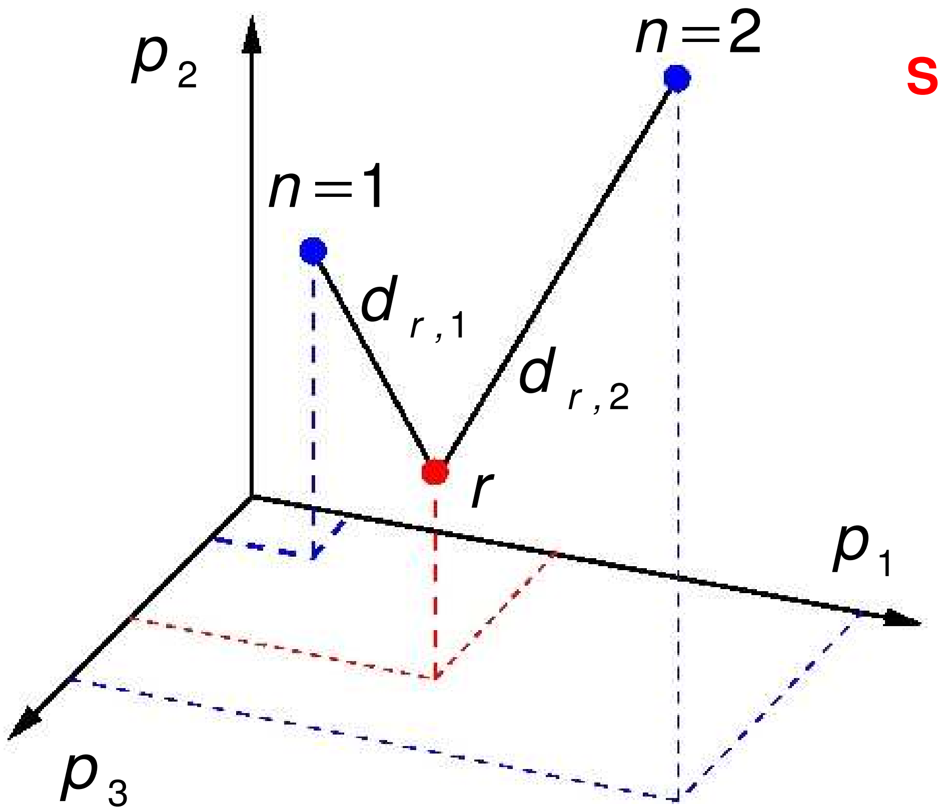
c central wavelength
w half width at half maximum
t fractional integration time (of total available for all filters)



Generalised Gaussian profile with $\gamma = 8$

$$y = \exp(-\ln 2\ [(\lambda - c)/w]^{\gamma})$$

Fitness: SNR distance & AP gradient



SNR distance of star r from neighbour n:

$$d^2_{r,n} = \sum_i \frac{(p_{i,n} - p_{i,r})^2}{\sigma^2_{i,n} + \sigma^2_{i,r}}$$

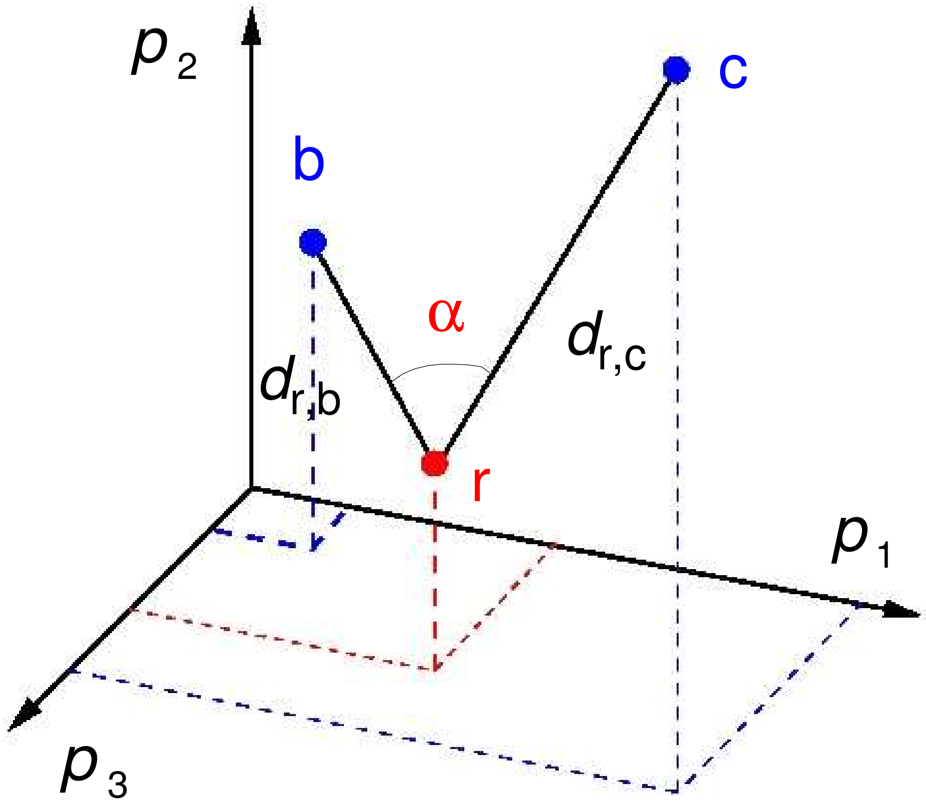
$p_{i,n}$ = photon counts in filter *i* for star *n*

$\sigma_{i,n}$ = expectation of error in $p_{i,n}$

AP gradient: $$\frac{d_{r,n}}{\Delta\phi_{j,n,r}}$$

$\Delta\phi_{j,n,r}$ = difference in AP j between star n and r

Fitness: vector separation



Find nearest neighbours to r, each of which differs in only 1 AP

For each pair of neigbours, calculate angle, α, between their vectors:
Nearer to 90° => better separation
(less degeneracy)

Fitness at r for this AP pair is

$$f_{r,b,c} = \sin\alpha \frac{d_{r,b}\, d_{r,c}}{\Delta\phi_{r,b}\, \Delta\phi_{r,c}}$$

Total fitness = $\sum_r \sum_{b,c} f_{r,b,c}$

J APs => J NNs per source
=> J(J-1)/2 vector pairs

Genetic operators

Selection

Select from parent population with probability proportional to fitness *(+ Elitism)*

Mutation *(modified)*

$c_i\,(g+1) = c_i\,(g) + N(0,\sigma_c)$

$h_i\,(g+1) = h_i\,(g)\,\exp\,[N(0,\sigma_h)]$

$t_i\,(g+1) = t_i\,(g)\,\exp\,[N(0,\sigma_t)]$

Merging *(new)*

Merge similar filters (reinitialize spare)

Stellar grid

logg			T_{eff} / K (SpT)				
4.5	3500 MV	4750 KV	5750 GV				
4.0				6750 FV	8500 AV	15000 BV	35000 OV
3.5							
3.0				6000 RRLyr	8500 BHB		
2.5			5500 GIII			15000 BIa	
2.0		4500 KIII	5500 FIa				
1.5					8500 AI		
1.0	3500 MIII		5000 GI				
0.5							
0.0	3500 MIa						
	[Fe/H]:		+0.5 = 0.0 if $T_{eff} > 10000$ K	0.0	−0.5	−1.5	−2.5
	$\mathbf{A_V}$:		0.0	0.3	0.6	0.9	1.2

17 T_{eff} / logg combinations at each of 5 [Fe/H] and extinction values
=> 360 sources

BaSeL 2.2 library
+ Fitzpatrick (1999) extinction curves

noise-free data

!! WARNING !!

Ångstroms in use!

