MODELING DARK ENERGY WITH GENETIC ALGORITHMS

Benasque 2012
9th August 2012

arXiv:1205.0364

Juan García-Bellido
IFT-UAM/CSIC
Madrid

Outline/Summary

In collaboration with Savvas Nesseris [1205.0364]

- **A decade since the discovery of acceleration, and we still don't know what Dark Energy is!**
- **We need a model-independent and bias-free reconstruction of the observables that we deduce from data.**
- **We propose a new method: Genetic Algorithms**
- **Error estimation with Path Integral formalism**
- **Show how it works for SnIa, BAO & growth data**
- **Conclusions: ΛCDM works pretty well…**

How can we infer the DE dynamics?

- Characterize observables as functions of redshift
- Operate on functions (differentiate, integrate)
- Search for relationships between functions
- Derive other physical quantities from them
- Look for consistency relations
- Only then compare with Dark Energy models

Genetic Algorithm does that for you!

Evolution at work!

How does evolution work?

- Individuals have a genome (a pool of genes)
- The genome mutates and mixes (via sex)
- There is a rate of mutation and crossover
- Next generation’s genome is different
- Natural selection “crops the tree”: speciation
- Only a few individuals (species) survive
- Some species live for a long time (survival)

How do Genetic Algorithms work?

Population (individuals)	⟺	**Functions** (cont. & differen.)
Generations (genome)	⟺	**Grammar** (poly, exp, trig, log)
Fitness (survival)	⟺	χ^2 (minimized)
Mutations (rand.)	⟺	**Operations** (rand.)
Crossover (sex)	⟺	**Recombination**
Lineage (species)	⟺	**Convergence** (limit)

How do Genetic Algorithms work?

- The grammar and the allowed operations determines how the Genetic Algorithm works
- There are two important parameters:
  - The mutation rate (5 – 10%) is the probability that an arbitrary part of a genome changes
  - The selection rate (10 – 20 %) is the fraction of individuals that finally produce offsprings
- If rates are larger then GA may not converge
- If rates are smaller then GA converges too slowly
- Important: jumps between different minima, in order to explore the whole space of functions

Advantages of Genetic Algorithms

- **No "a priori" assumptions on the DE model (e.g. # free parameters)**
- **In general DE models are not well understood (parameter space too large or too complex)**
- **OUR GOAL: minimize a functional $\chi^2(f(z))$ thru stochastic processes based on GA evolution.**
- **Similar to a Path Integral, but where dynamics is dictated by the data w/o `a priori' assumptions.**
- **Use the DATA to find the MODEL**
- **The end result doesn't have any free parameters**

An example of Genetic Algorithm

- **Assume our grammar includes basic functions: polynomials, trigonometric, exponentials, logs, …**

1.- Set up random initial population M(0), e.g.

$$f_{GA,1}(x) = \ln x$$

$$f_{GA,2}(x) = -1 + x + x^2$$

$$f_{GA,3}(x) = \sin x$$

[Note: The number of functions in a generic population is typically of a few hundreds!]

An example of Genetic Algorithm

2.- Given some data points w/ errors (x_i, y_i, σ_i), the algorithm measures the fitness of each function

$$\chi^2(f) \equiv \sum_{i=1}^{N} \left(\frac{y_i - f(x_i)}{\sigma_i} \right)^2$$

In our example:

$$\chi_1^2(x) = 200$$

$$\chi_2^2(x) = 500$$

$$\chi_3^2(x) = 1000$$

An example of Genetic Algorithm

3.- Selection is done via "tournament method", sorting population w.r.t. fitness χ^2 each individual

A fixed % total population, adjusted by selection rate, are chosen out (in our example only 2 of 3).

Reproduction = crossover & mutation

- Crossover will randomly combine parts of parent solutions

$$\begin{aligned} f_{GA,1}(x) \oplus f_{GA,2}(x) &\rightarrow (\bar{f}_{GA,1}(x), \bar{f}_{GA,2}(x), \bar{f}_{GA,3}(x)) \\ &= \left(\ln(x^2), -1 + \ln(x^2), -1 + \ln(x)\right) \end{aligned}$$

An example of Genetic Algorithm

- Mutation will randomly change the functions, according to mutation rate, e.g. exponent of some terms or the value of a coefficient.

$$\bar{f}_{GA,3}(x) = -1 + \ln(x) \rightarrow -1 + \ln(x^3)$$

Then the next generation becomes:

$$\begin{aligned} M(1) \;&=\; (\bar{f}_{GA,1}(x),\ \bar{f}_{GA,2}(x),\ \bar{f}_{GA,3}(x)) = \\ &\qquad (\ln(x^2), -1 + \ln(x^2), -1 + \ln(x^3)) \end{aligned}$$

An example of Genetic Algorithm

3.- Back to beginning: evaluate fitness χ^2 of each individual:

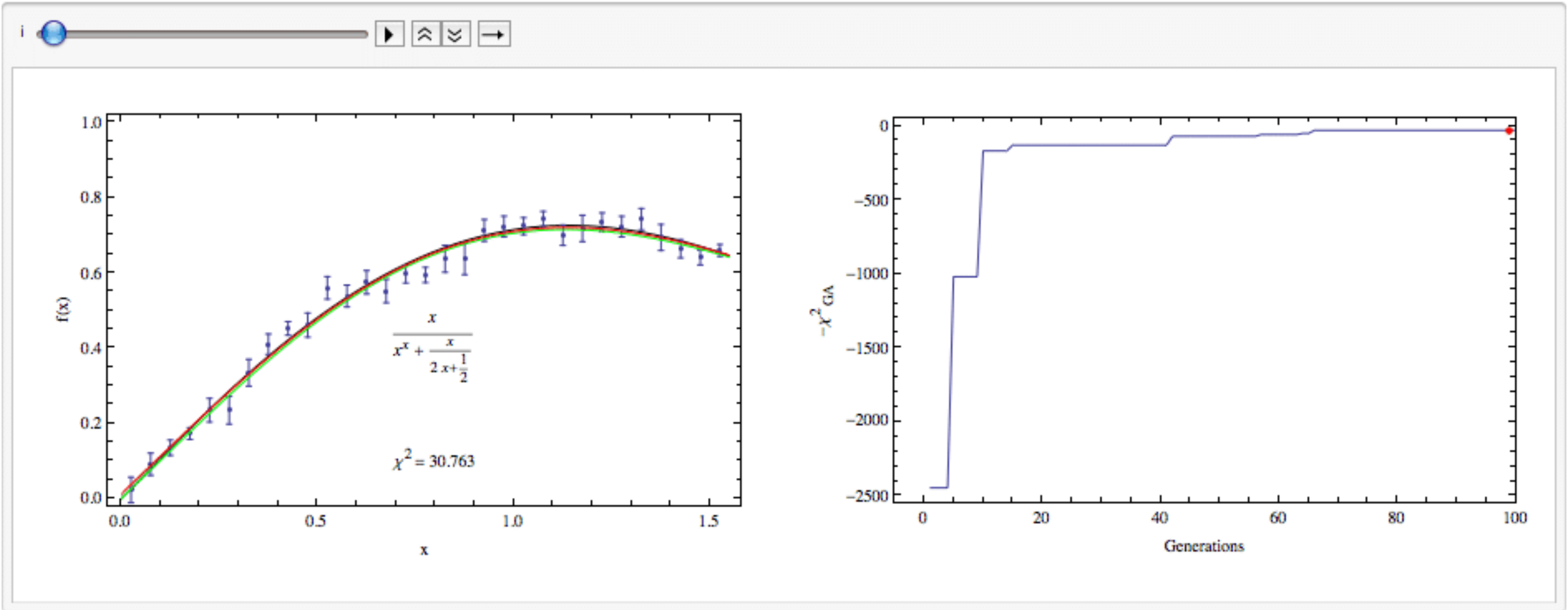
$$\chi_1^2(x) = 150, \quad \chi_2^2(x) = 300, \quad \chi_3^2(x) = 400$$

Even after a single generation, the combined χ^2 of each individual has decreased significantly!

