A Statistical Study of the Heterotic Landscape

Michael Lennek
University of Arizona
Santa Fe Workshop

with K. Dienes, D. Sénéchal, & V. Wasnik
hep-th/0608nnn (to appear)

Why do Statistical Studies?

- Guide for Model Builders
  - determine the easiest ways to obtain certain space-time properties
- Possible method to extract phenomenological predictions from string theory
- Hypothesis generation for new string properties and correlations

Limitations of statistical studies

- Results only Statistical (not absolute results)
- “Lamppost” Problem (can only explore certain parts of Landscape)
- “Bull’s Eye” Problem (not always clear what the target is)
- Statistical Bias effects (may not even explore space randomly)

Landscaping

- Lots of theoretical speculation on the form of the String Landscape
- Few actual statistical studies of the landscape (Dijkstra **et al.** hep-th/0411129, Blumenhagen **et al.** hep-th/0510170, Dienes. hep-th/0602286)
- What do we find when we look at the space of actual string models that can be constructed and analyzed?

Why study the Heterotic String Landscape?

- Models generically more constrained than Type I models
- Lots of positive phenomenological features (gauge coupling unification, rich massless spectrum)
- Very different mechanism for generating gauge groups, thus correlations are expected to be different

How do we distinguish models?

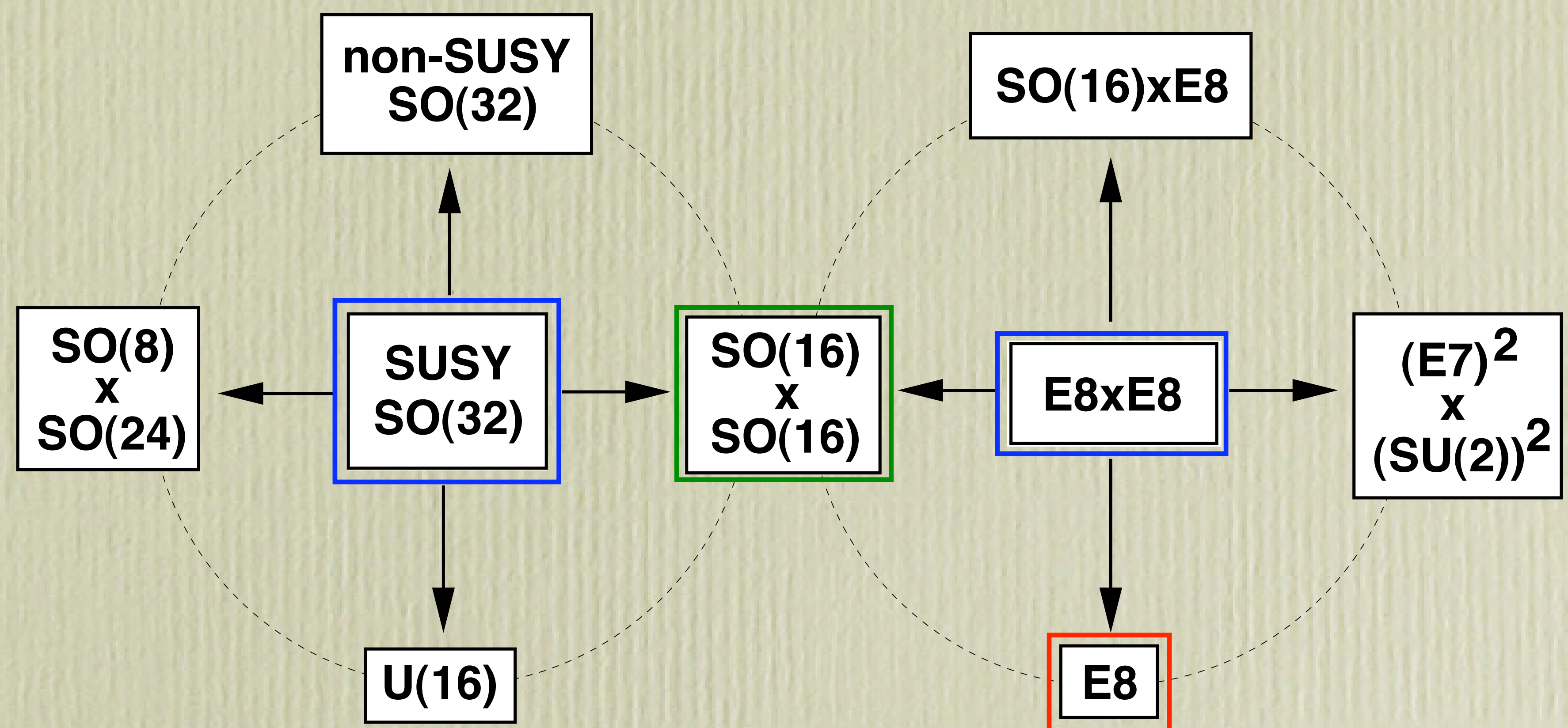
- Characteristics in space-time:
  - Particle spectrum
  - Gauge group
  - Number of SUSY Generators

If any quantity is ***different*** then the model is considered distinct

Heterotic String Models in D=10

- Only nine unique models
- Maximal SUSY is N=1
- Large variation in gauge groups
- Rank of gauge group is ≤16

Orbifold Relations amongst models



N=1 SUSY

Rank ≠ 16

→ = Z_2 Orbifold

N=0 tachyon-free

Shatter

Shatter level:	$N = 1$ SUSY	$N = 0$ SUSY Tachyon-free	$N = 0$ SUSY Tachyonic
1	$SO(32)$		$SO(32)$, E_8
2	$E_8 \times E_8$	$SO(16) \times SO(16)$	$SO(16) \times E_8$, $SO(24) \times SO(8)$, $SU(16) \times U(1)$
4			$E_7^2 \times SU(2)^2$

Shatter = # of gauge group factors

- Can be used as an organizing principle for the heterotic landscape
- Not every possible shatter present
- Different levels of SUSY have different possible shatter levels

Lessons from D = 10

- Only some gauge groups realizable
- Orbifold techniques utilized for this study will not find every model
- Correlations can exist between quantities which are formally independent in Quantum Field Theory (e.g. gauge group and supersymmetry)

Now let's go to D = 4!

Quick Introduction to D=4 Heterotic Strings

- Many more than nine distinct models
- Maximal SUSY is N=4
- Rank of gauge group is ≤ 22

Main Characteristics of this Study

- Perturbative Heterotic Strings (main area for string phenomenology in 80's and 90's)
- Millions of models randomly generated and analyzed by computer, all satisfying worldsheet self-consistency constraints
- Models with all of the levels of space-time SUSY realizable in D=4 N=0,1,2,4
- Uses Free Fermionic Construction partially overlaps with Narain bosonic lattice compactifications and orbifolds with arbitrary Wilson Lines

all gauge groups rank 16+6=22

The Free Fermionic Construction Method (very quickly)

- String is taken to be two CFTs (left-movers are conformal, right-movers are super-conformal)
- CFTs are made of tensor products of free non-interacting complex fermionic fields
- Create different models by changing the boundary conditions of the fields around the worldsheet torus while also changing the phase for the spin-structure's contribution to the string partition function. (Phases are +/- signs for GSOs)

Kawai et al, NPB **288**, 1 (1987)
Antoniadis et al, NPB **289**, 87 (1987)

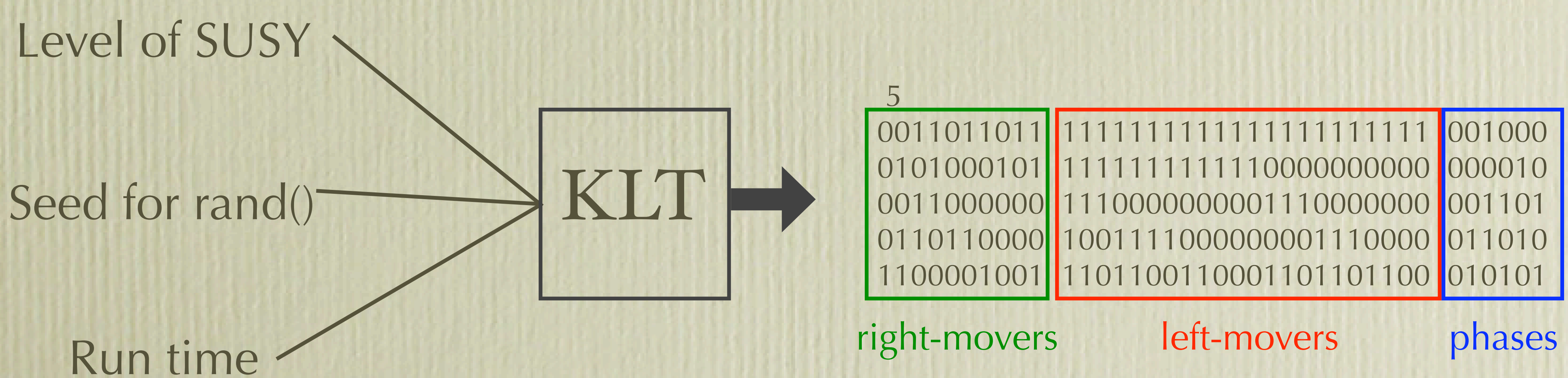
Advantages of Free Fermionic Method

- Models which are geometrically complex may be realized relatively easily
- Can get the full spectrum of the string model
- Can be put on the computer easily

D. Sénéchal, **PRD** 39, 3717 (1989).

Code I: Generating Models

- Give desired level of SUSY, seed for rand, and run-time and KLT generates self-consistent sets of vectors which correspond to models
- For vectors, 0(1) = (anti-)periodic



Code II: Analyzing Models

Steps to analyze any particle in spectrum:

- Determine all possible string excitations
- Verify that excitation is level-matched
- Verify that excitation satisfies GSO constraints

Steps to classify model:

determine N → determine gauge group, G → determine spectrum

find gravitinos → identify & analyze gauge bosons → all other states classified by “charges” under G and grouped into suitable multiplets