Conflict with Gaia conventions
Do not repeat this at home

HFD setup

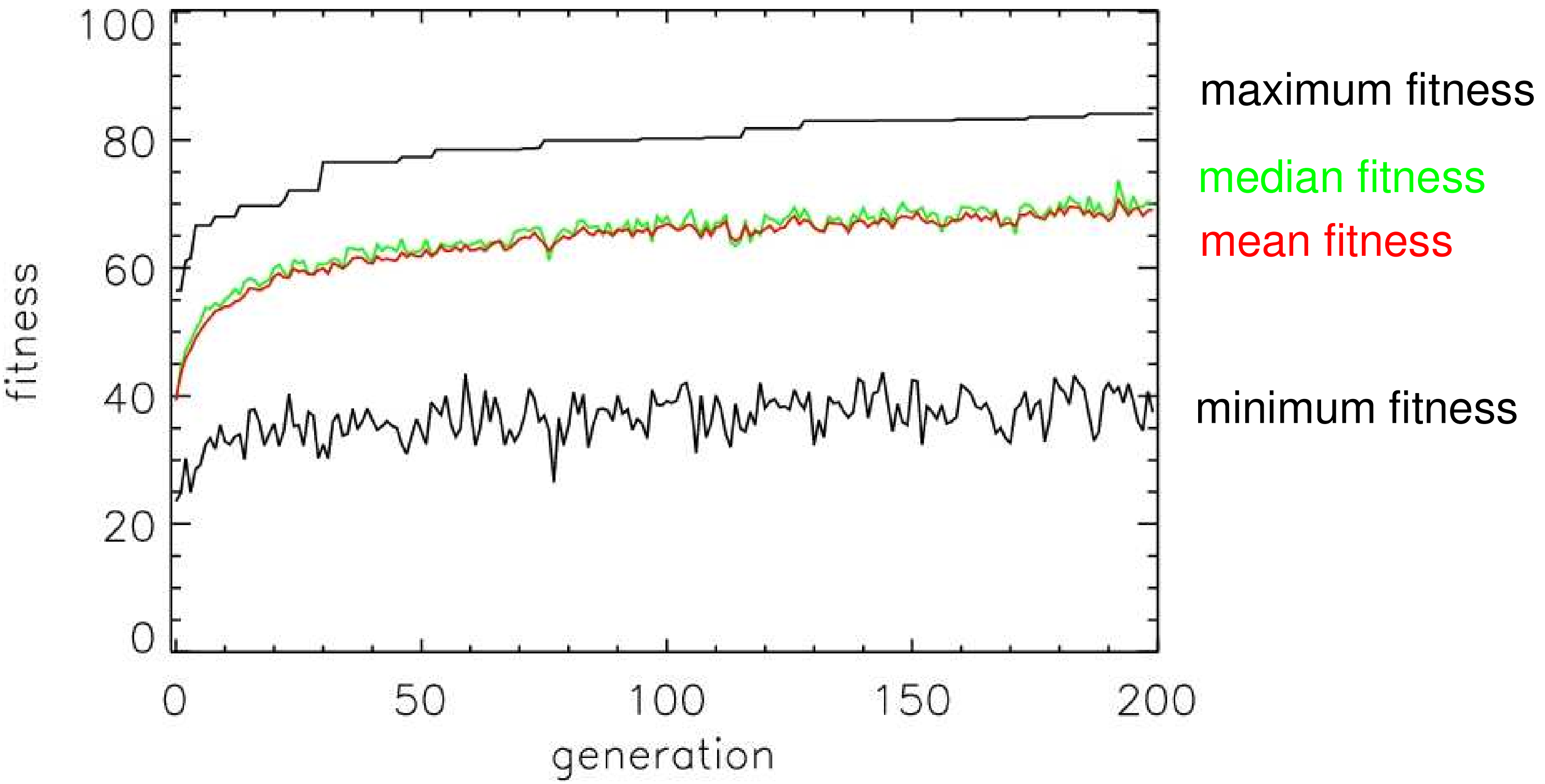
		BBP	MBP
Free parameters			
central wavelength / Å	c	$\times I$	
full-width at half maximum / Å	b	$\times I$	
fractional integration time	t	$\times I$	
Fixed parameters: fitness measure		**BBP**	**MBP**
stellar population		as in BJ04	grid3d
magnitude of stars (in G band)		20	15
AP weight A_V		1.5/128	1.5/128
AP weight [Fe/H]		75.0/128	75.0/128
AP weight $\log g$		50.0/128	50.0/128
AP weight $\log T_{eff}$		1.5/128	1.5/128
Fixed parameters: evolutionary algorithm			
number of filter systems (= population size)	K	200	200
size of elite	E	50	50
number of generations		200	200
number of runs		10	10
probability of recombination		0.0	0.0
probability of mutation for c		0.4	0.4
probability of mutation for b		0.4	0.4
probability of mutation for t		0.0	2/15
std. dev. of mutation for c / Å	σ_c	200	400
std. dev. of mutation for b	σ_b	0.2	0.2
std. dev. of mutation for t	σ_t	0.2	0.2
filter merging threshold	M_{limit}	n/a	0.7
Fixed parameters: instrumental			
filter profile		as in BJ04	as in BJ04
number of filters	I	4	12
telescope aperture area / m^2		0.7	0.25
total integration time / s		1205	16560
CCD & instrument response		as in BJ04	as in BJ04
CCD readout noise / e^-		251	167
effective background / G mag		22.37	18.29
min.$(c-fb)$, max.$(c+fb)$ / Å ($f = 1.162$)		3900, 10100	2500, 10400
min. b, max. b / Å		300, 1500	80, 250
min. t, max. t	t_{min}, t_{max}	n/a	0.03, 0.2

MBP:
- I=8,10,12,14,16
- fixed RVF
- @ G=15

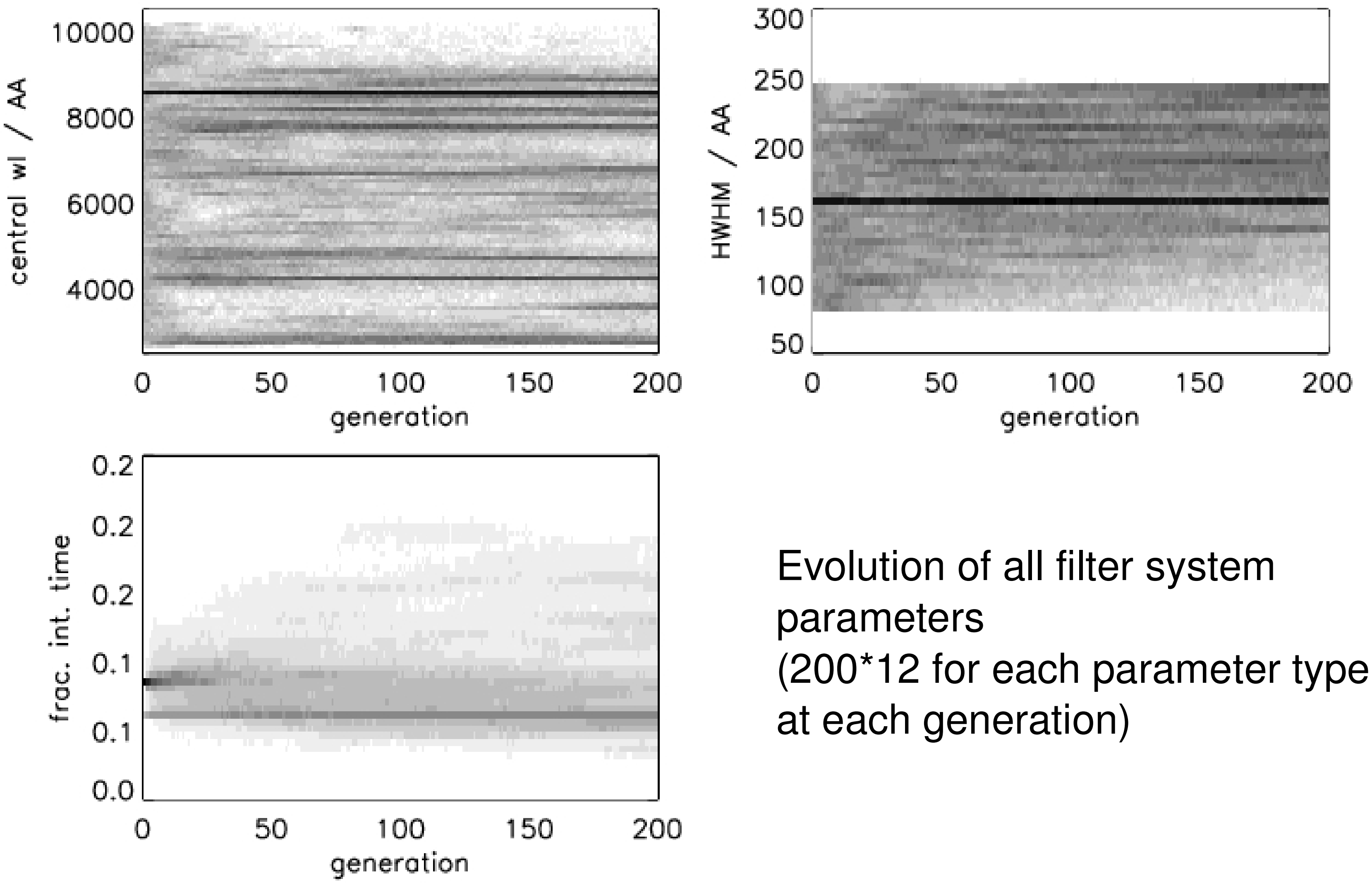
MBP variations tested:
- no merging
- broader filters
- only w.r.t orthovariance
- ...