It usually takes ~ 100 generations to reach a reasonable χ^2 and ~ 500 generations to converge to a minimum χ^2, and ~ 1000 generations until a predetermined goal has been reached.

An example of Genetic Algorithm

Input function $$f(x)=\frac{1}{4}+\left(x-\frac{1}{4}\right)\exp\left[-\frac{1}{2}x^2\right]$$



Best χ^2 function $$f_{best\ \chi^2}(x)=x/\left(x^x+2x/(4x+1)\right)$$

Error estimation for GAs

Construct a functional $\mathcal{L} = \mathcal{N} \exp\left(-\chi^2(f)/2\right)$

where $$\chi^2(f) \equiv \sum_{i=1}^{N} \left(\frac{y_i - f(x_i)}{\sigma_i} \right)^2$$

Properly normalized functional

$$\int \mathfrak{D}f \; \mathcal{L} = \int \mathfrak{D}f \; \mathcal{N} \exp\left(-\chi^2(f)/2\right) = 1$$

where $$\mathcal{N} = \left((2\pi)^{N/2} \prod_{i=1}^{N} \sigma_i \right)^{-1}$$

Error estimation for GAs

The 1σ error around the best-fit function (BFF)

$$
\begin{aligned}
CI(x_i, \delta f_i) &= \int_{f_{bf}(x_i)-\delta f_i}^{f_{bf}(x_i)+\delta f_i} df_i \frac{1}{(2\pi)^{1/2}\sigma_i} \exp\left(-\frac{1}{2}\left(\frac{y_i - f_i}{\sigma_i}\right)^2\right) \\
&= \frac{1}{2}\left(\mathrm{erf}\left(\frac{\delta \mathrm{f_i} + \mathrm{f_{bf}}(\mathrm{x_i}) - \mathrm{y_i}}{\sqrt{2}\sigma_\mathrm{i}}\right) + \mathrm{erf}\left(\frac{\delta \mathrm{f_i} - \mathrm{f_{bf}}(\mathrm{x_i}) + \mathrm{y_i}}{\sqrt{2}\sigma_\mathrm{i}}\right)\right) \\
&= \mathrm{erf}\left(1/\sqrt{2}\right)
\end{aligned}
$$

at each point $\mathbf{x_i}$

Then minimize to get δf(x) – a band around BFF

$$
\chi^2_{CI}(\delta f_i) = \sum_{i=1}^{N}\left(CI(x_i, \delta f_i) - \mathrm{erf}\left(1/\sqrt{2}\right)\right)^2
$$

Concrete Example for DE

Data sets and models

Dataset	Information	FRW/DE models	LTB models
SnIa	$d_L(z)$	$(1+z)\int_0^z \frac{dx}{H(x)/H_0}$	$(1+z)^2 A(r(z),t(z))$
BAO	$H(z), d_A(z)$	$\frac{l_{BAO}(z_{drag})}{D_V(z)}$	$(1+z)\xi(z)\frac{l_{BAO}(r_\infty,t_0)}{D_V(z)}$ [a]
Growth-rate $f=\frac{dln(\delta)}{dlna}$	γ,$\Omega_{\rm m}(a)$	$\Omega_{\rm m}(a)=\frac{\Omega_{0\rm m}a^{-3}}{H(a)^2/H_0^2}$	$(1-\Omega_M^{-1}(r))\frac{A(r,t)}{r}$ [b]

[a] where in both cases $D_V(z)=\left((1+z)^2 D_A(z)^2 \frac{z}{H(z)}\right)^{1/3}$ and $H(z)=H_L(z)$ for the LTB
[b] where $\Omega_M(r)$ is the mass radial function.

Genetic Algorithm data functions

SnIa – luminosity distance

$$GA_1(z) \equiv D_L(z) = d_L(z)/H_0^{-1}$$

BAO – angular diameter distance

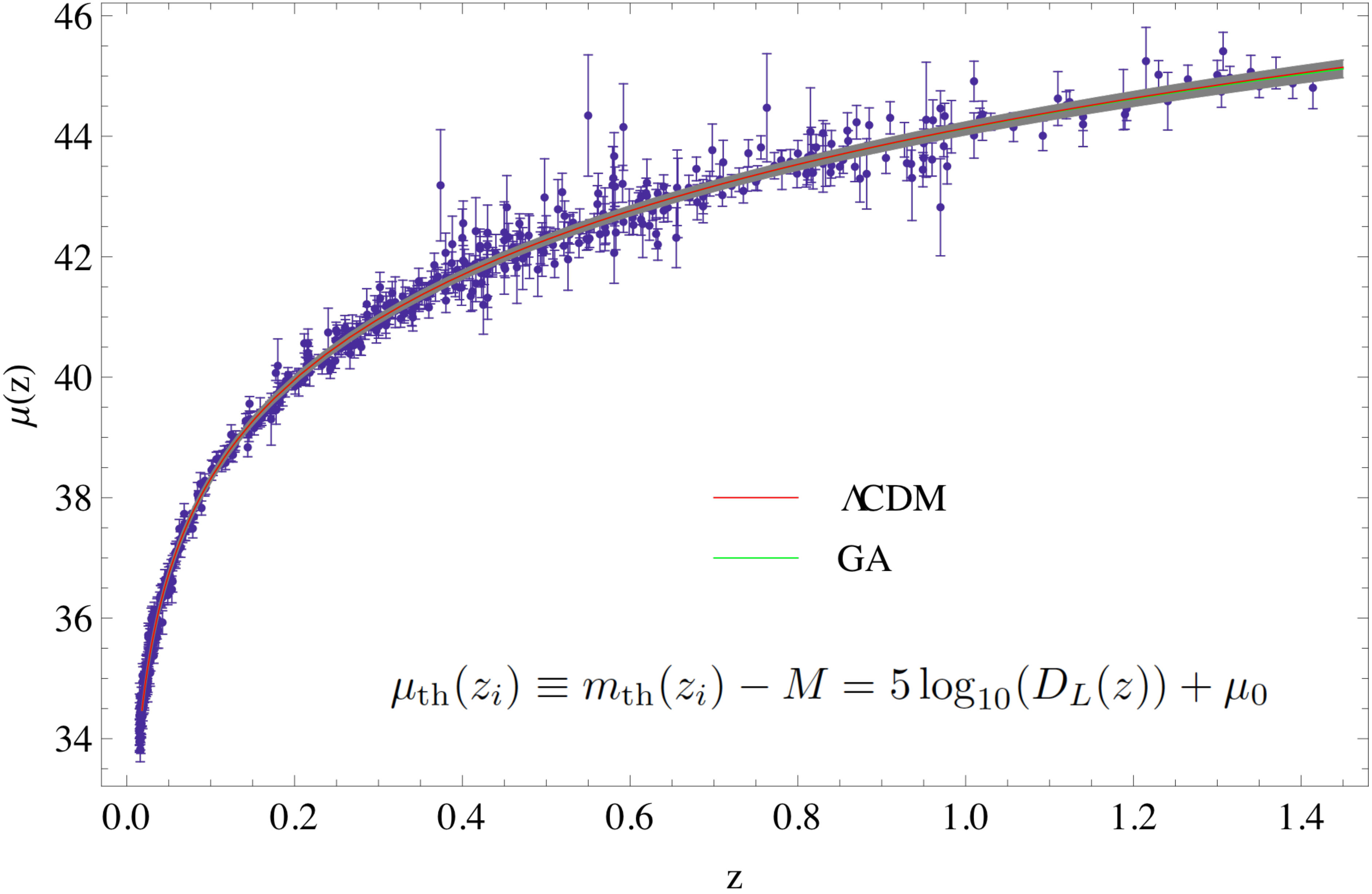
$$d_z(z) = \frac{l_{BAO}(z_{drag})}{D_V(z)} = \frac{l_{BAO}(\Omega_{0\mathrm{m}} = 1)}{z} GA_2(z)$$