Sample Output for each model:

```
Supersymmetry N = 0

57 gauge bosons in SU(4) x SU(2)^14 x U(1)^5

34 Fermions irreps:
24    010 1 1 0 0 0 0 0 0 0 0 0 0 0 0   0 0 0 0 0  r
24    010 0 0 0 0 1 1 0 0 0 0 0 0 0 0   0 0 0 0 0  r
24    010 0 0 0 0 0 0 0 0 1 1 0 0 0 0   0 0 0 0 0  r
24    010 0 0 0 0 0 0 0 0 0 0 1 1 0 0   0 0 0 0 0  r
16    000 1 1 1 1 0 0 0 0 0 0 0 0 0 0   0 0 0 0 0  r
16    000 1 1 0 0 0 0 0 0 0 0 0 0 1 1   0 0 0 0 0  r
16    000 0 0 1 1 1 1 0 0 0 0 0 0 0 0   0 0 0 0 0  r
16    000 0 0 1 1 0 0 0 0 1 1 0 0 0 0   0 0 0 0 0  r
...
4     000 1 1 0 0 0 0 0 0 0 0 0 0 0 0   0 0 0 0 2  c
4     000 0 0 0 0 1 1 0 0 0 0 0 0 0 0   0 0 0 0 2  c
4     000 0 0 0 0 0 0 1 1 0 0 0 0 0 0   0 2 0 0 0  c
4     000 0 0 0 0 0 0 1 1 0 0 0 0 0 0   0 0 0 2 0  c
...
35 Scalar irreps:
24    010 0 0 1 1 0 0 0 0 0 0 0 0 0 0   0 0 0 0 0
...
```

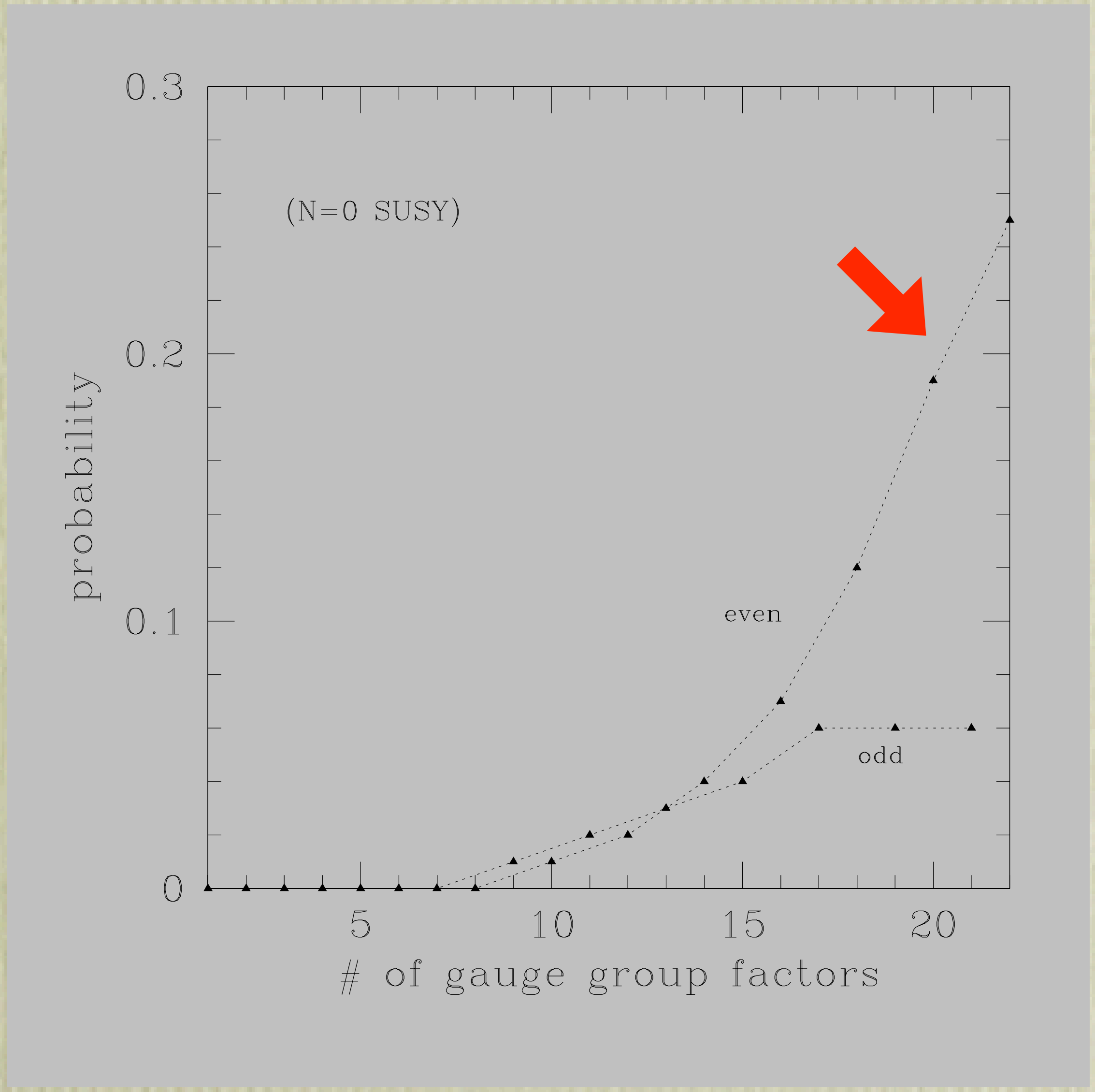
How many models were analyzed?

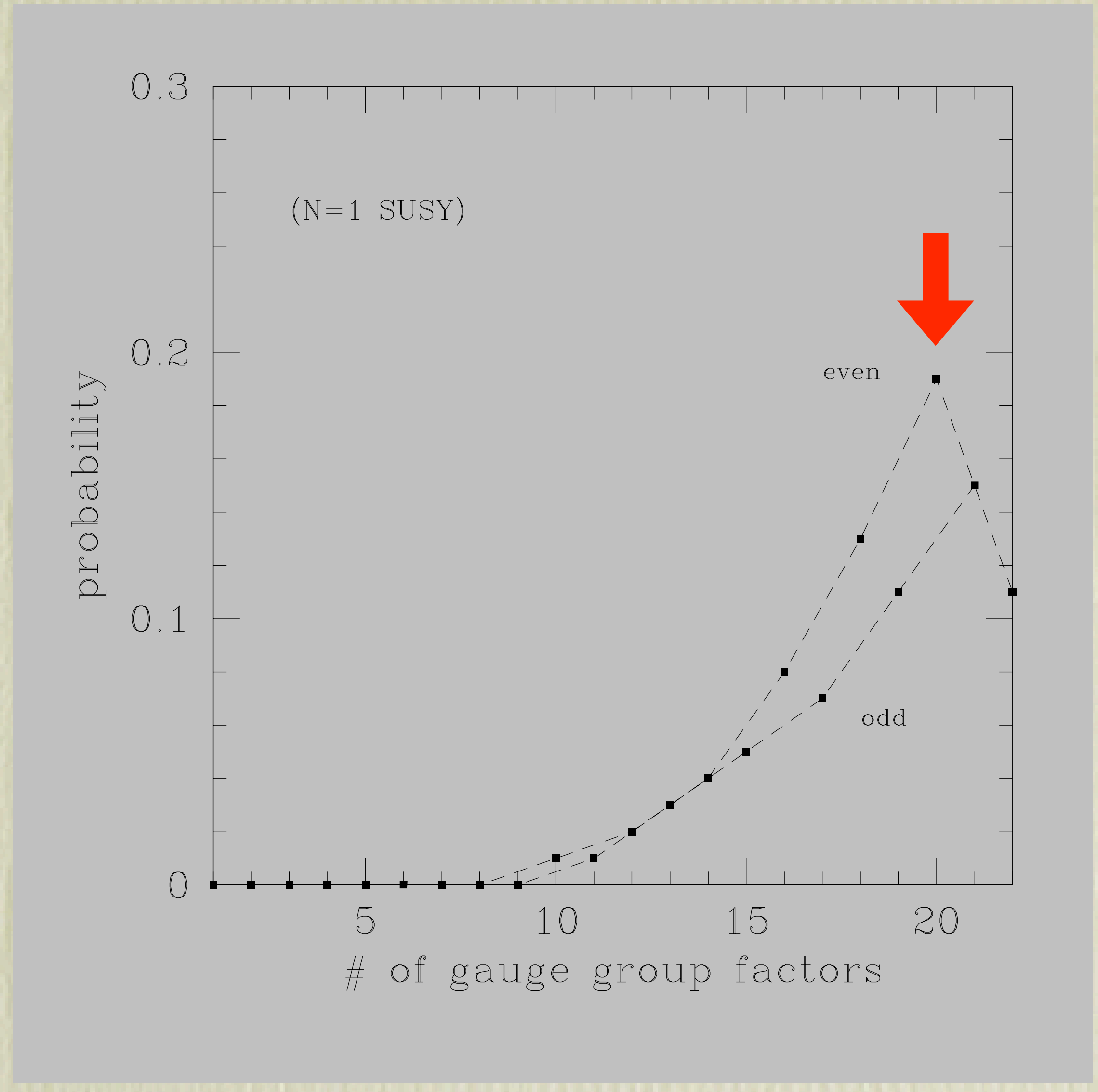
Class of model	Number of models	Attempts/Model
N=0	1.6×10^6	2.76
N=1	1.25×10^6	3.40
N=2	0.5×10^6	28.15
N=4	900	420.70

Shatter in D = 4

- Only one distinct gauge group with a shatter of one: SO(44) (just like SO(32) in D = 10)
- Lots of distinct gauge groups with a shatter of two (but they all consist of SO(44-n) x SO(n))
- Highest level of shatter is 22 and gauge groups at that level are $U(1)^{22-n} \times SU(2)^n$

What level of shatter do we expect?