Selection criteria:
- individual fitness terms (10)
- `appearance'
- performance at G>15

12-filter MBP: fitness evolution



12-filter MBP: filter system evolution

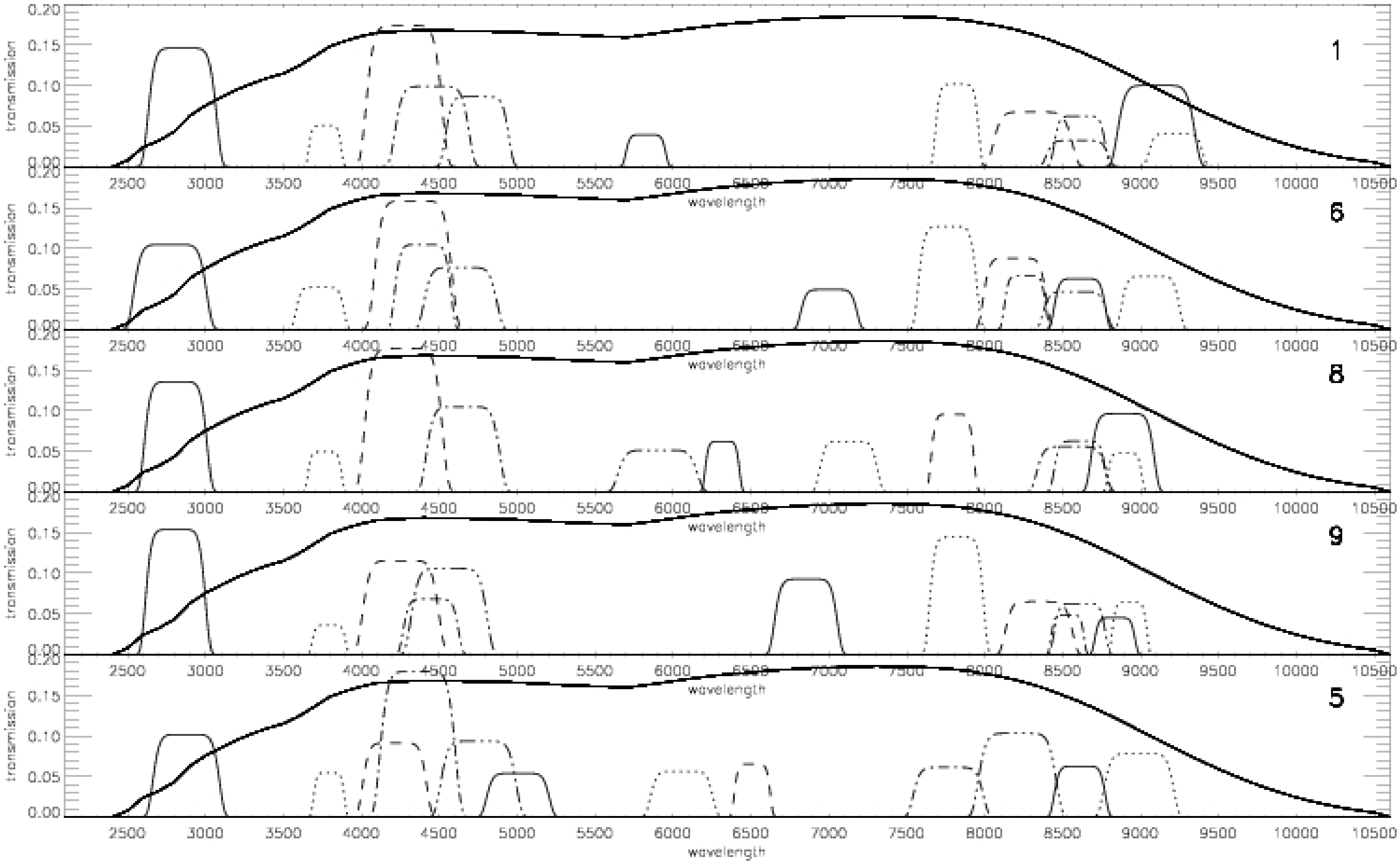


Evolution of all filter system parameters
(200*12 for each parameter type at each generation)