Growth rate of matter structures

$$f\sigma8(z) \equiv f(z) \cdot \sigma_8(z) = GA_3(z)$$

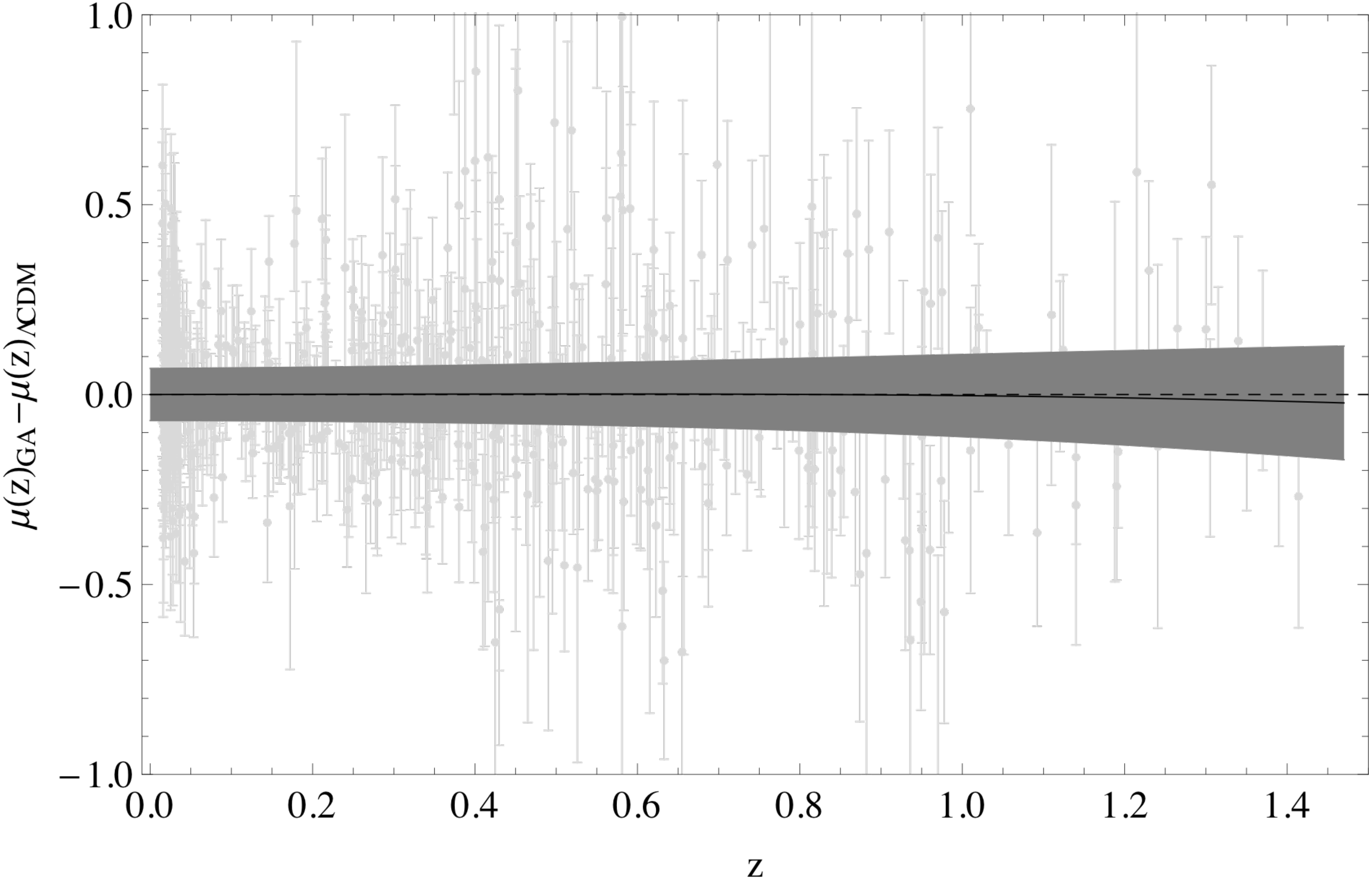
Genetic Algorithms & SnIa data

Union2.1, Suzuki et al. ApJ 746, 85 (2012)