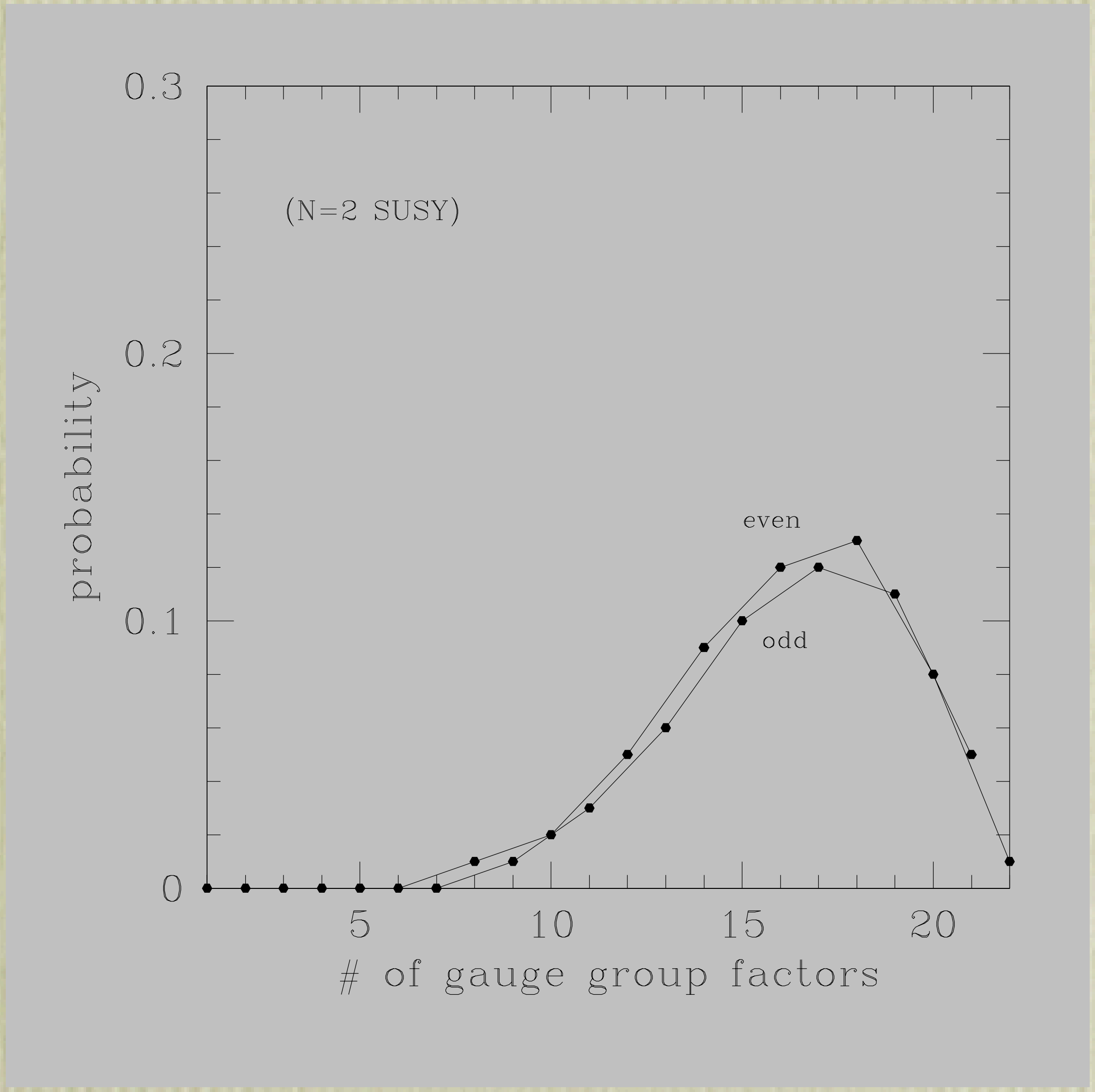


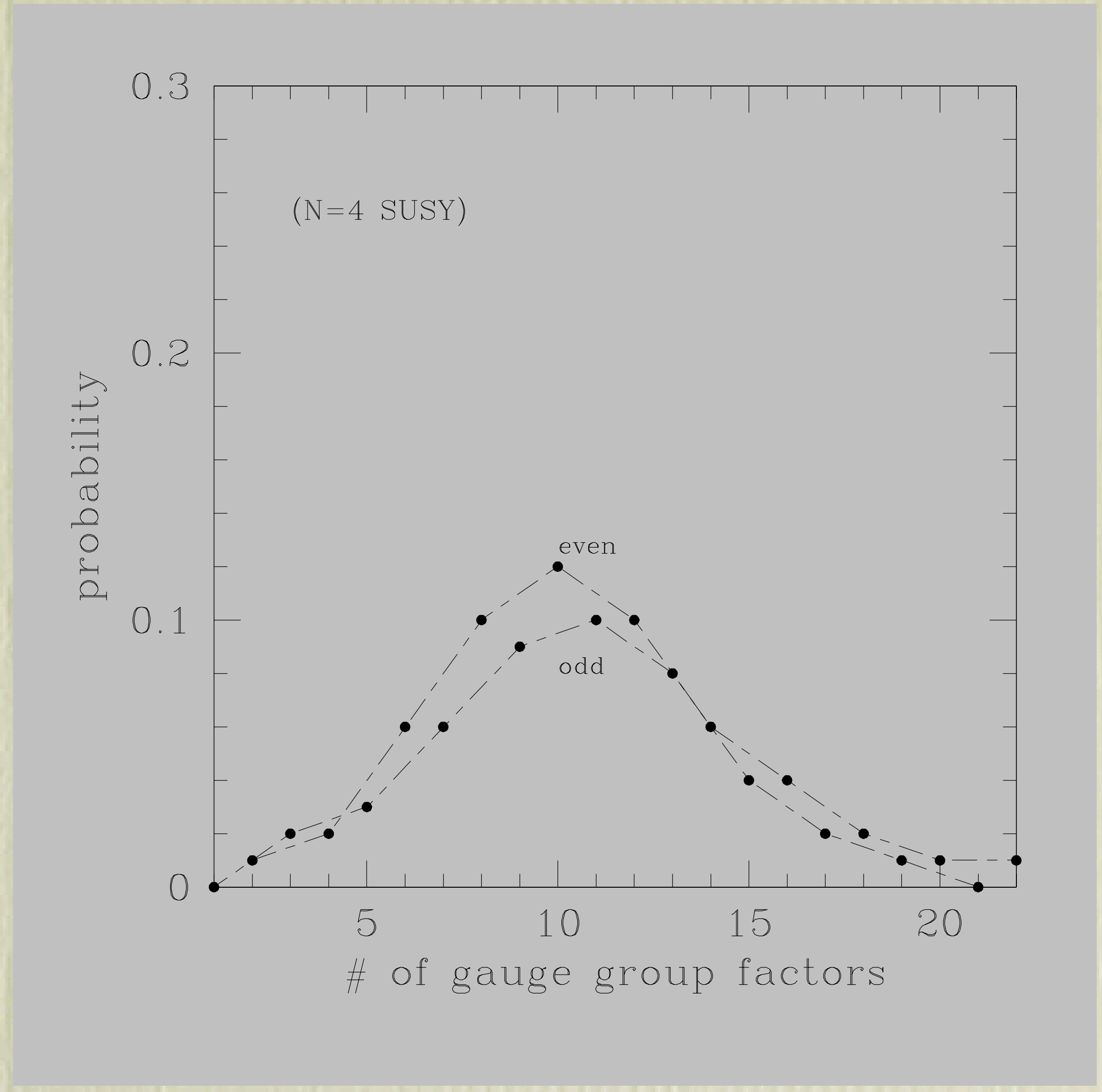


even numbers of gauge factors dominate

large numbers of factors dominate

With more SUSY...

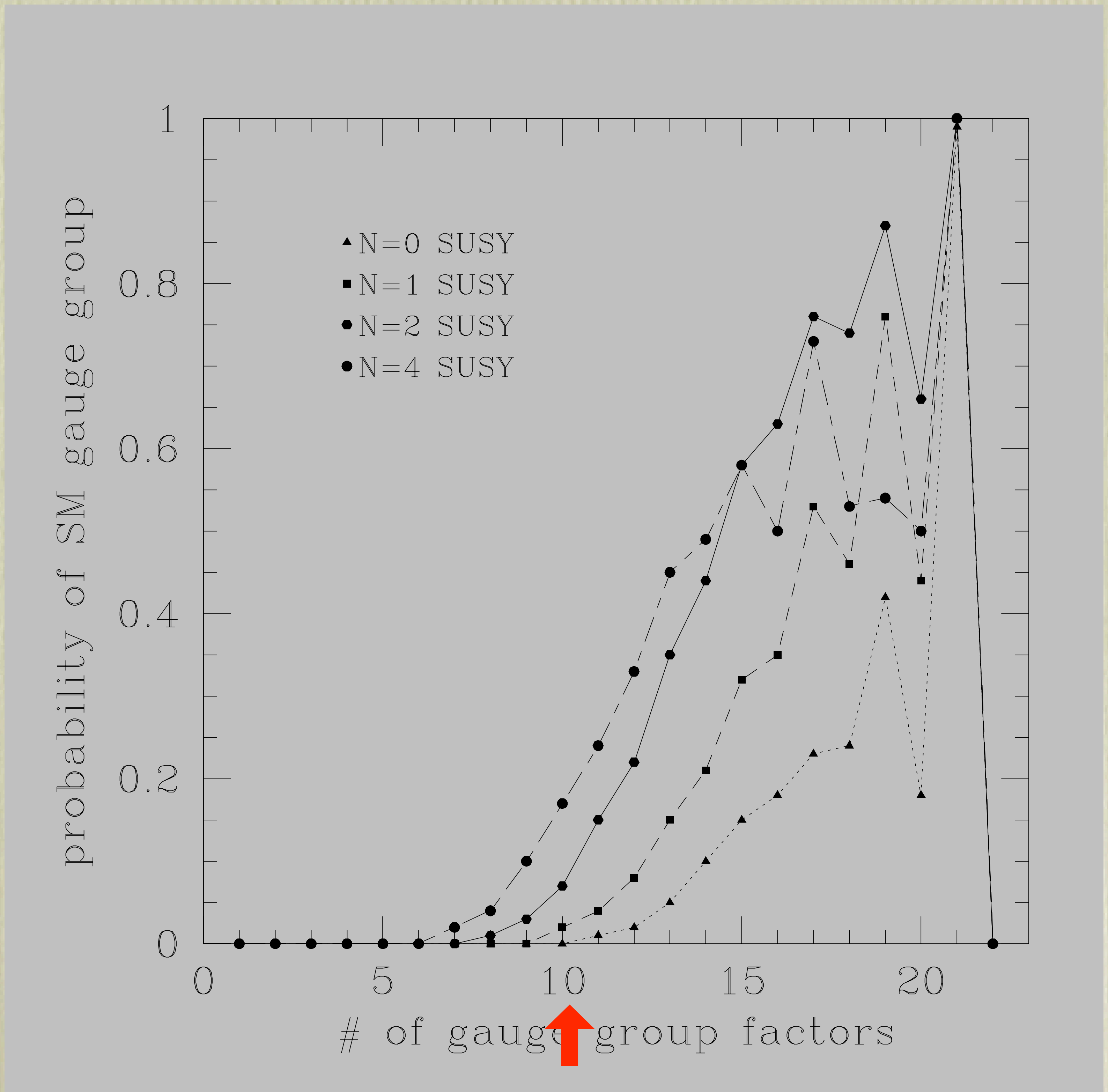




As SUSY increases, even/odd difference disappears!