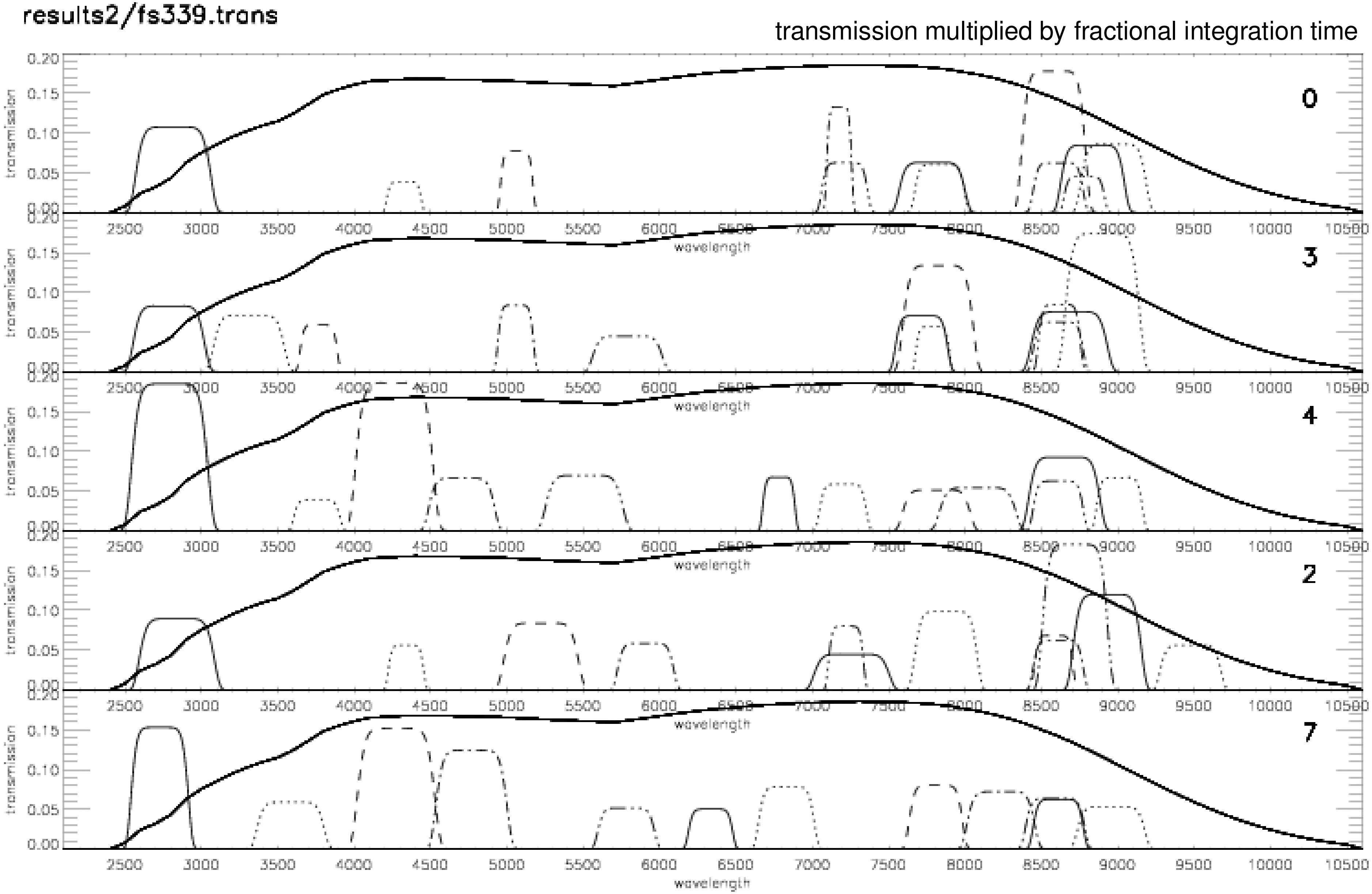
12-filter MBP: optimized systems

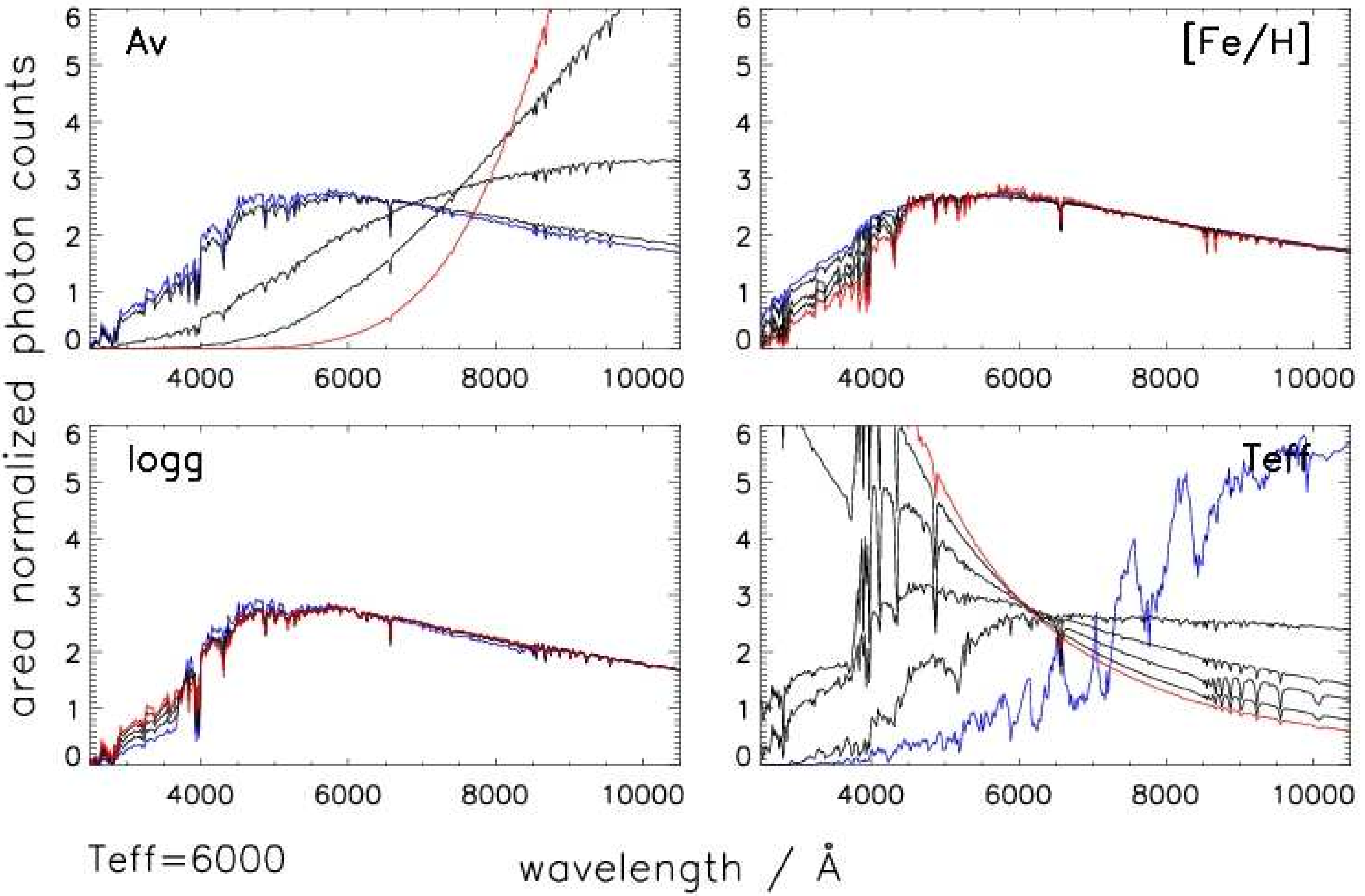
results2/fs339.trans

transmission multiplied by fractional integration time

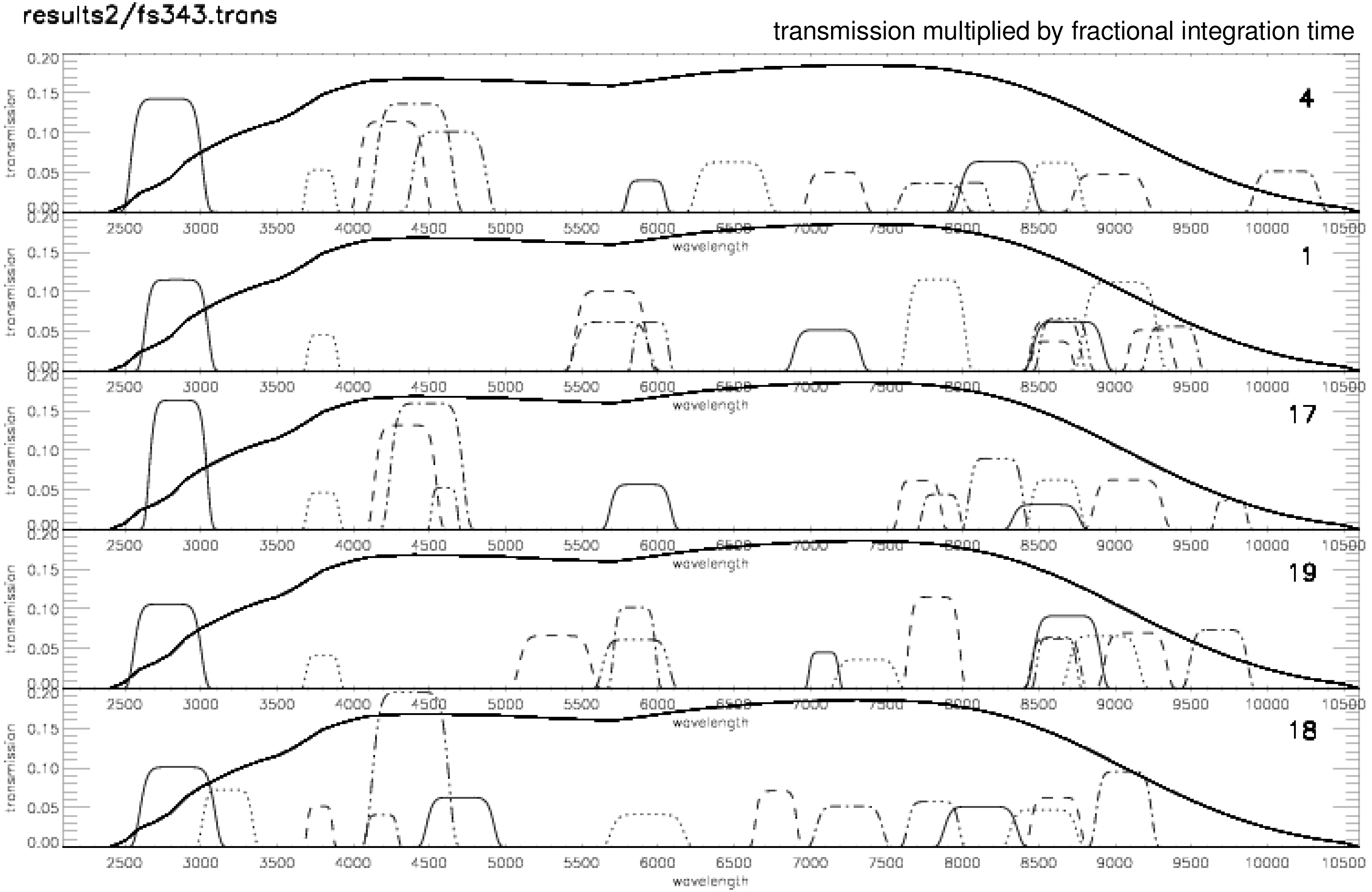


12-filter MBP: optimized systems

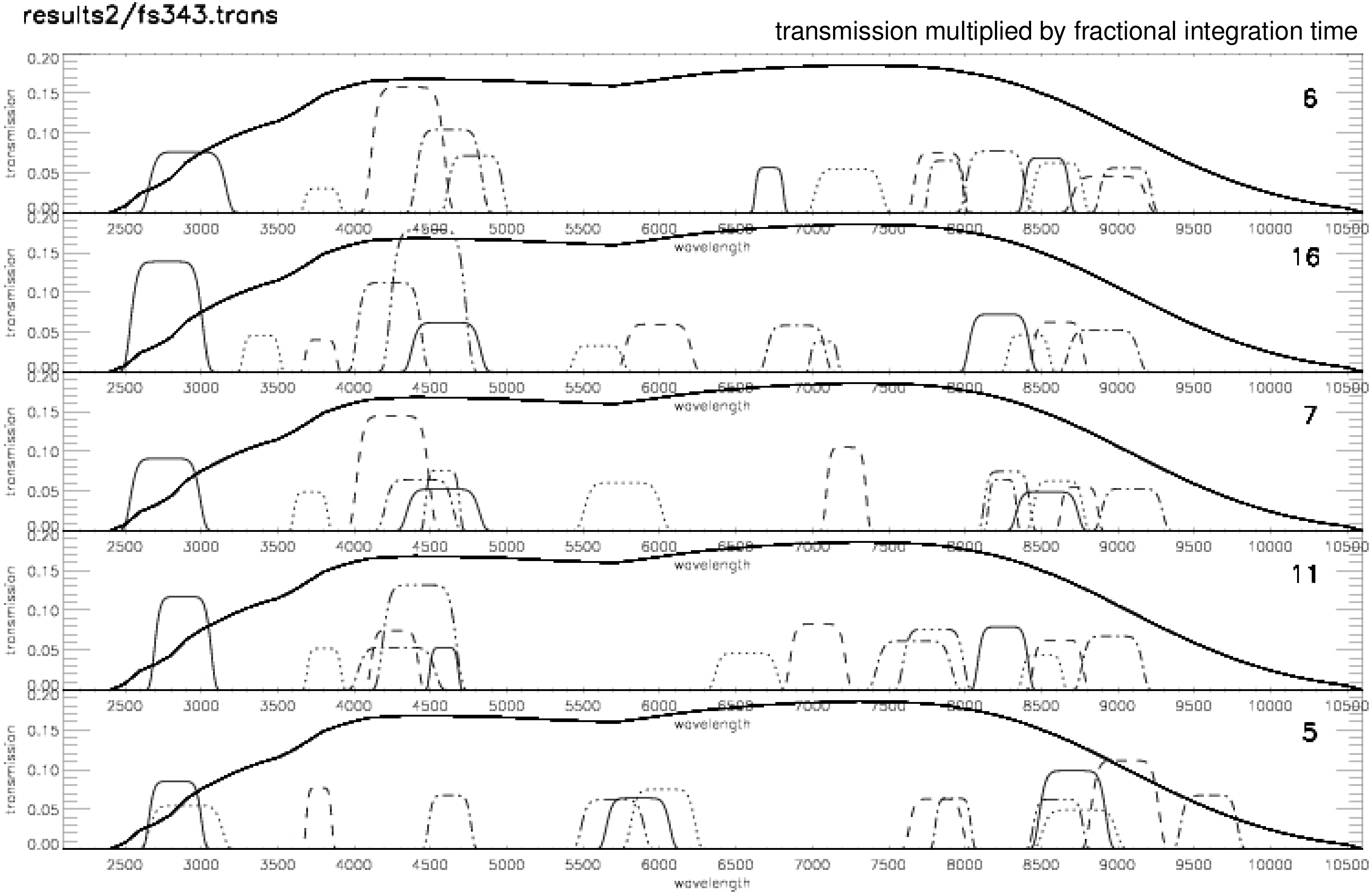



Av
[Fe/H]
logg
Teff
Teff=6000
area normalized photon counts
wavelength / Å
0
1
2
3
4
5
6
4000
6000
8000
10000

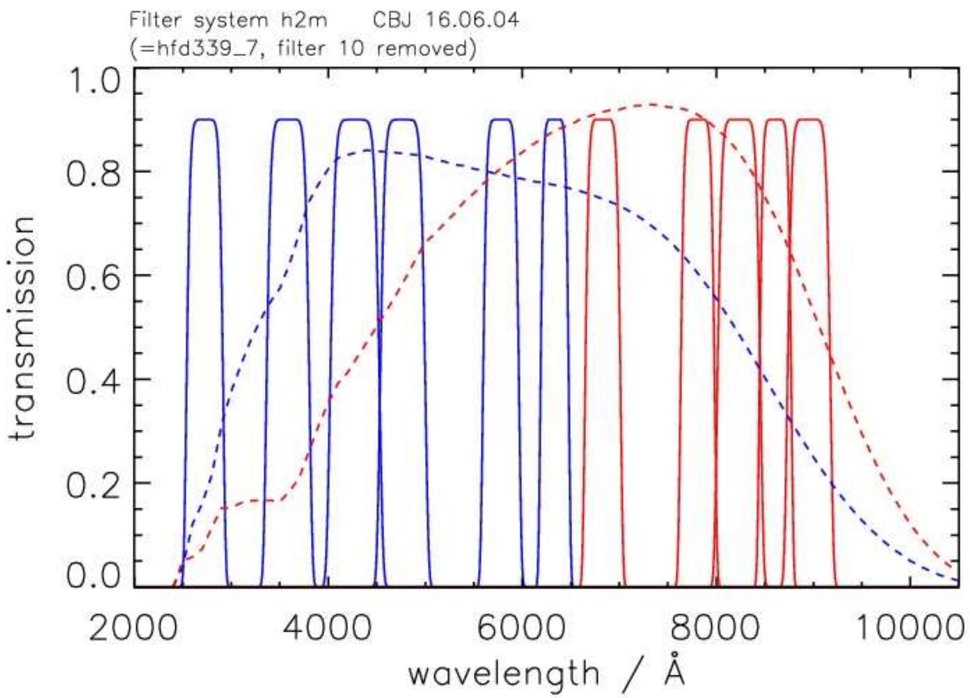
14-filter MBP: optimized systems



14-filter MBP: optimized systems



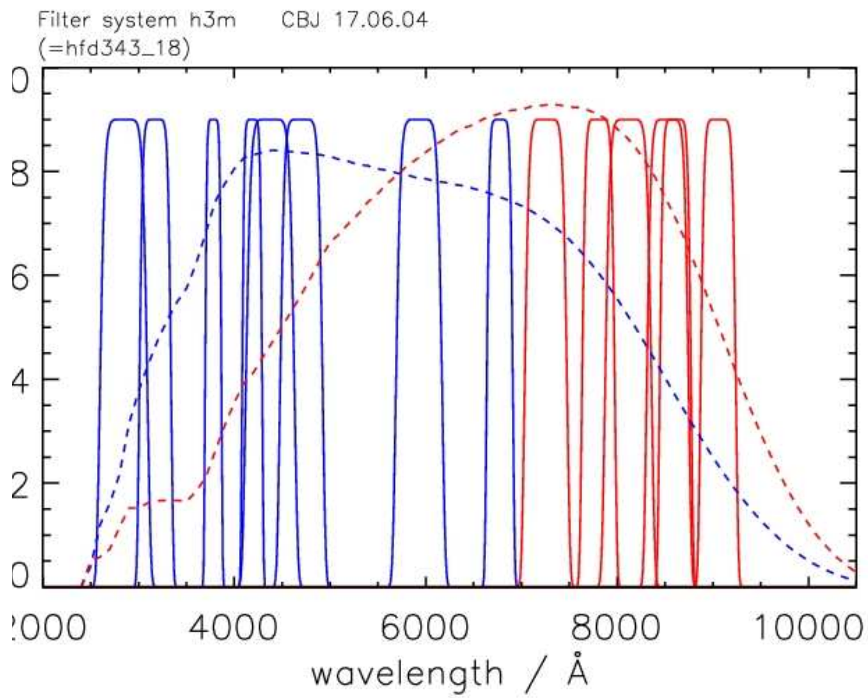
H2M: 11 filters



Allocation: 3 1 2 2 1 1 ; 1 1 1 2 1

- 12 filter optimization
- 1 duplicate filter removed: itime given to RVF
- itime discretized

H3M: 14 filters



Allocation: 2 1 1 2 1 1 1 1 ; 1 1 1 1 1 1

- 14 filter optimization
- itime discretized

System performance

- fitness, AP-gradients and orthovariances calculated on discrete systems
- checked at G>15
- different discretizations checked
- verification with standard parametrization techniques (MDM, ANN etc.) not yet done
- addition of RVS, MBP+BBP, parallax not done

	H2M	H3M	H2B	H3B
Fitness	81.1	72.1	0.45	0.49
$h(A_V)$	547	530	64	72
h([Fe/H])	295	286	2.5	2.8
$h(\log g)$	277	311	3.0	3.7
$h(T_{eff})$	5628	5421	61	73
$\sin\alpha(A_V,[Fe/H])$	0.86	0.86	0.65	0.58
$\sin\alpha(A_V,\log g)$	0.83	0.83	0.61	0.57
$\sin\alpha(A_V,T_{eff})$	0.30	0.30	0.24	0.23
$\sin\alpha([Fe/H],\log g)$	0.78	0.83	0.65	0.59
$\sin\alpha([Fe/H],T_{eff})$	0.83	0.83	0.61	0.58
$\sin\alpha(\log g,T_{eff})$	0.76	0.75	0.56	0.57