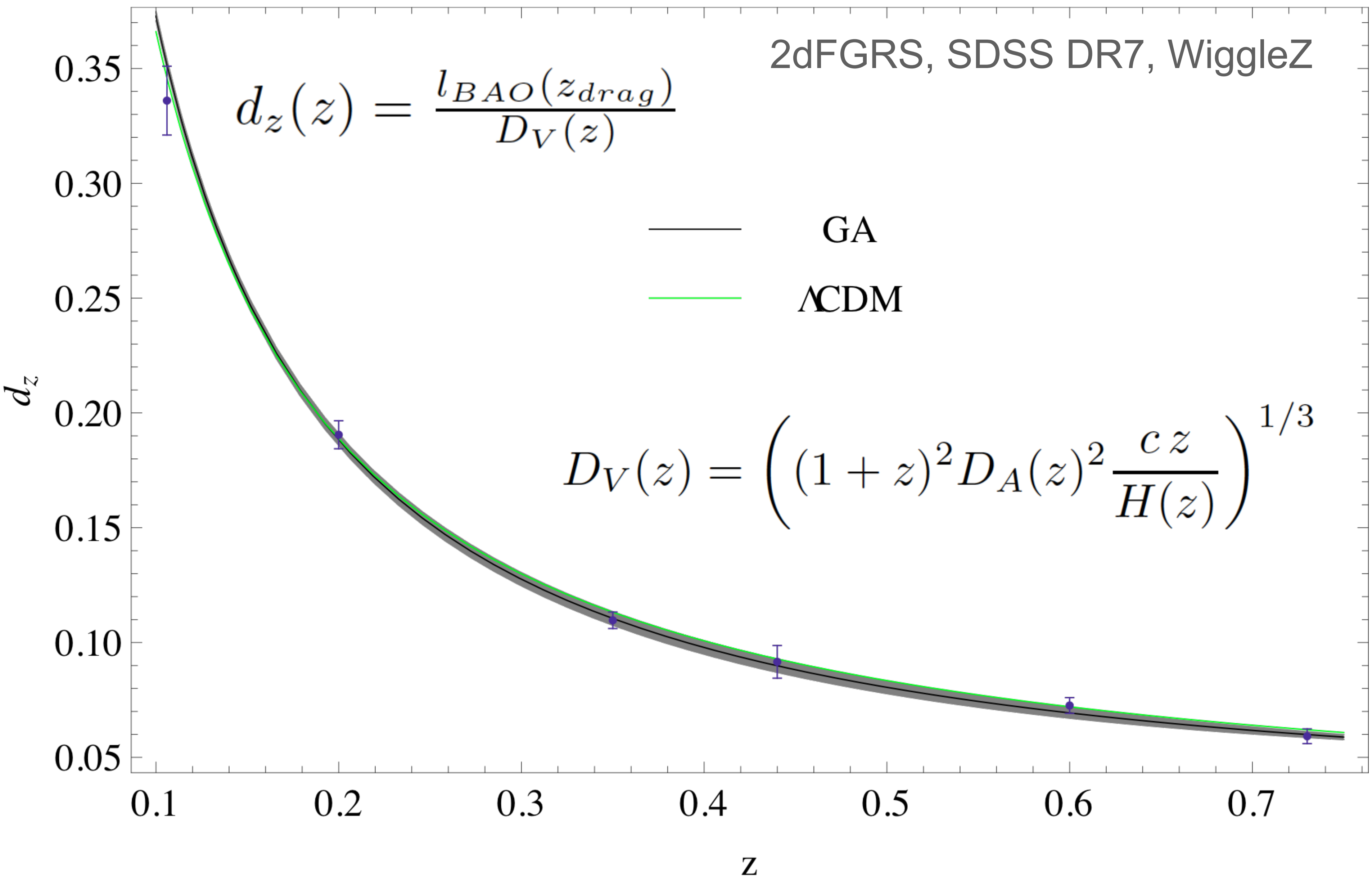


$$\mu_{\rm th}(z_i) \equiv m_{\rm th}(z_i) - M = 5\log_{10}(D_L(z)) + \mu_0$$

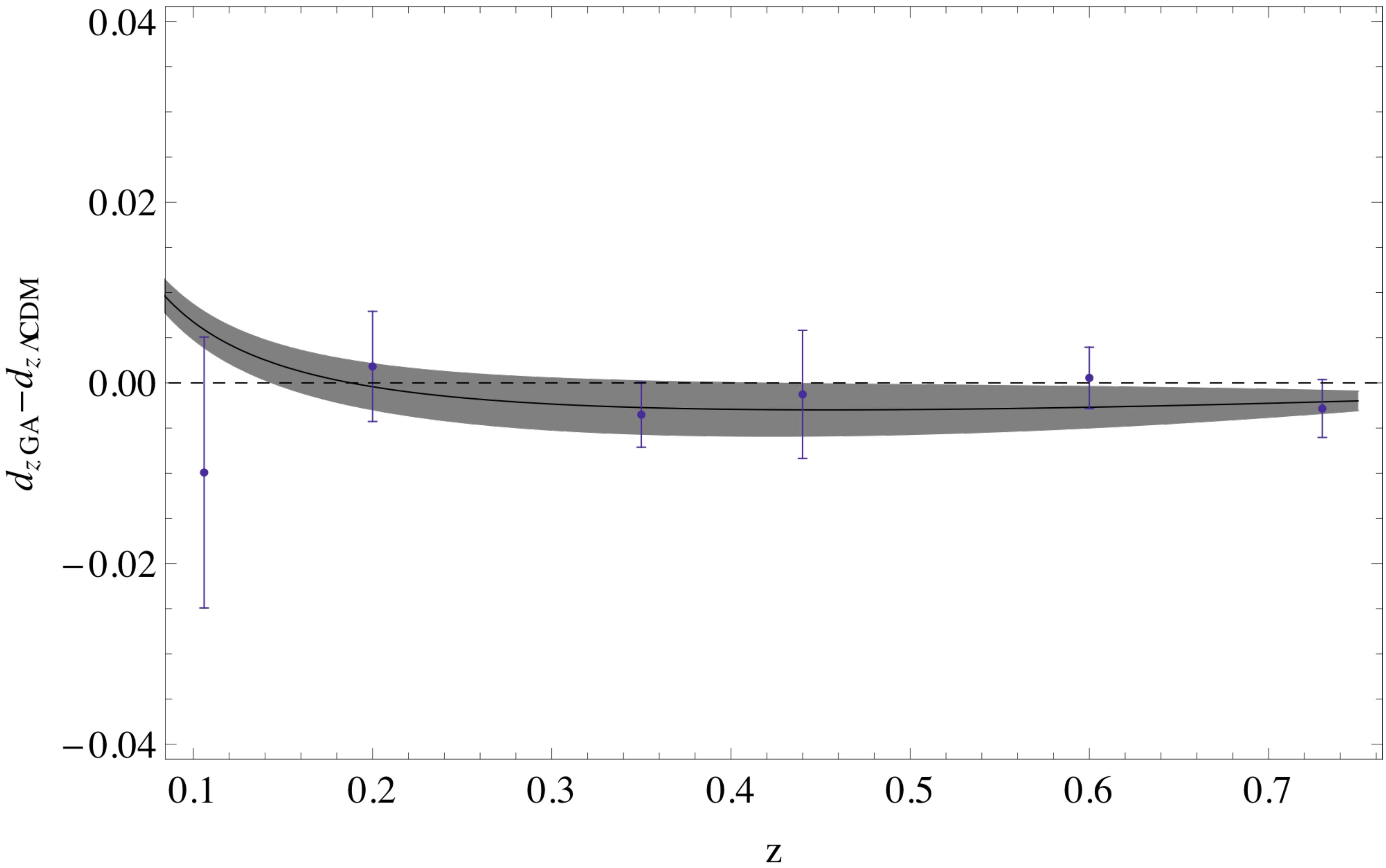
Genetic Algorithms & SnIa data