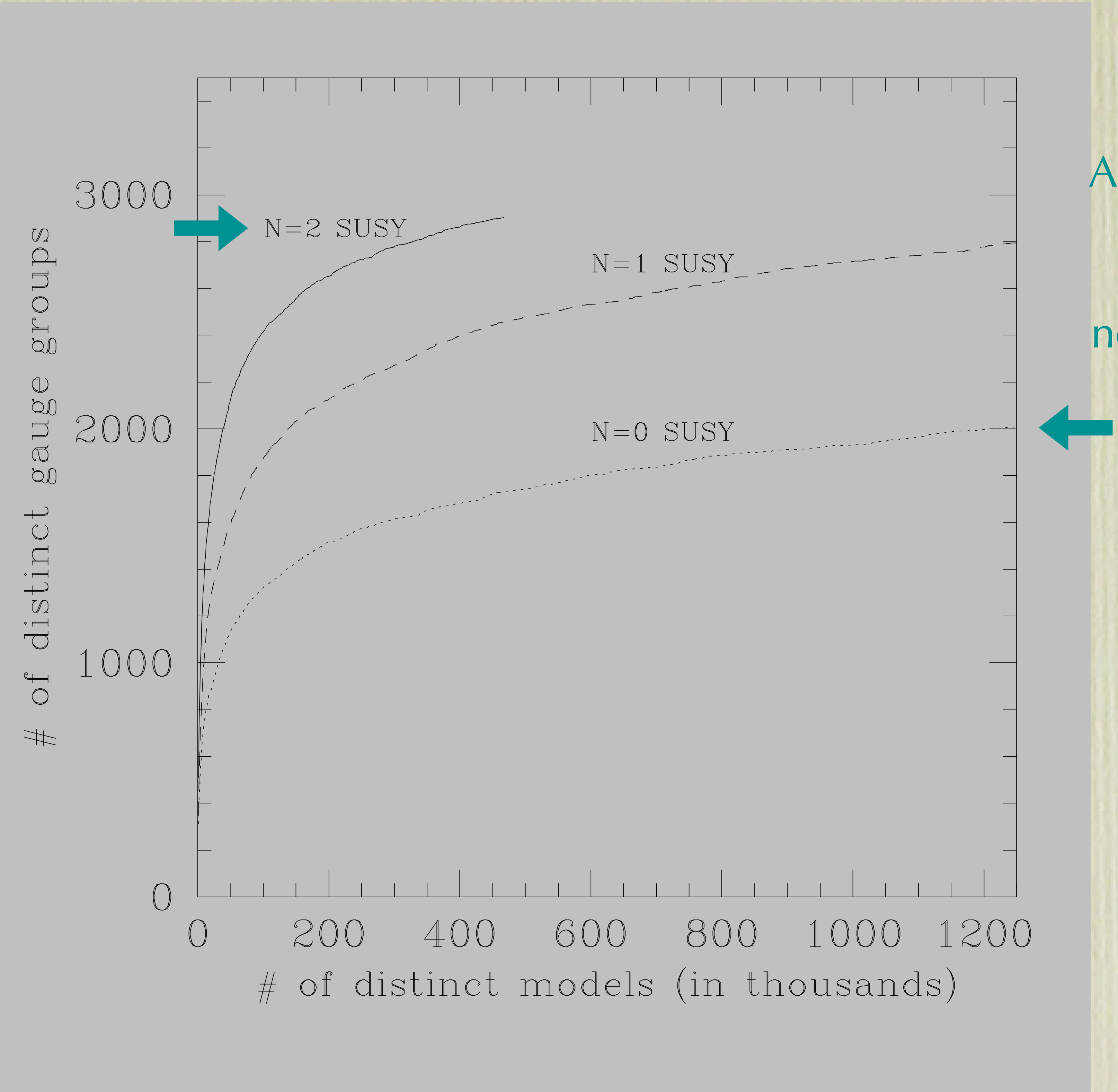
peak probability shifts towards smaller numbers of factors

When does the Standard Model gauge group appear?



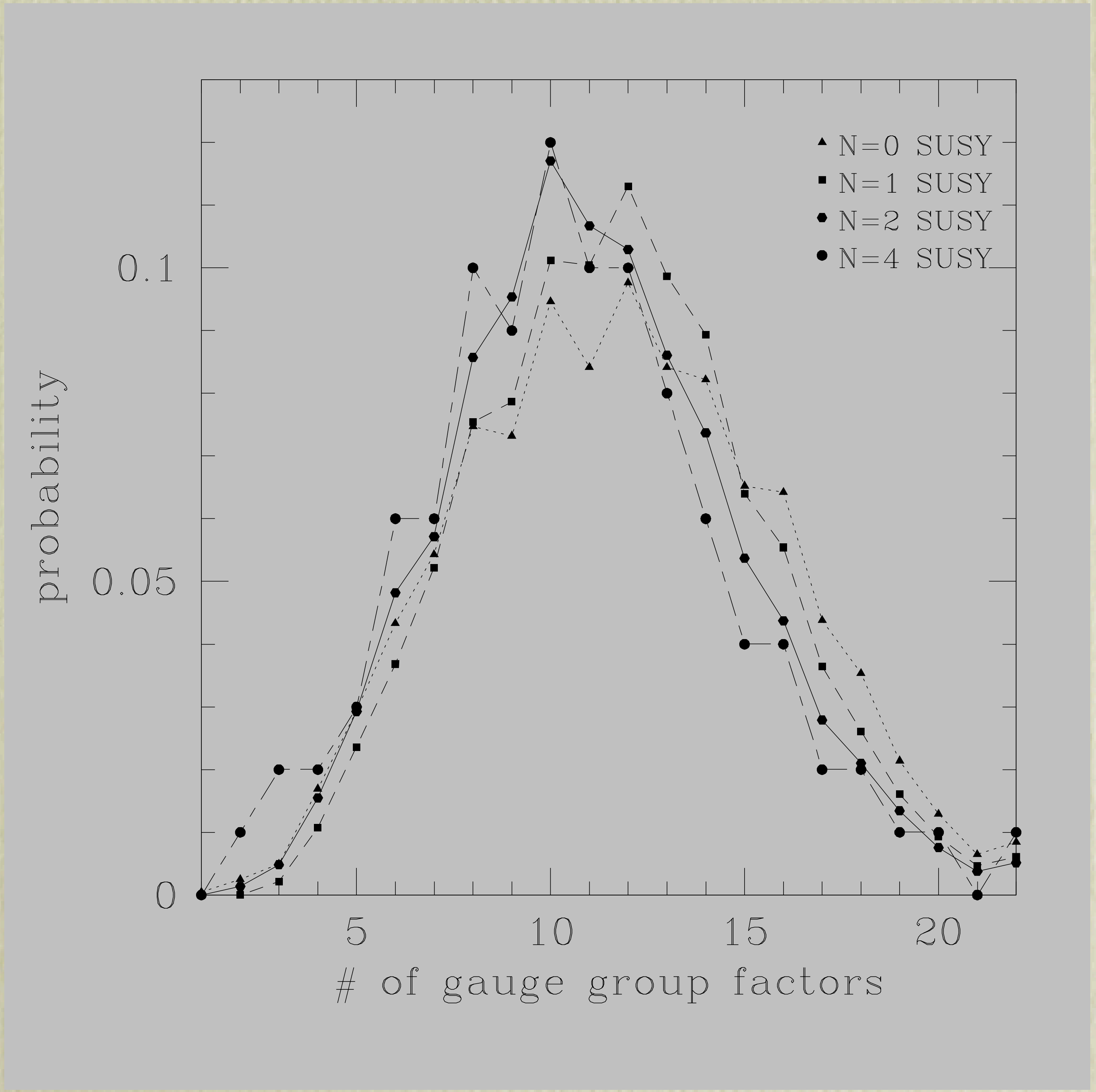
No Standard Model gauge group until a large number of gauge group factors

How often do we obtain unique gauge groups?



As SUSY increases, new models are forced to have new gauge groups.

How likely are different gauge groups?



For all levels of SUSY, same probability of production.