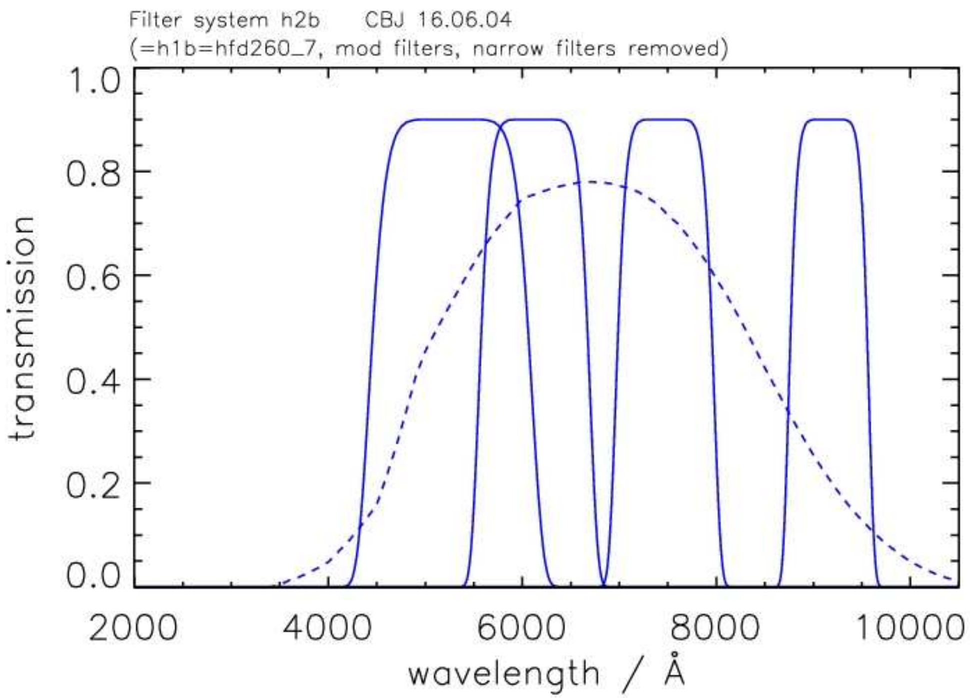
Generic MBP results (1)

- fittest filter systems are generally those with broad, overlapping filters (less constrained optimization)
- trade-off in fitness concerning filter width
  - broad: increases AP-gradients (scalar separation)
  - narrow: increases orthovariance (vector separation)
- many different filter systems with similar performance
  - small across-run variance in optima
  - numerous local optima of similar fitness
- some filters frequently reproduced in different runs (e.g. ~7 of 8) ...
- ... but overall convergence degrades with more filters

Generic MBP results (2)

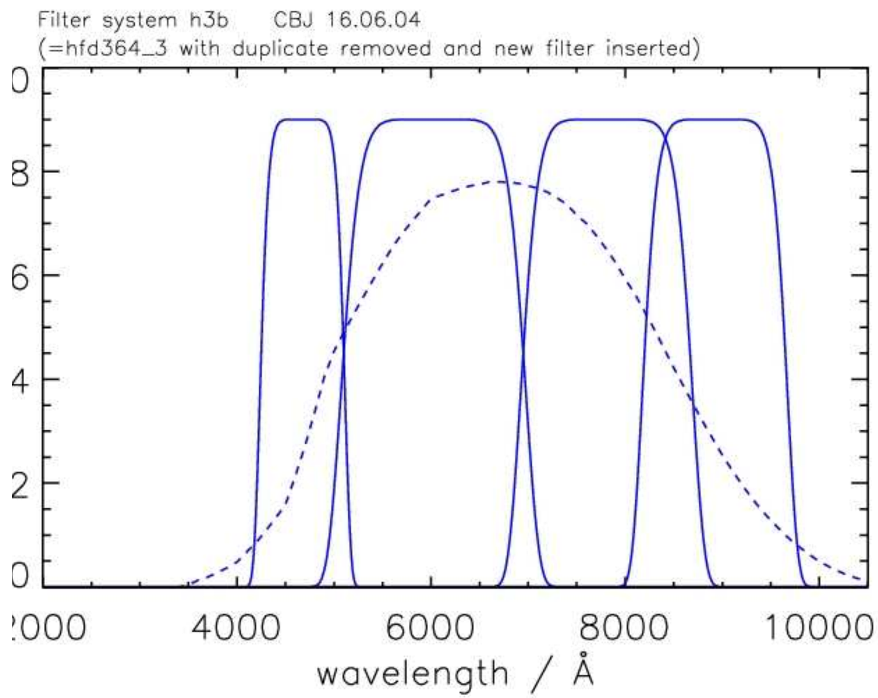
- universal preference for a filter as blue as possible
- universal preference for one or more filters extending to 9000Å or 9500Å, but not beyond
- a tendency for overlapping filters, esp. around 7500-9000Å
- generally a larger number of red (c > 5500Å) filters than blue
- not a uniform distribution of filters
  - tendency for 'crowding' in red and blue
  - tendency for a gap around 6000-8000Å (normalization?)
- flexibility with exact slot allocation

H2B



- = H1B with narrow filter removed
- H1B was a full optimization

H3B



- 2 extreme filters fixed
- itime fixed (1/4)
- optimize w.r.t sin α (A_V, T_{eff}) only
- remove duplicate and fill gap

Conclusions concerning HFD

Negative points

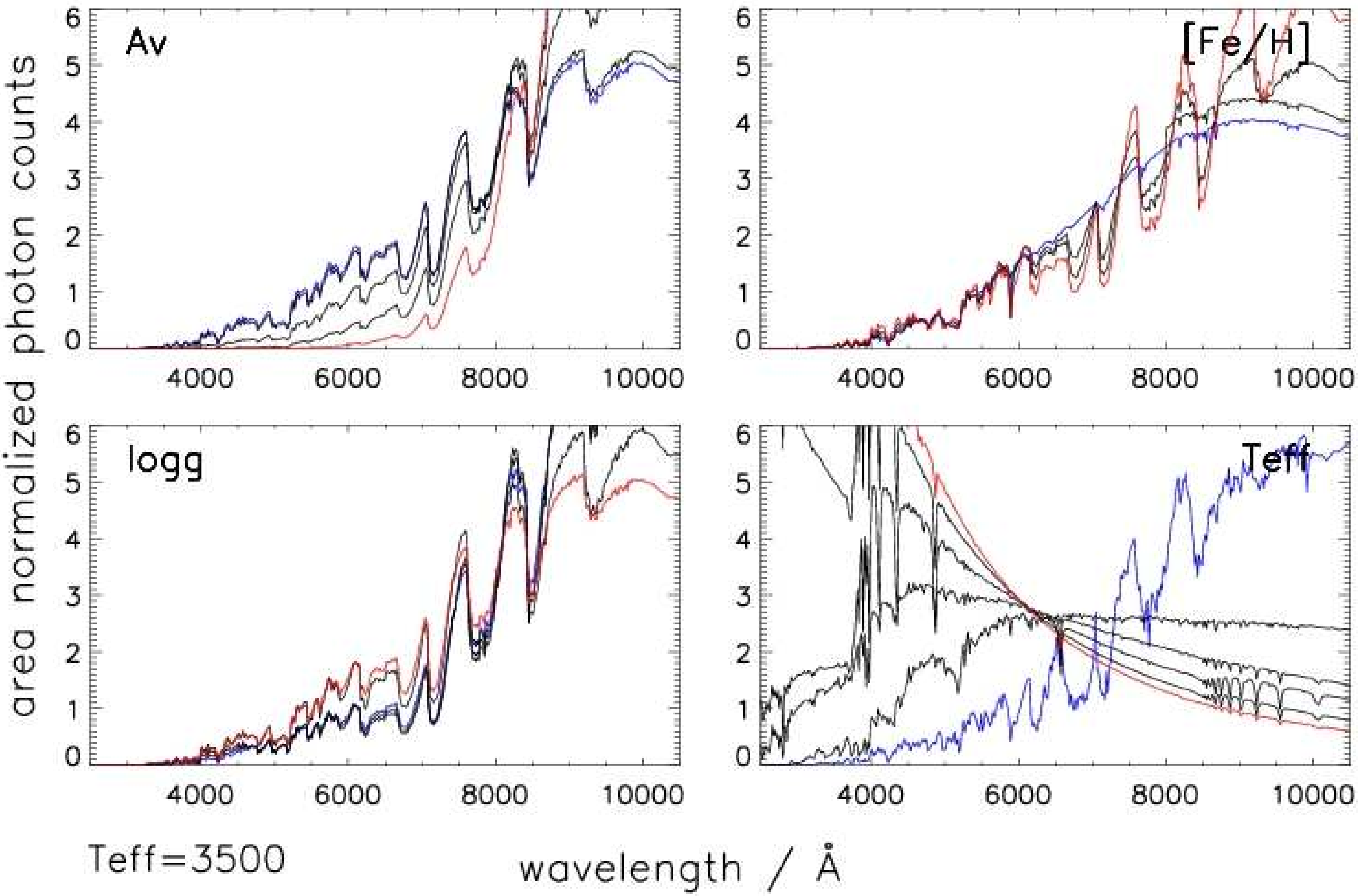
- fitness function imperfect
  - should also optimize weights for each filter
  - is single objective; problem is multiobjective
  - neighbours may not be near (use secondary grid)
  - global degeneracies not accounted for
- search mechanism could be more efficient (?)