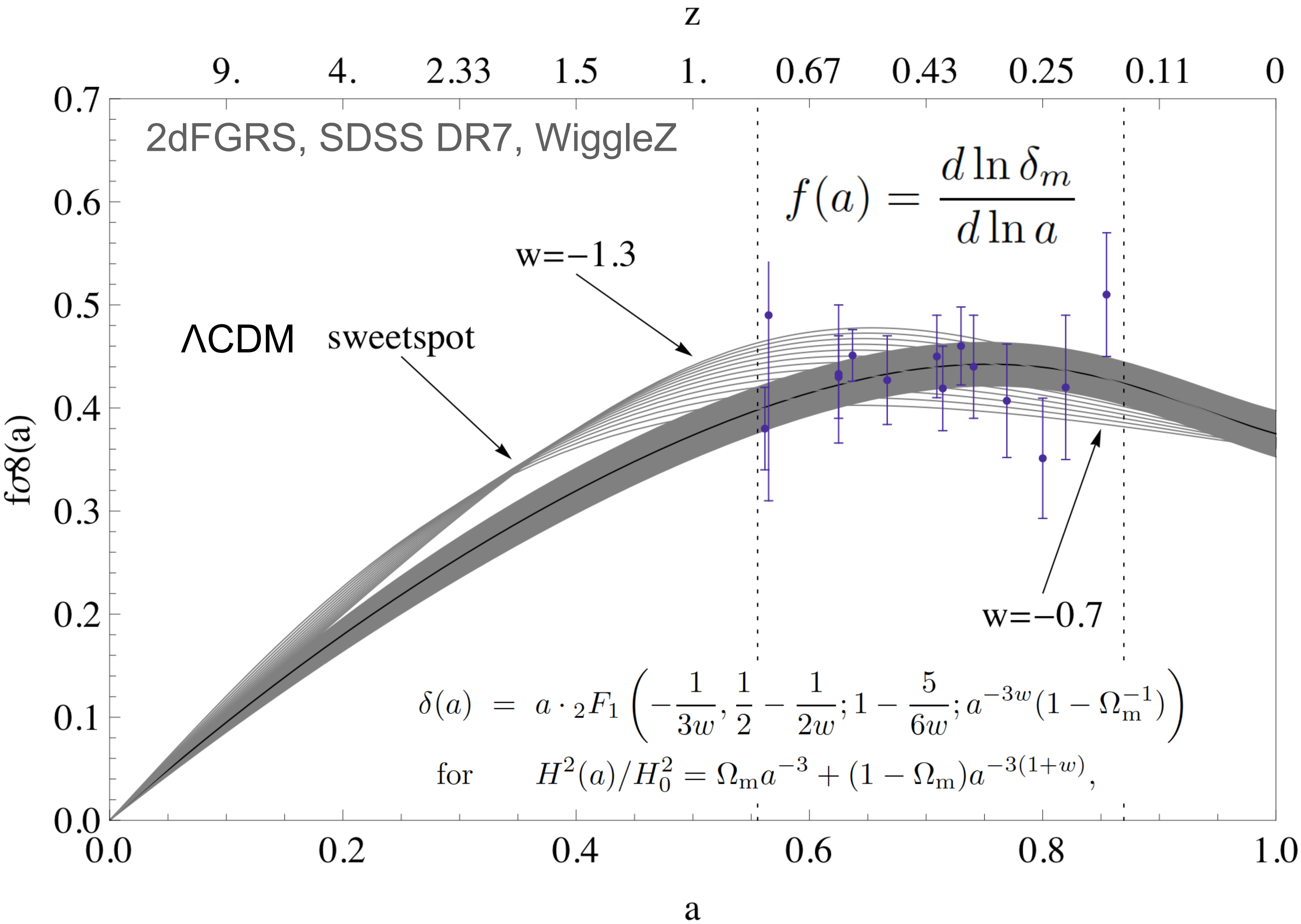
Genetic Algorithms & BAO data



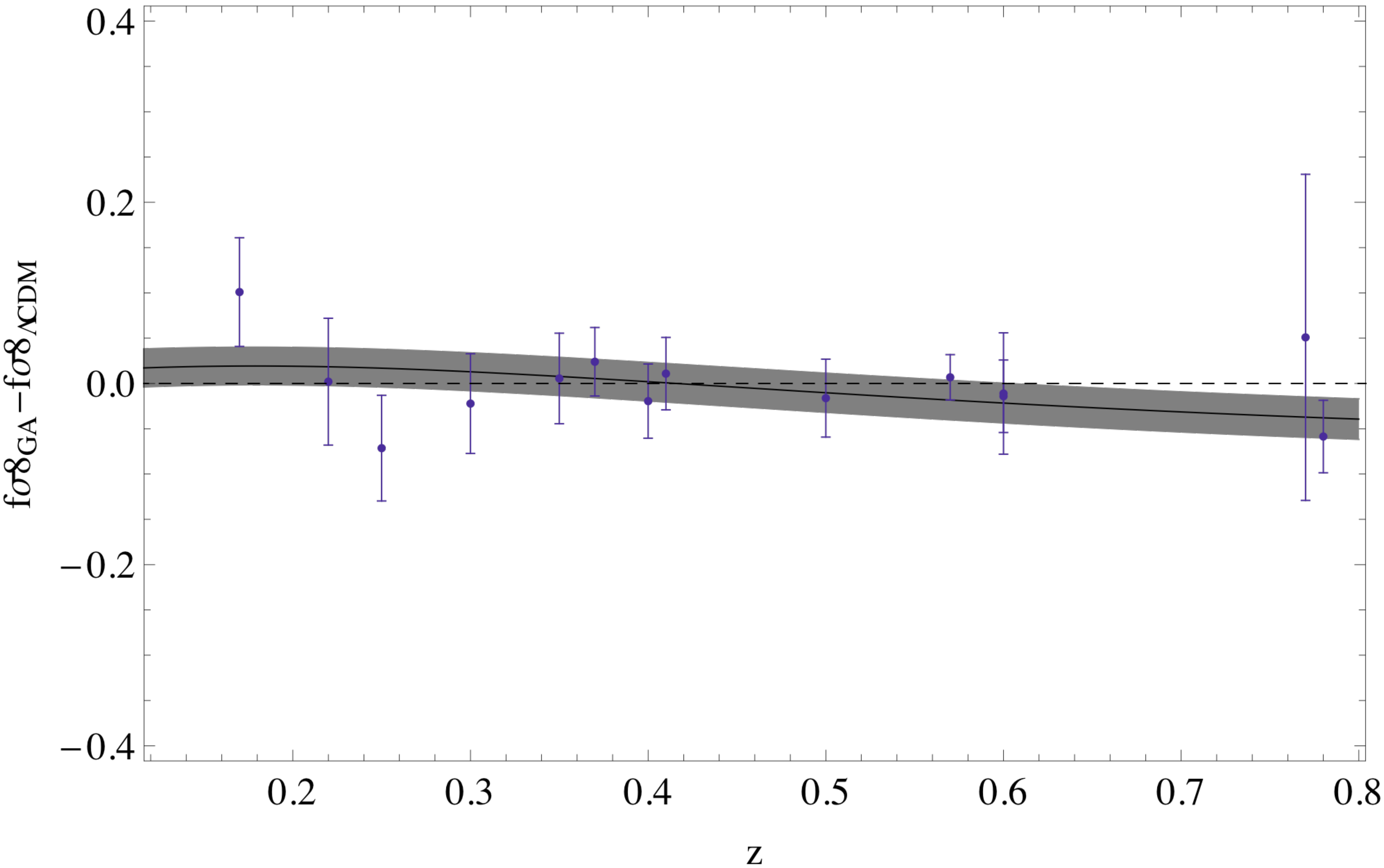
Genetic Algorithms & BAO data



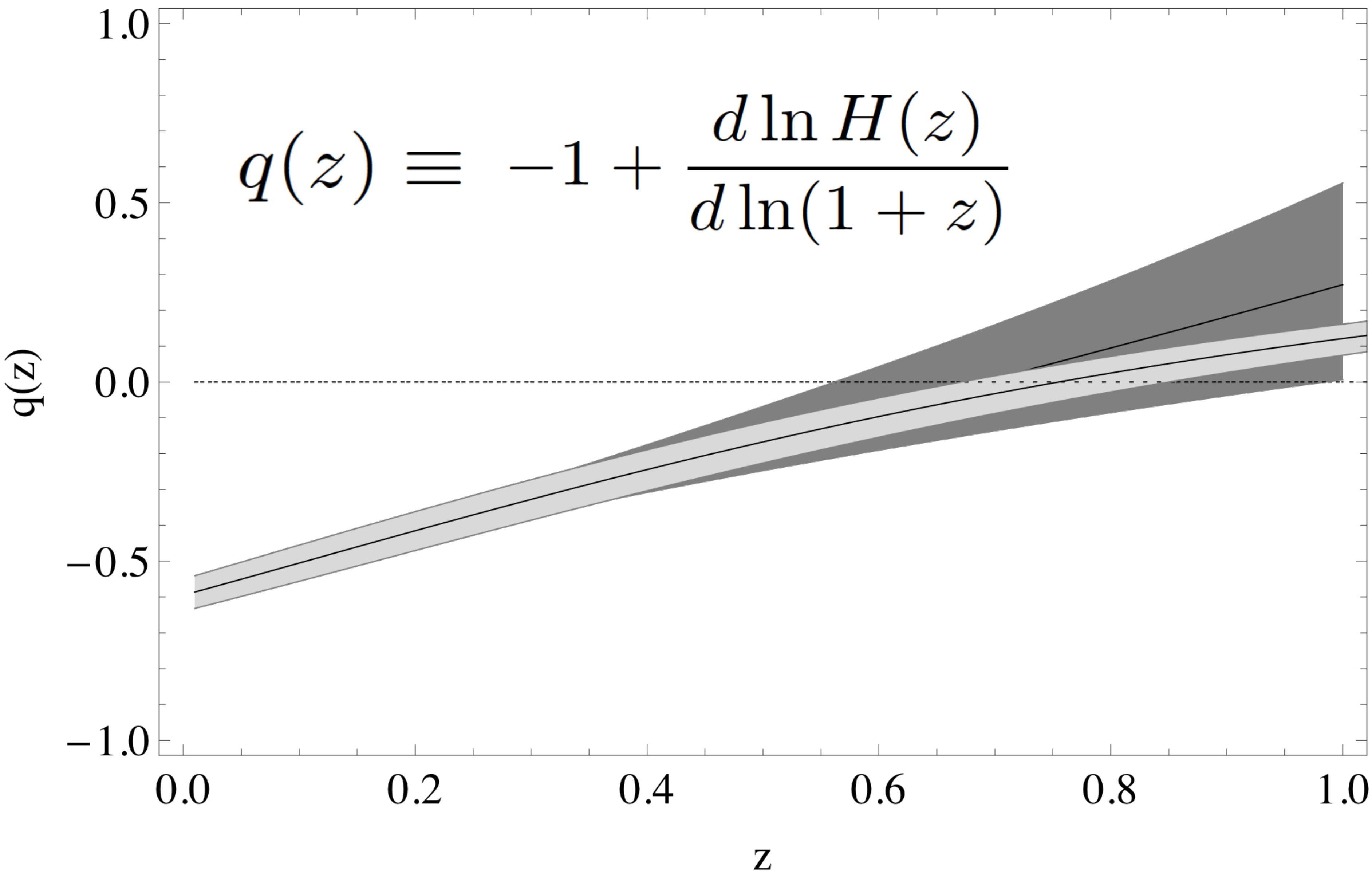
Genetic Algorithms & growth data