How likely are different gauge group factors?

smaller groups much more common

Standard Model limited by SU(3)

Group:	$N=0$	$N=1$	$N=2$	$N=4$
U_1	98.79	99.94	99.72	89.19
SU_2	97.72	97.44	96.57	70.17
SU_3	19.58	47.84	60.42	34.13
SU_4	53.61	51.04	62.76	42.94
SU_5	2.79	7.36	20.29	23.72
$SU_{>5}$	4.03	6.60	18.80	53.65
SO_8	19.83	13.75	19.73	19.82
SO_{10}	6.19	4.83	7.57	15.42
$SO_{>10}$	3.75	2.69	4.72	19.92
$E_{6,7,8}$	0.14	0.27	1.02	16.12
Standard Model: $3 \times 2 \times 1$	18.92	46.6	58.65	23.92
Pati-Salam: $4 \times 2 \times 2$	51.04	47.03	55.52	19.02

A brief interlude, Λ and the Heterotic Landscape

- Data comes from a different data set, **hep-th/0602286**
- Non-zero value indicates instability of vacuum beyond tree level but,
- Simplest one-loop amplitude for these models (many other amplitudes related through derivatives)

$$\Lambda \equiv \int_{\mathcal{F}} \frac{d^2\tau}{(\mathrm{Im}\,\tau)^2}\, Z(\tau)$$

$$Z(\tau) \equiv \mathrm{Tr}\,(-1)^F\, \overline{q}^{H_R}\, q^{H_L}$$

$$\mathcal{F} \equiv \{\tau : |\mathrm{Re}\,\tau| \leq \frac{1}{2}, \mathrm{Im}\,\tau > 0, |\tau| \geq 1\}$$

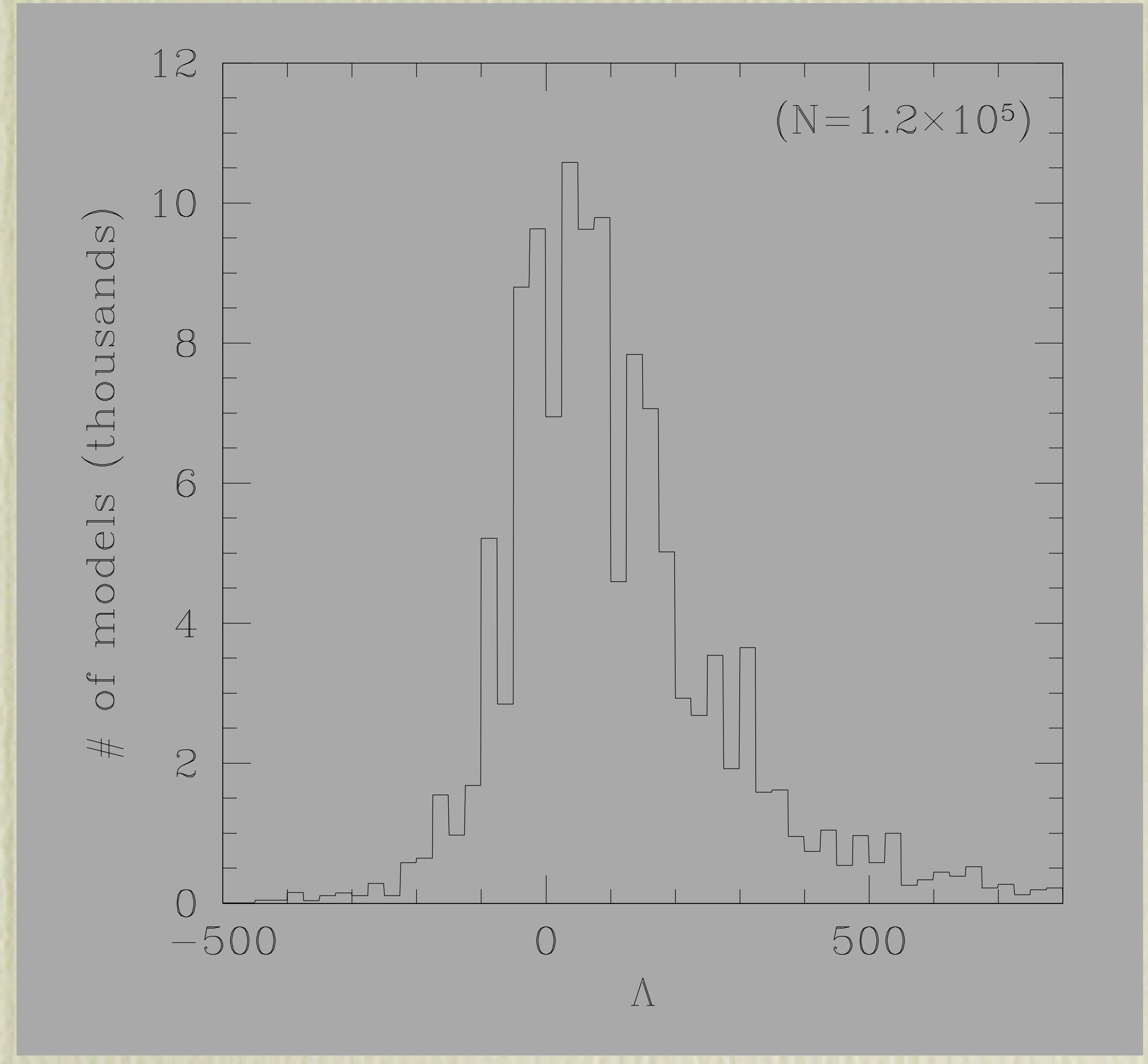
$$\Lambda \sim -C.\ C.$$

123,573
tachyon-free models
with N=0 SUSY

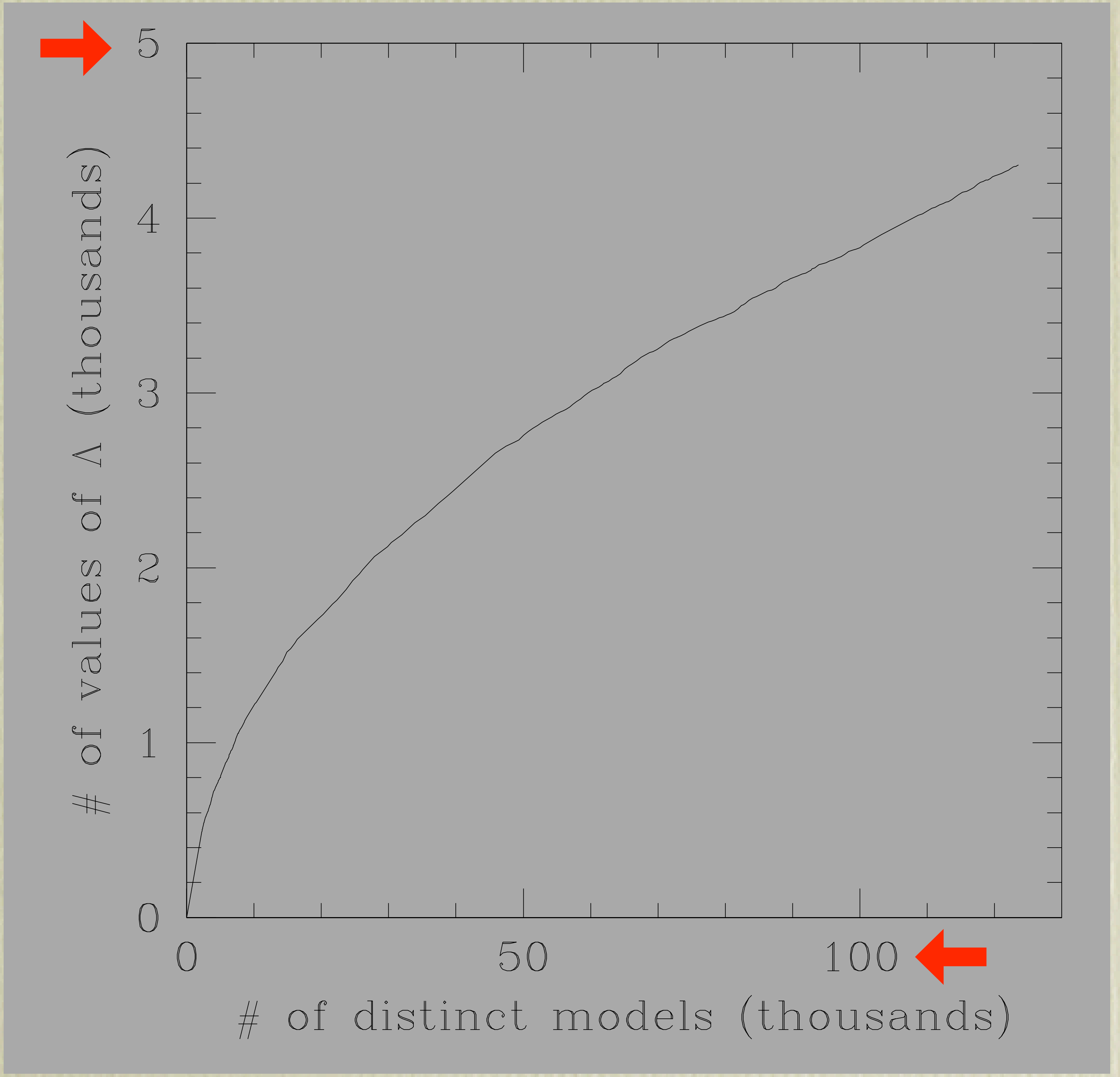
Dienes, K. PRD:73 106010, 2006.

How are the Λ's distributed?