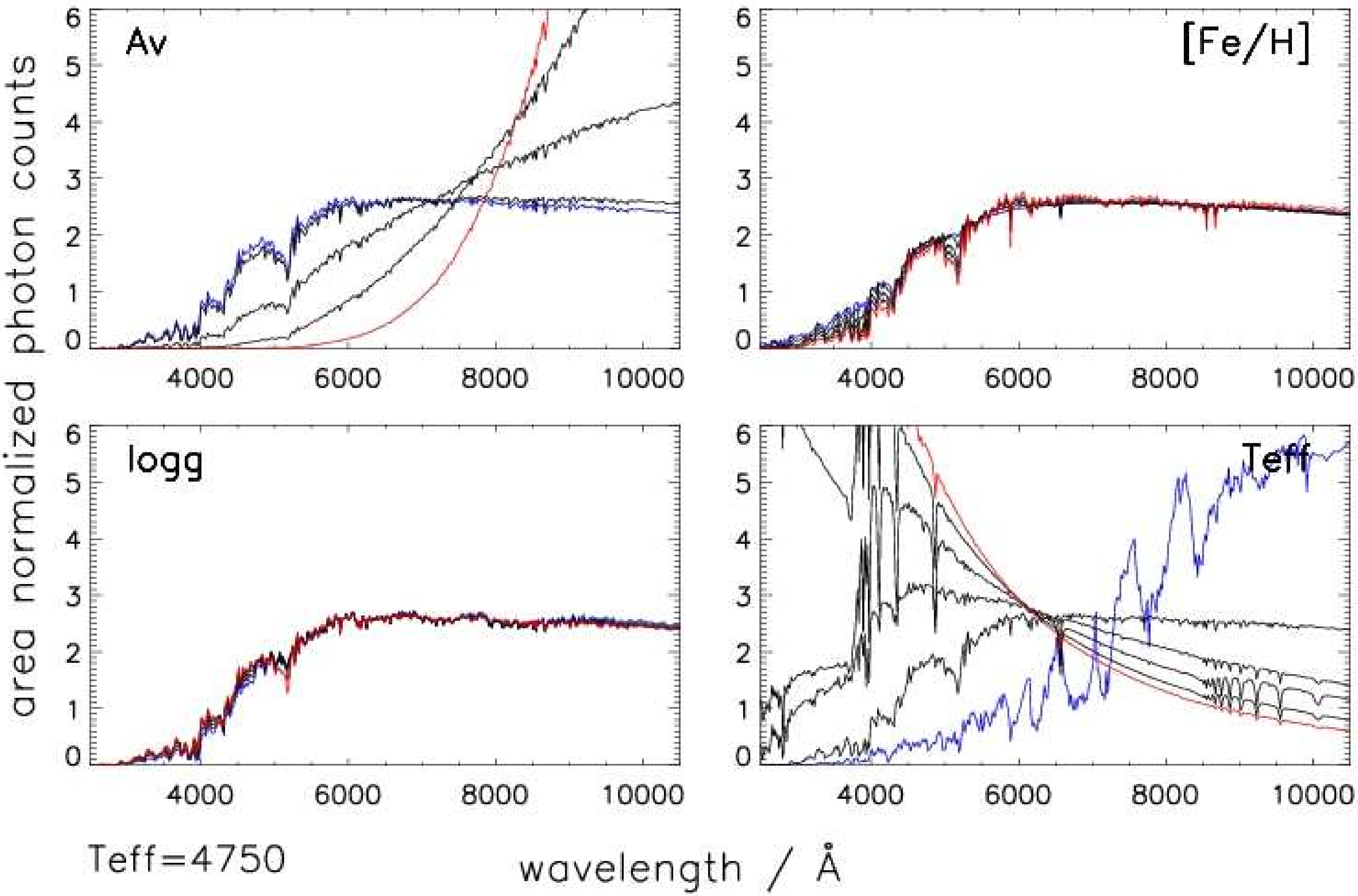
Positive points

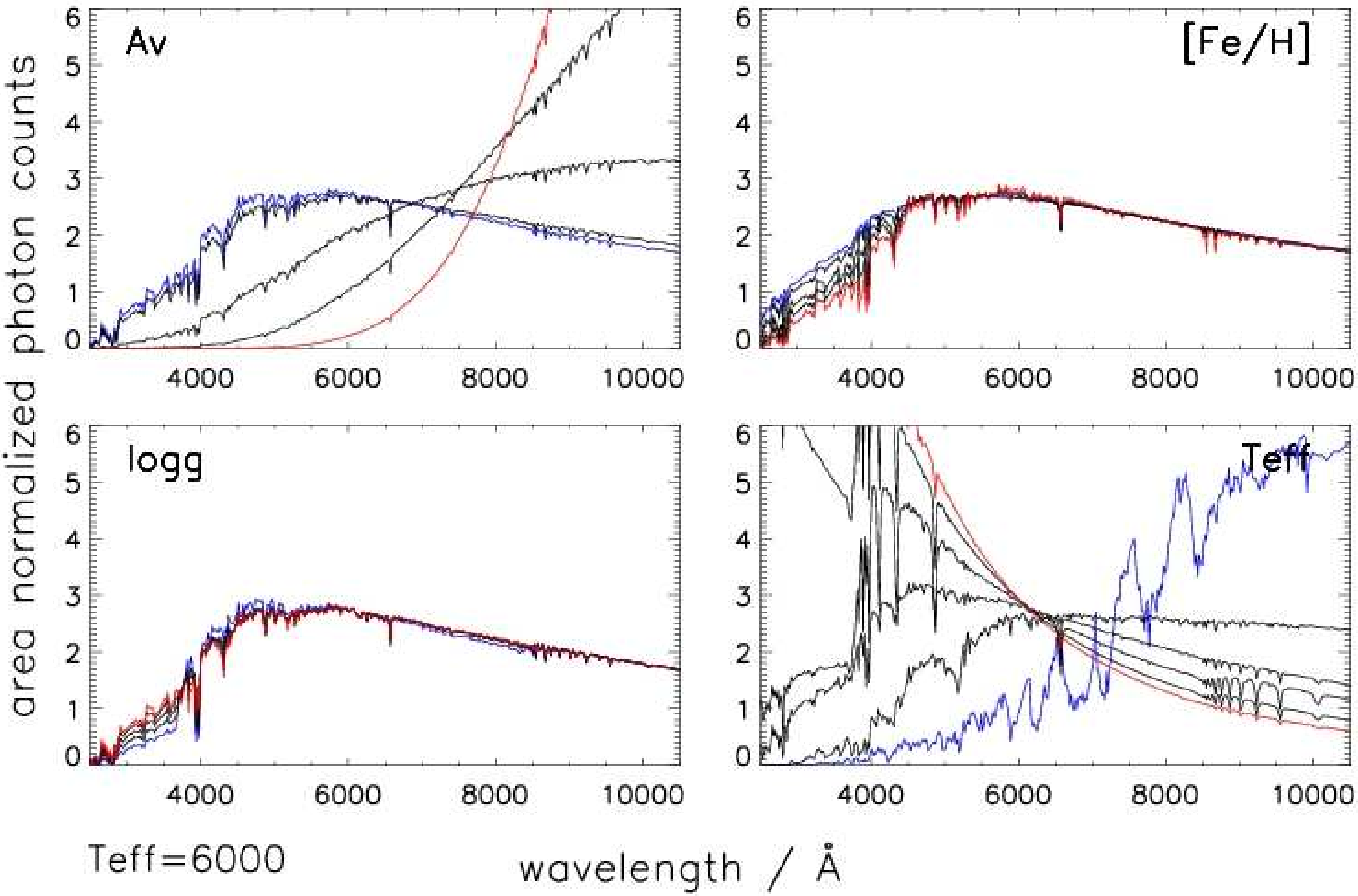
- systematic, extendable approach to filter design
- many systems tested; allows general properties to be studied
- allows simultaneous optimization of parametrizer