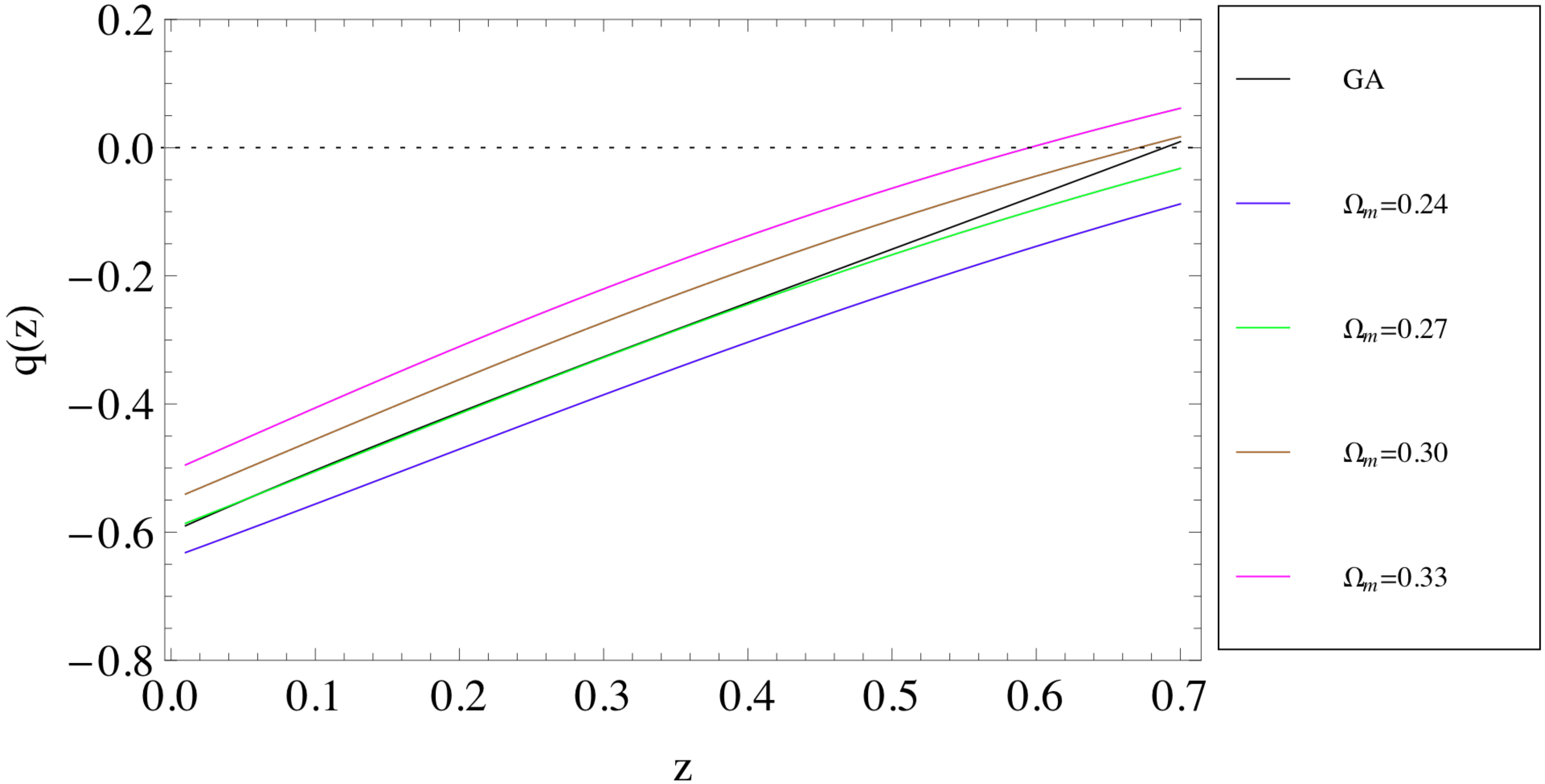
Genetic Algorithms & growth data



Genetic Algorithms & derived param.

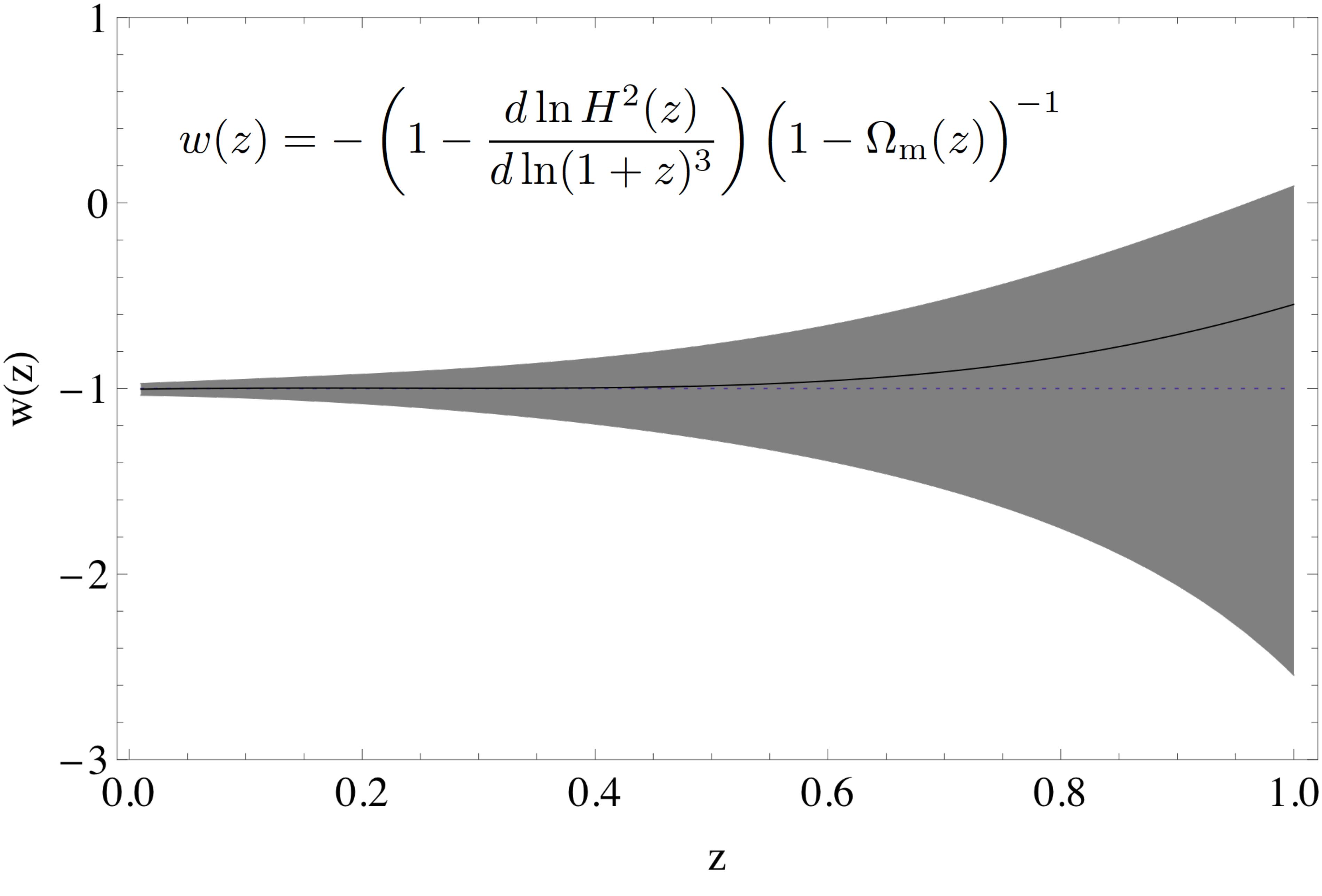


Genetic Algorithms & derived param.