- 73% of models have $\Lambda > 0$ (AdS space)
- Many different models with completely different gauge groups and particle contents nevertheless have the exact same Λ

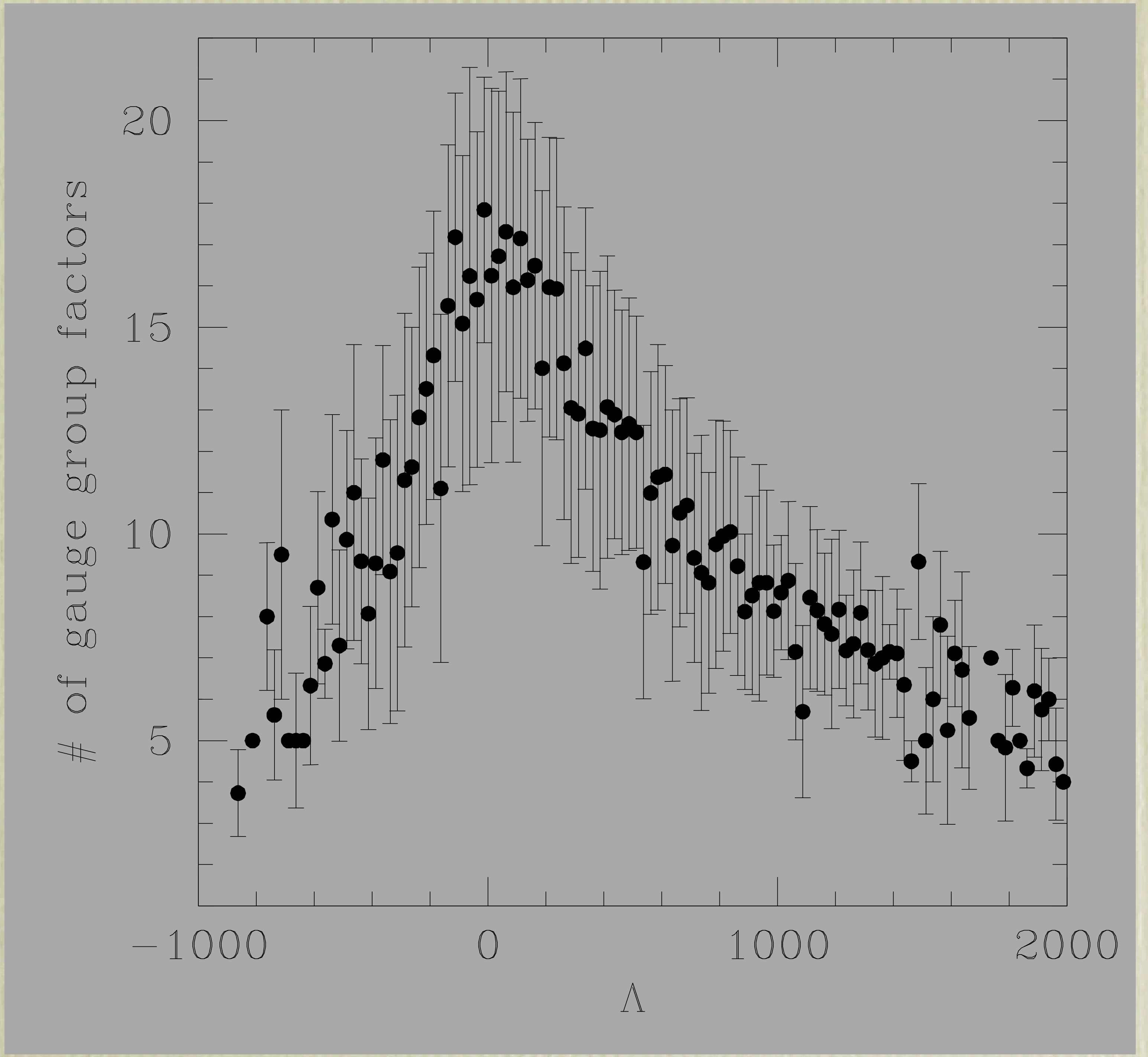


How many different values of Λ were found?



High degree of redundancy in Λ, thus N models does **not** imply N different values of Λ, consequences for Bousso-Polchinski?

How are shatter and Λ correlated?



Highly shattered groups have smaller Λ

And the list goes on...

- numbers of quark generations
- numbers of lepton generations
- Chirality
- Hypercharge normalizations
- Yukawa couplings
- etc.

However, there's one BIG issue:

Some of these correlations are ***NOT STABLE***.
MANY of these correlations will "*float*", i.e., evolve as the sample size increases!

WHY????

Simply put, as we explore more of the model space, it gets harder to find new models. Therefore, "rare" properties MUST appear more often in the process of obtaining additional **distinct** models!

So what can we do?

The Problem of “Float”

- Generic to any landscape study where full exploration of model space impossible/impractical
- Need to find method to overcome “float” so that true continuum limits may be found

This issue has never been addressed before in the literature, but plays a huge role in obtaining meaningful statistical results.

Need to re-examine the entire process of random model generation.

Start from basic probability analysis

General Features of Random Searches

- Probability to find new model proportional to amount of model space already explored
- Most model generation methods have biases which favor certain space-time properties over others

$$P_{\mathrm{new}} = 1 - \frac{x}{N}$$

x found models

N total models

bias

$$P_j = \frac{p_j N_j}{\Omega_{\mathrm{prob}}}$$

$$\Omega_{\mathrm{prob}} = \sum_{i=1}^{n} p_i N_i$$

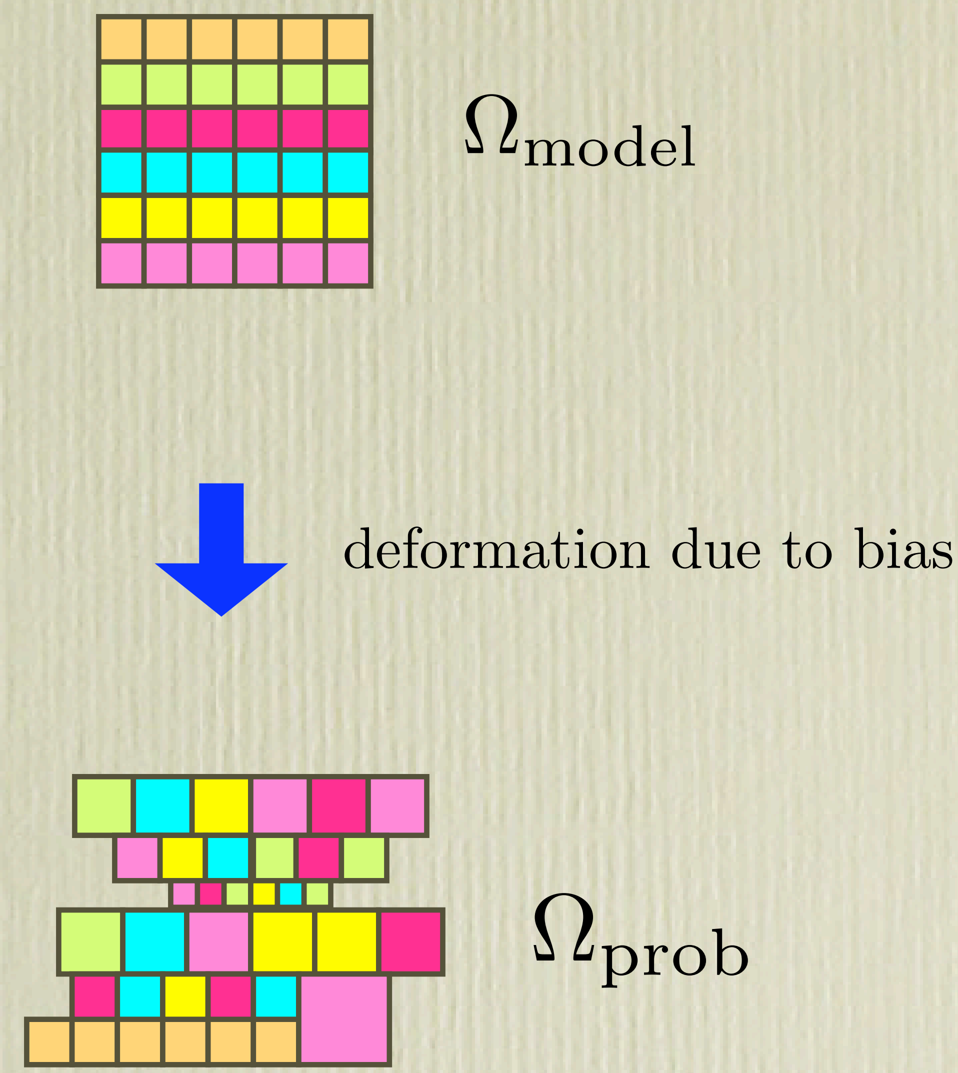
Thus,

$$P_{\mathrm{new}\ j} = \frac{p_j}{\Omega_{\mathrm{prob}}} (N_j - x_j)$$

assuming a finite number of models

Definition of spaces

- Ω_{model} is a space where every model occupies the same volume
- Ω_{prob} is a space where every model occupies a volume proportional to the probability of production



analogous to US map rescaled by population

Comparison of Spaces

Ω_{model}	Ω_{prob}
every model occupies same volume	models generally occupy different volumes
volume relations determine correlations	volume relations determine biases