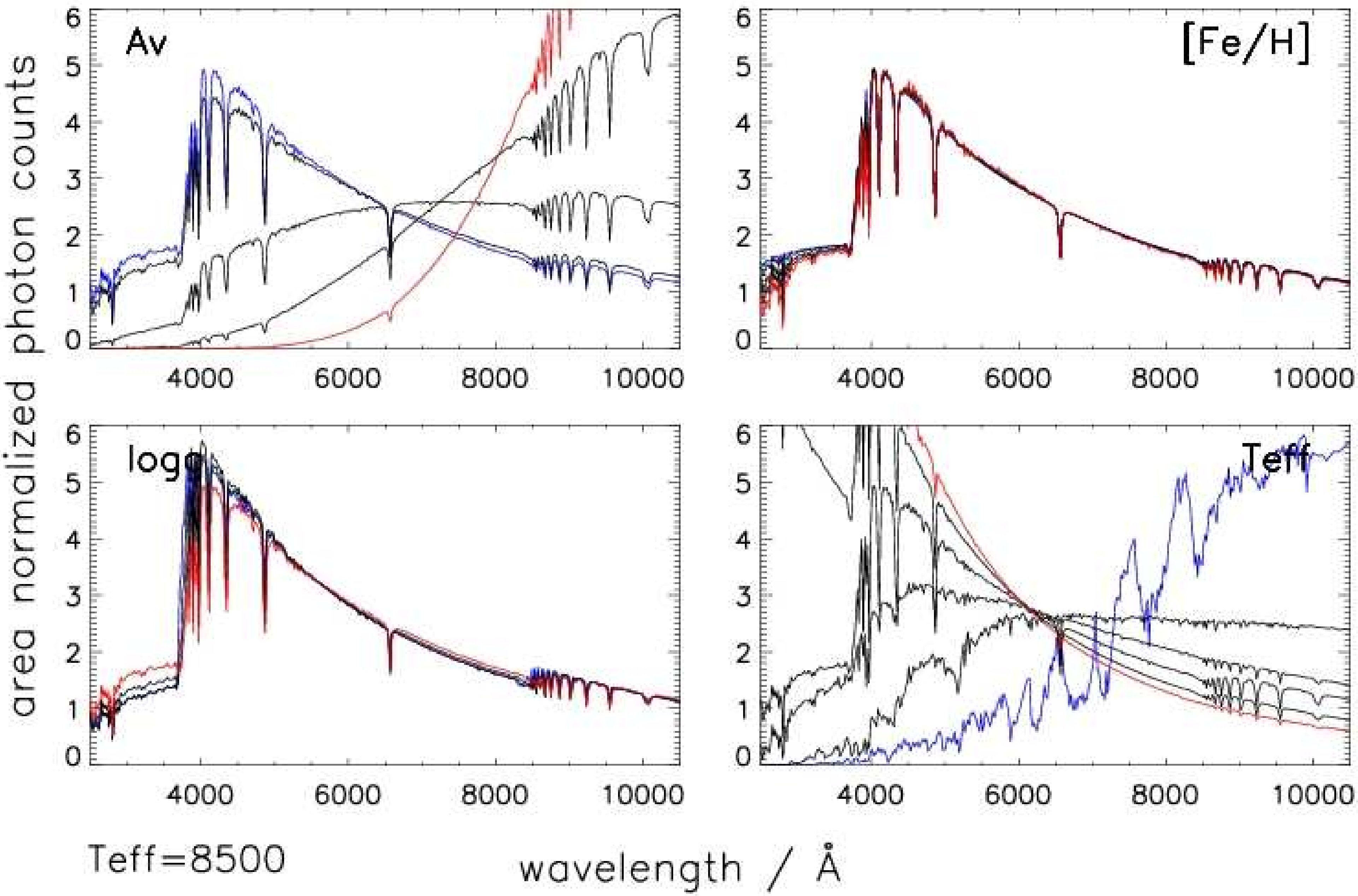
Conclusions concerning MBP

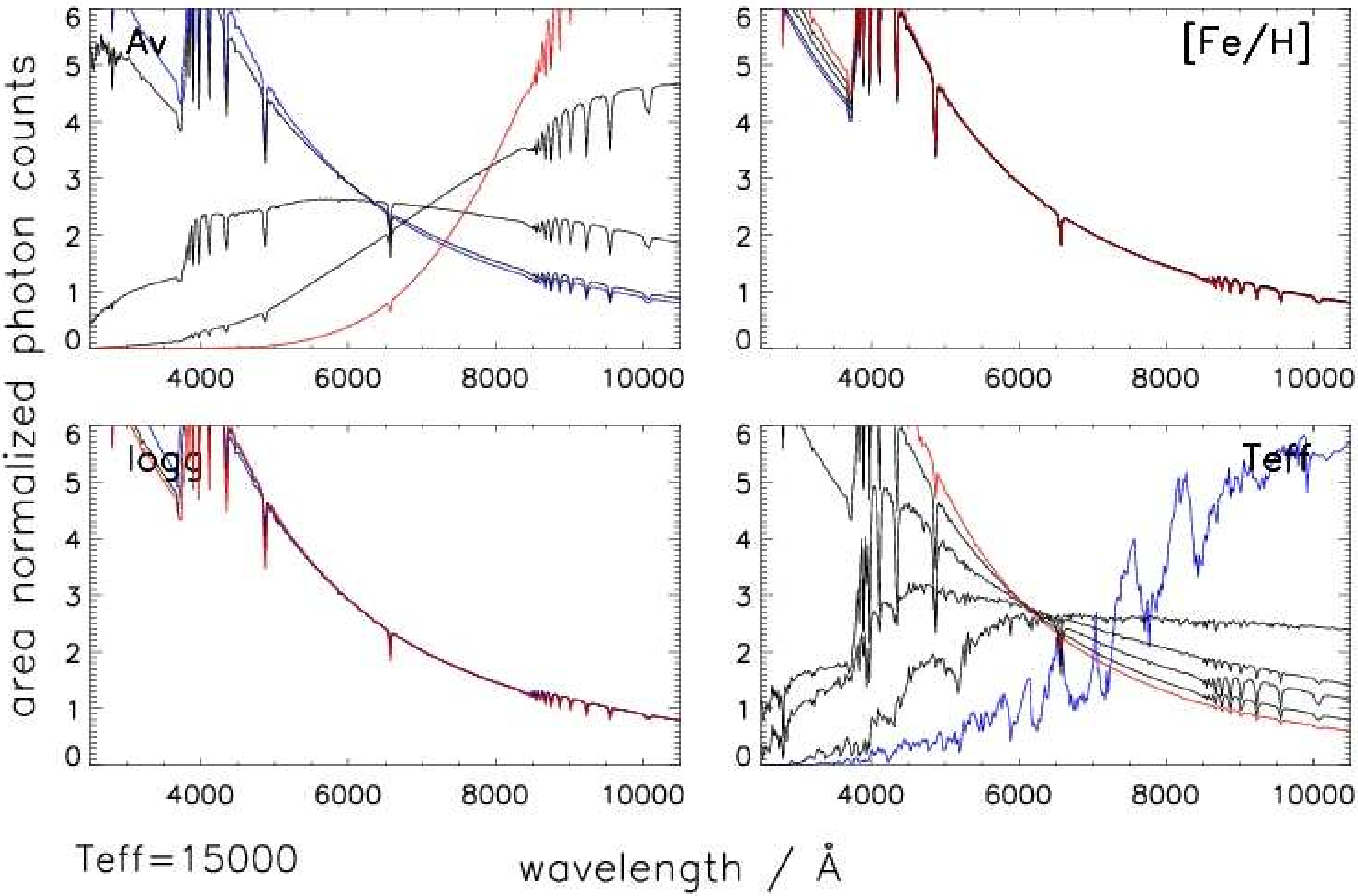
- a filter as blue as possible required with at least 2 slots
- filters extending to 9000-9500Å (but not beyond)
- not simply a uniform coverage
  - should take account of normalization (i.e. G band)
- consideration of the benefits of broad, overlapping filters
- current ratio of 10:6 blue:red CCD slots may not be ideal
- many different filter systems may have similar performance
  - therefore consider other essential requirements of MBP (e.g. discrimation against `contaminants', characterizing new objects)
- what is the optimal number of filters?


Av
[Fe/H]
logg
Teff
area normalized photon counts
wavelength / Å
Teff=3500
0
1
2
3
4
5
6
4000
6000
8000
10000


Av
[Fe/H]
logg
Teff
area normalized photon counts
wavelength / Å
Teff=4750
0
1
2
3
4
5
6
4000
6000
8000
10000


Av
[Fe/H]
logg
Teff
Teff=6000
area normalized photon counts
wavelength / Å
0
1
2
3
4
5
6
4000
6000
8000
10000


Av
[Fe/H]
logg
Teff
area normalized photon counts
wavelength / Å
Teff=8500
0
1
2
3
4
5
6
4000
6000
8000
10000


Av
[Fe/H]
logg
Teff
area normalized photon counts
wavelength / Å
Teff=15000