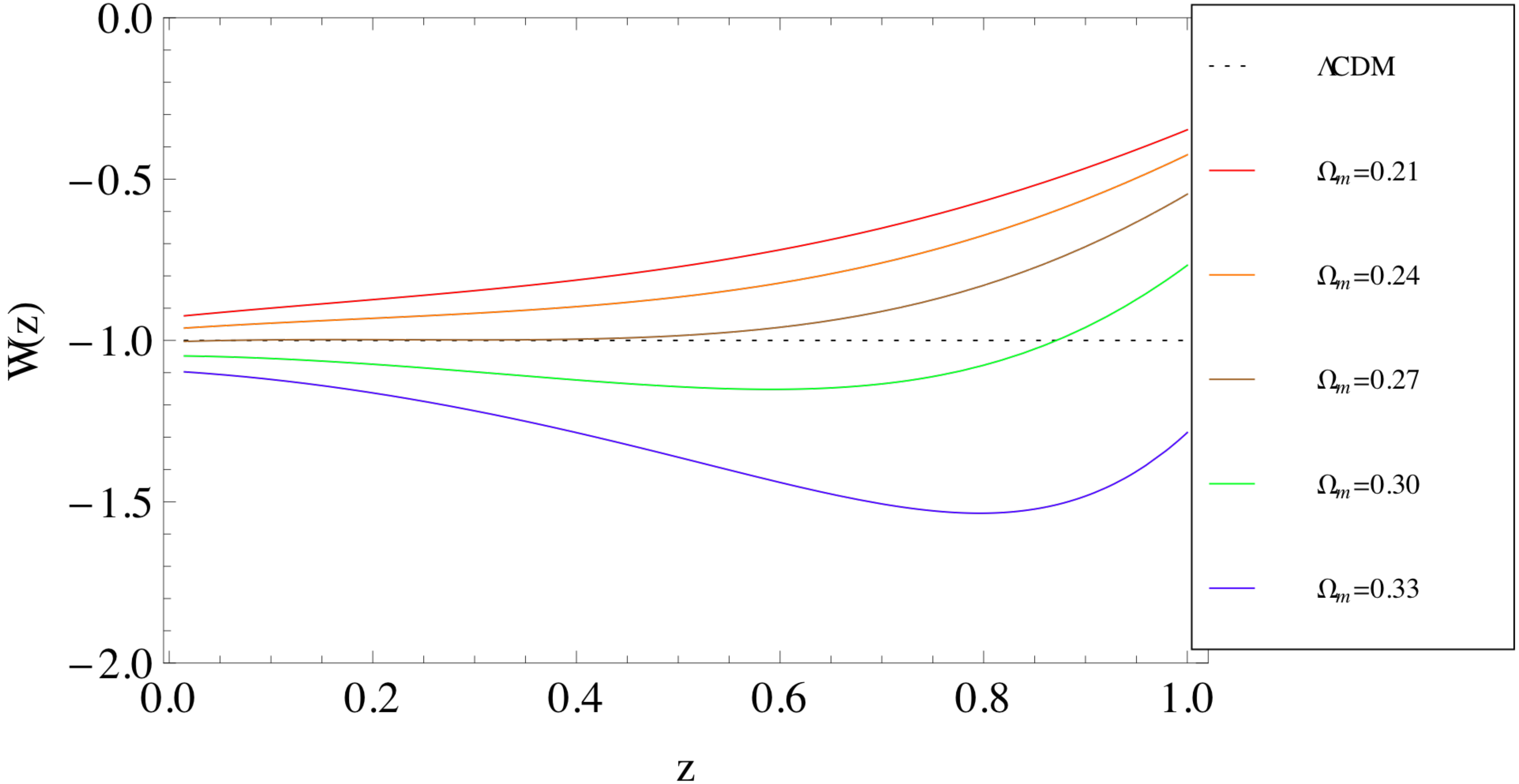


$$q(z) \equiv -1 + \frac{d \ln H(z)}{d \ln(1+z)}$$

Genetic Algorithms & derived param.

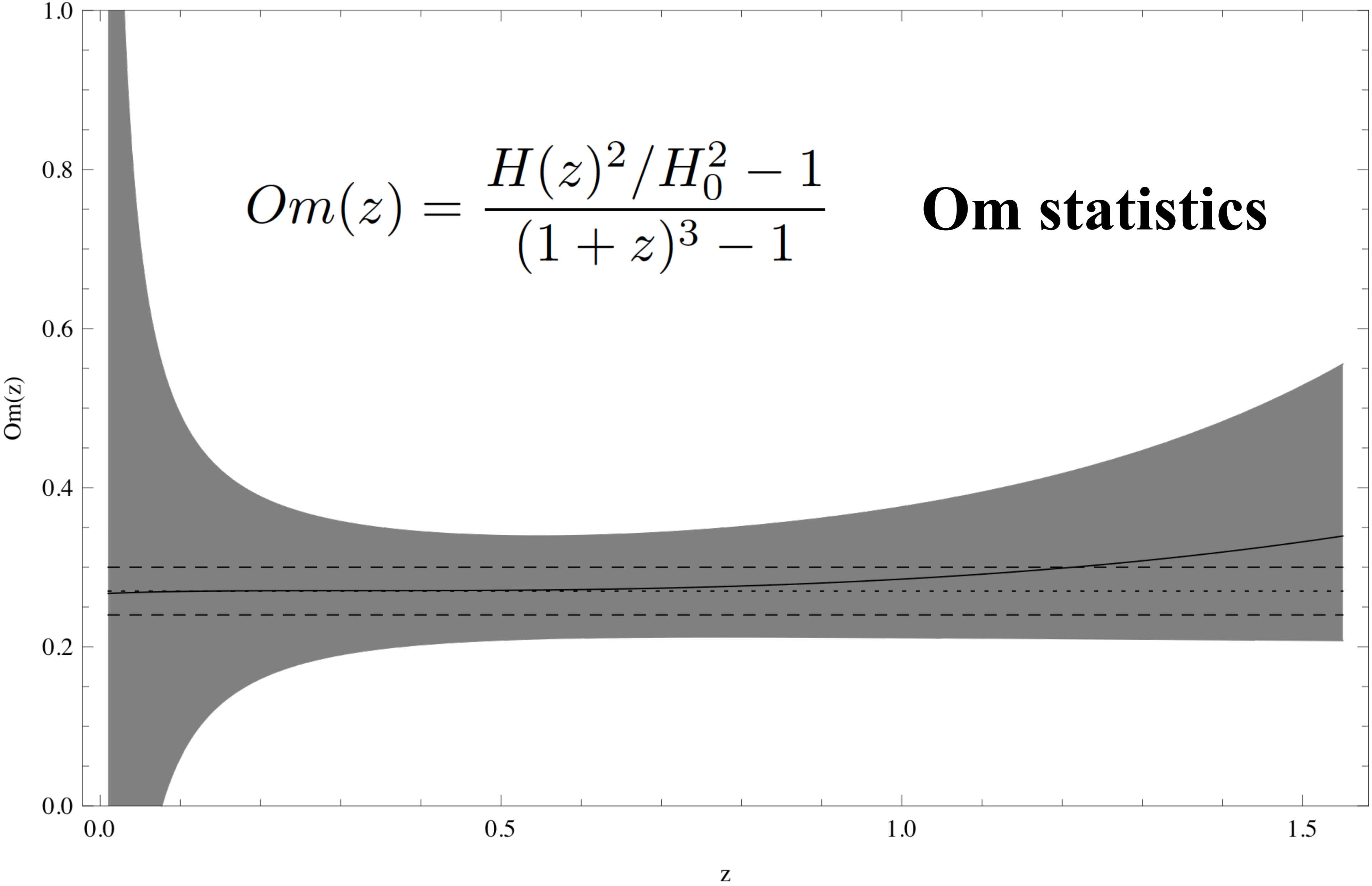


Genetic Algorithms & derived param.