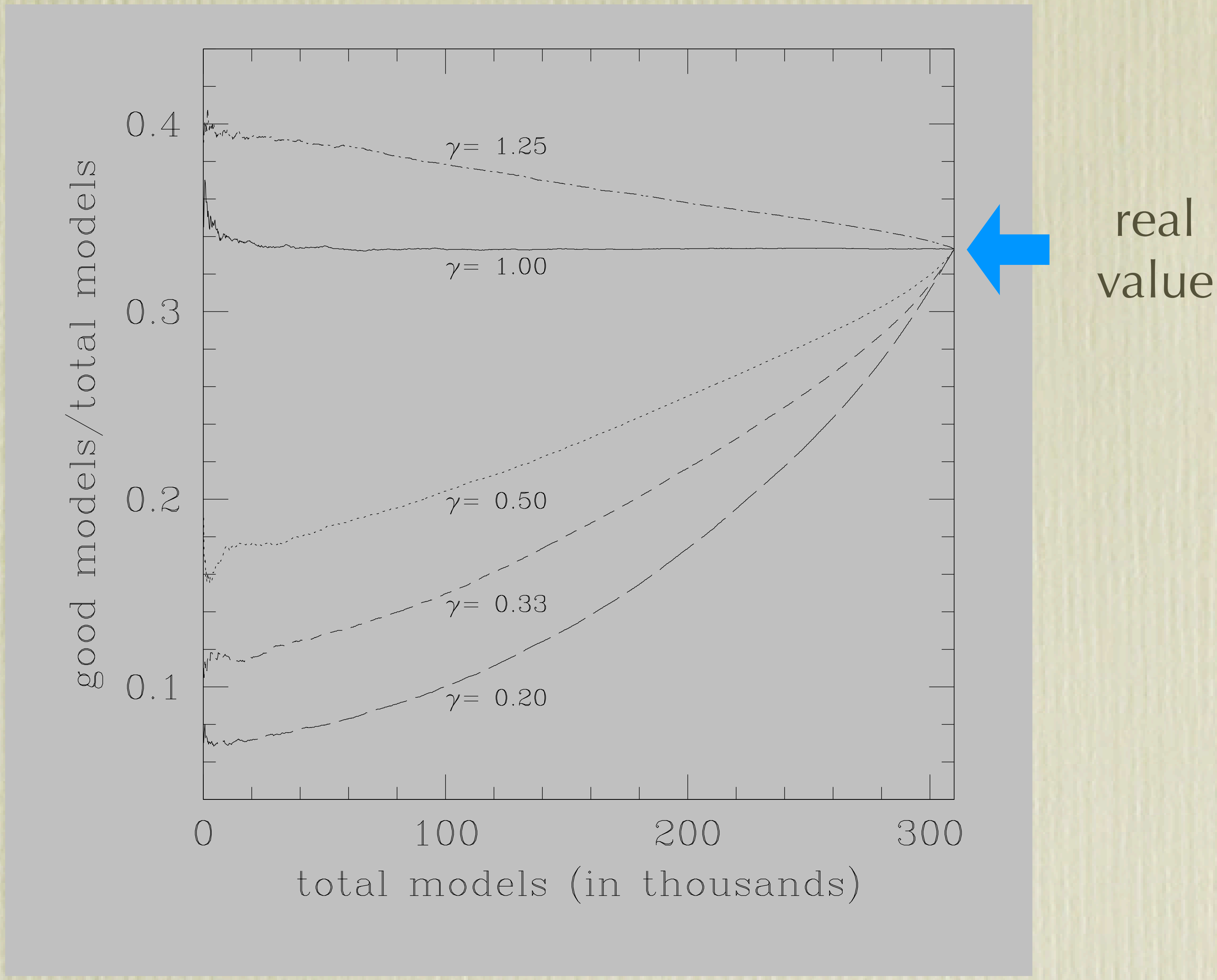
Need to simulate model production

- Label each model by an integer
- “Randomly” choose an integer to simulate model production
- Look at correlations among groups of integers (i.e. all integers which are divisible by 3 are “good”)

Illustration of Bias Danger

- Apparent correlation between good model/ total models ***floats***, or changes as the model space is explored

- Very hard to distinguish between physical correlation and biases inherent in model generation method



$$\gamma = \frac{p_{\text{good}}}{p_{\text{bad}}} \qquad \text{Recall,} \qquad P_j = \frac{p_j N_j}{\Omega_{\text{prob}}}$$

So now what?

- Problem 1: Model Spaces sizes are unknown and are possibly different
- Problem 2: Need to eliminate bias
- Solution 1: Find way to compare differently sized model spaces
- Solution 2: Find way to restore equal probability of production to each model

Solution to Problem 1

Recall: $P_{\text{new}} = 1 - \boxed{\frac{x}{N}}$

↓

measure of how explored space is → number of tries to get new model ***increases*** as model space explored

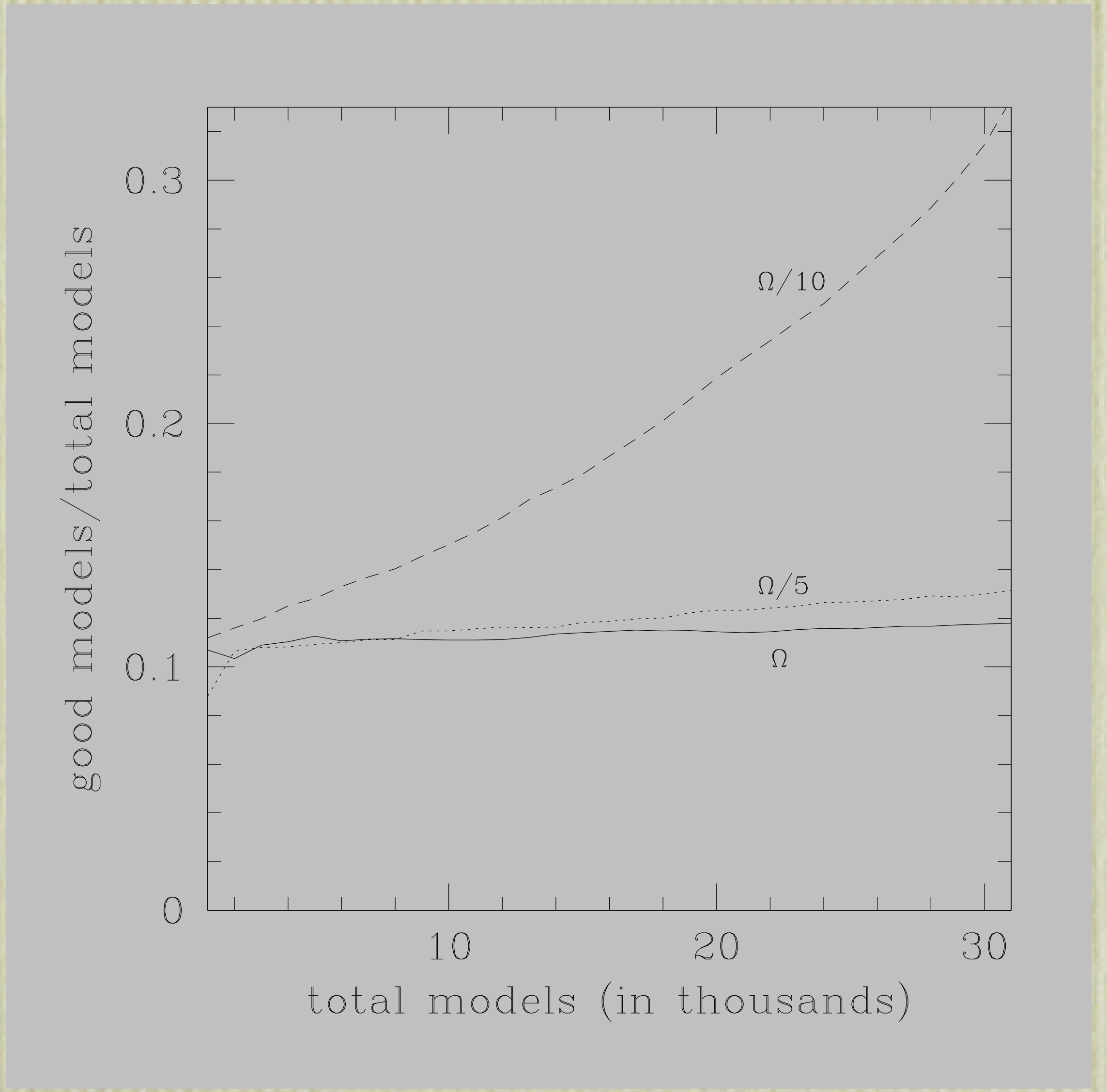
↓

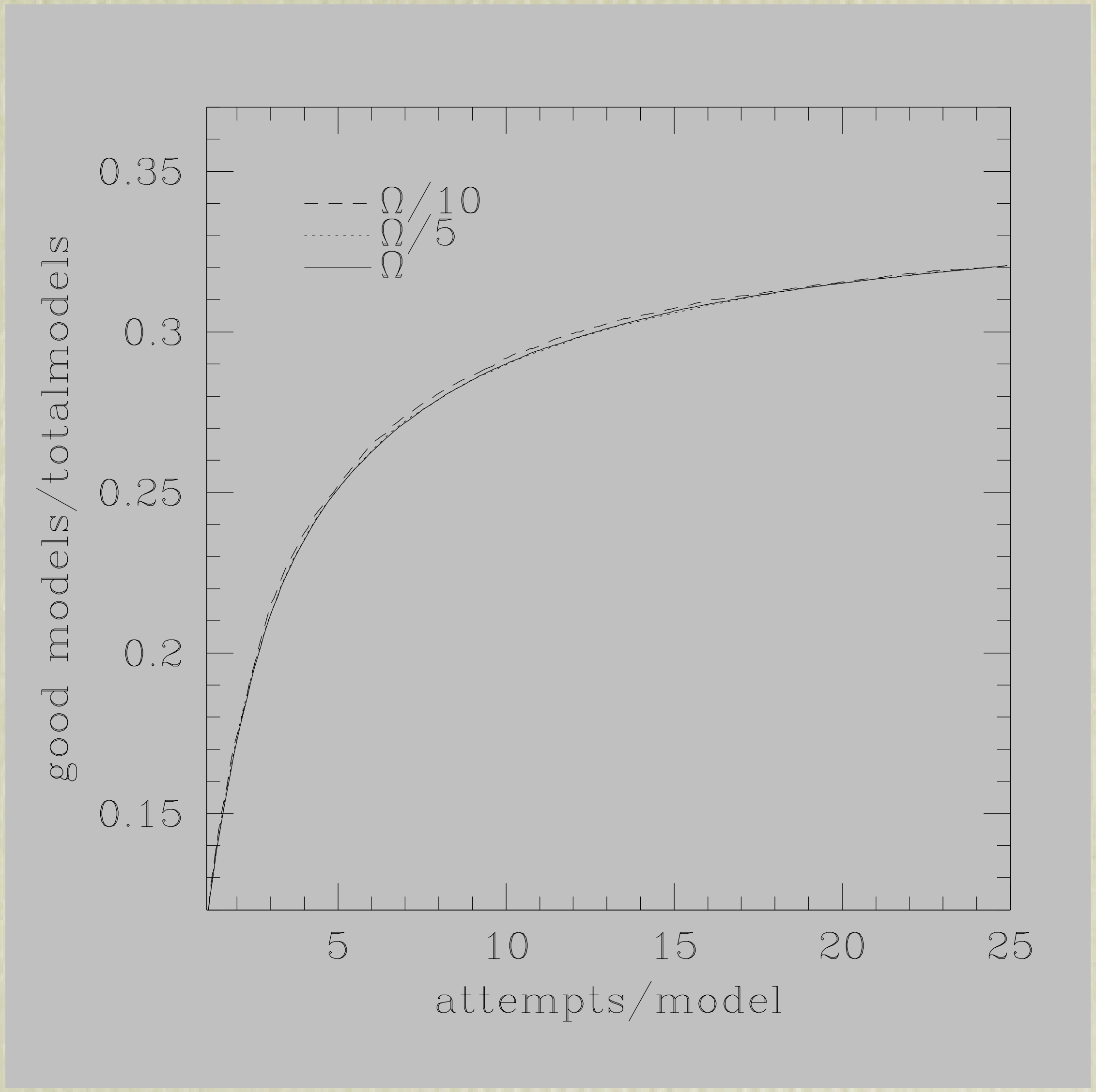
$$\frac{N_{\text{attempts}}}{N_{\text{models}}}$$ ← use this to compare model spaces at equal levels of exploration