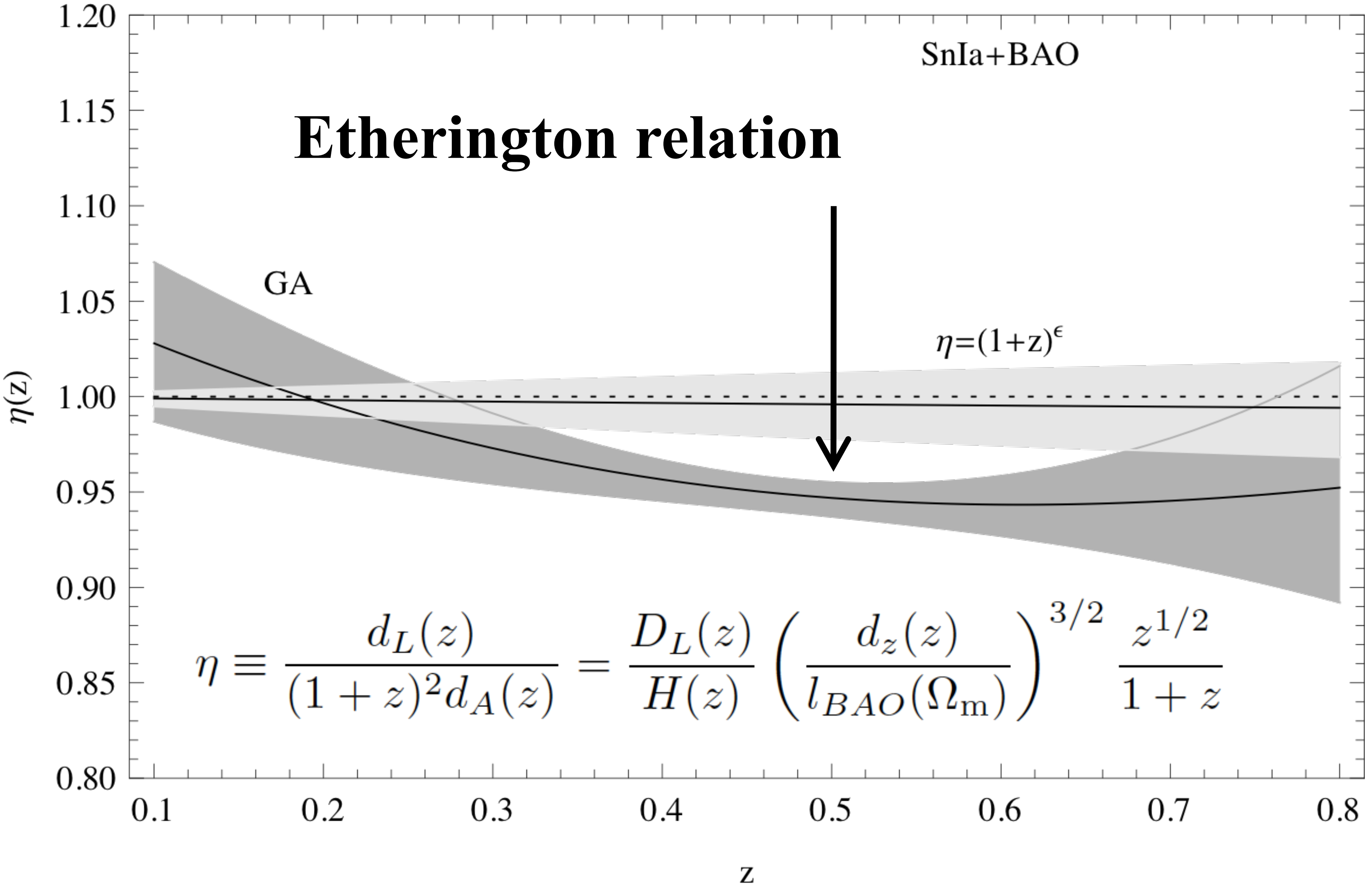


$$w(z) = -\left(1 - \frac{d\ln H^2(z)}{d\ln(1+z)^3}\right)\left(1 - \Omega_\mathrm{m}(z)\right)^{-1}$$

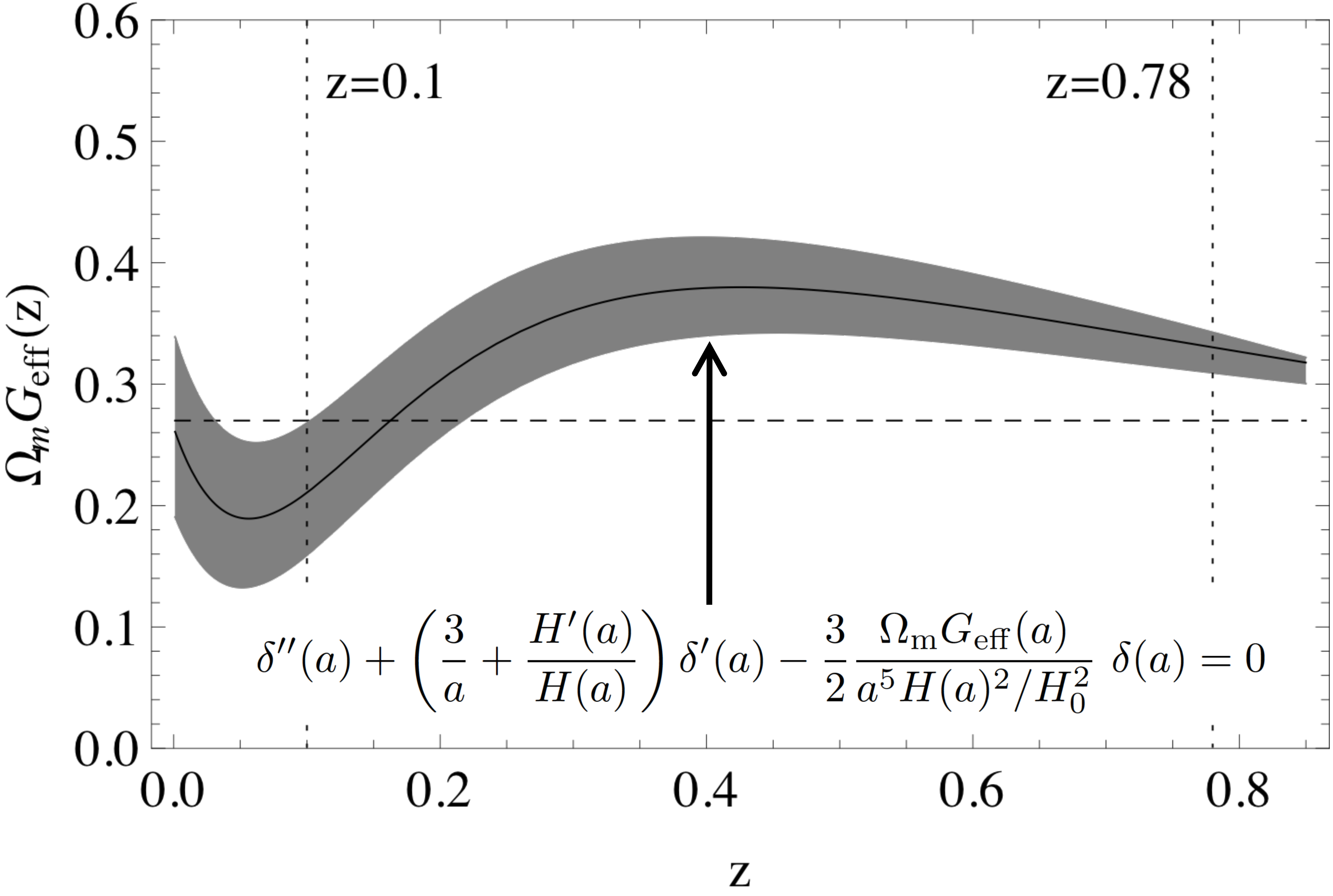
Genetic Algorithms & derived param.



Genetic Algorithms & derived param.



Genetic Algorithms & derived param.



Conclusions

- **A decade since the discovery of acceleration, we still don't know what Dark Energy is!**
- **We need a model-independent and bias-free reconstruction of the observables that we deduce from data.**
- **We propose a new method: Genetic Algorithms**
- **Error estimation with Path Integral formalism**
- **Show how it works for SnIa, BAO & growth data**
- **ΛCDM works pretty well…**
- **Find slight deviations at 3σ on $\Omega_m G_{eff}$ @ z=0.5 and the Etherington relation**