||

new way to measure sample sizes

Illustration of solution 1

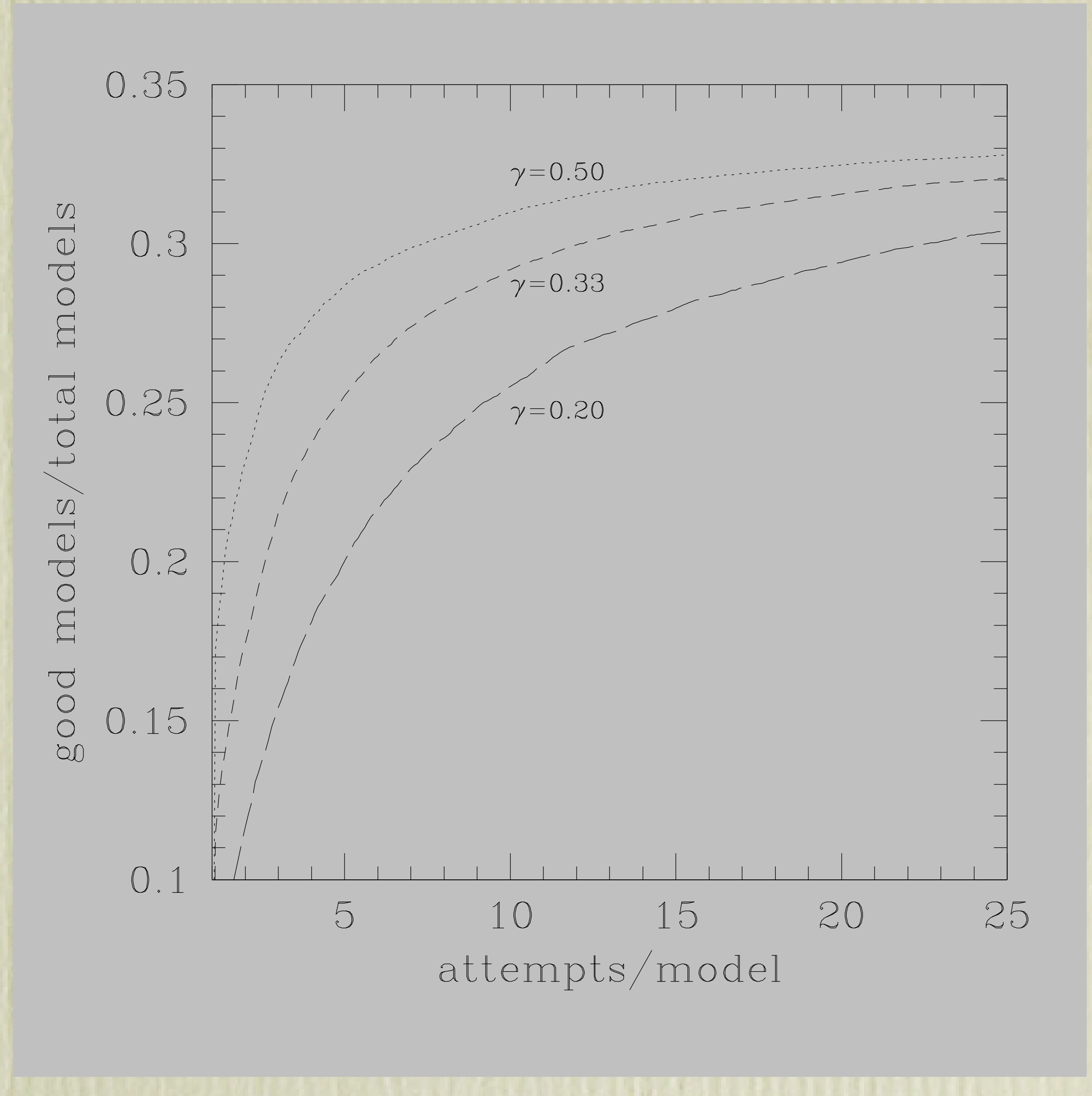




Three different model space sizes, all with the same bias for good models vs bad models

Does comparison method overcome the bias issue?

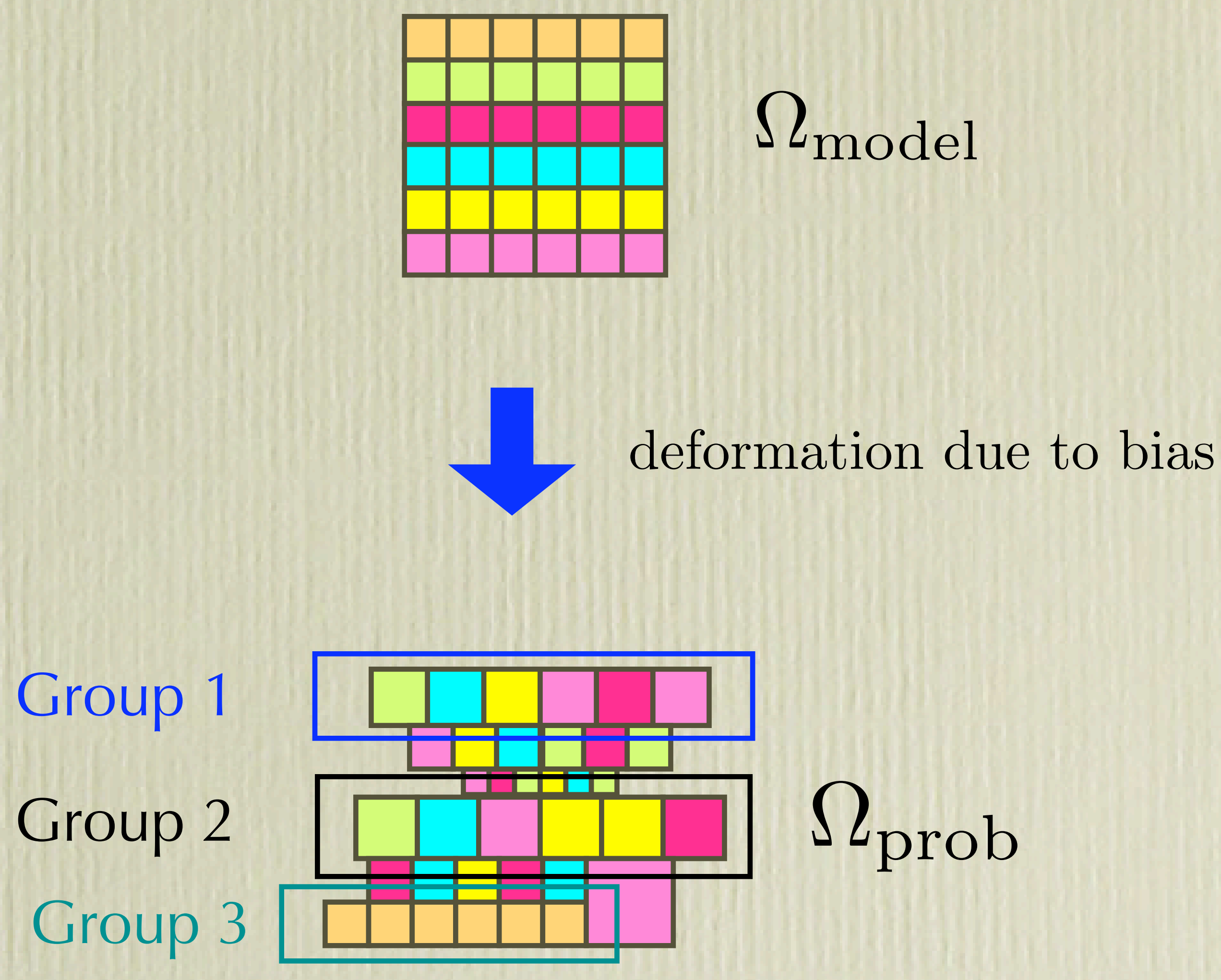
- Illustrates bias quite well
- Doesn't seem to be able to eliminate bias
- Can something else be found?



$$\gamma = \frac{p_{\rm good}}{p_{\rm bad}} \quad \text{Recall,} \quad P_j = \frac{p_j N_j}{\Omega_{\rm prob}}$$

Recall the Problem:

- All of the boxes are different sizes, thus some boxes are preferred!
- Can we restrict our attention to groups of all the same sized boxes?



New Method of Counting

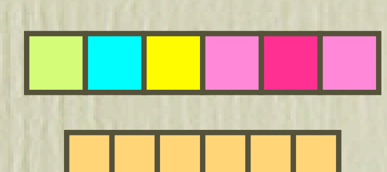
Within each group: $P_{\text{new}} = 1 - \frac{x}{N}$ Therefore, $\frac{N_{\text{attempts}}}{N_{\text{models}}}$

tells how explored each group is, at any point.

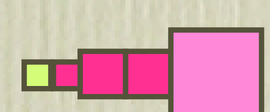
Thus,

$$\left.\frac{x_i}{x_j}\right|_{\left(\frac{N_{\text{attempts i}}}{N_{\text{models i}}} = \frac{N_{\text{attempts j}}}{N_{\text{models j}}}\right)} = \frac{N_i}{N_j}$$

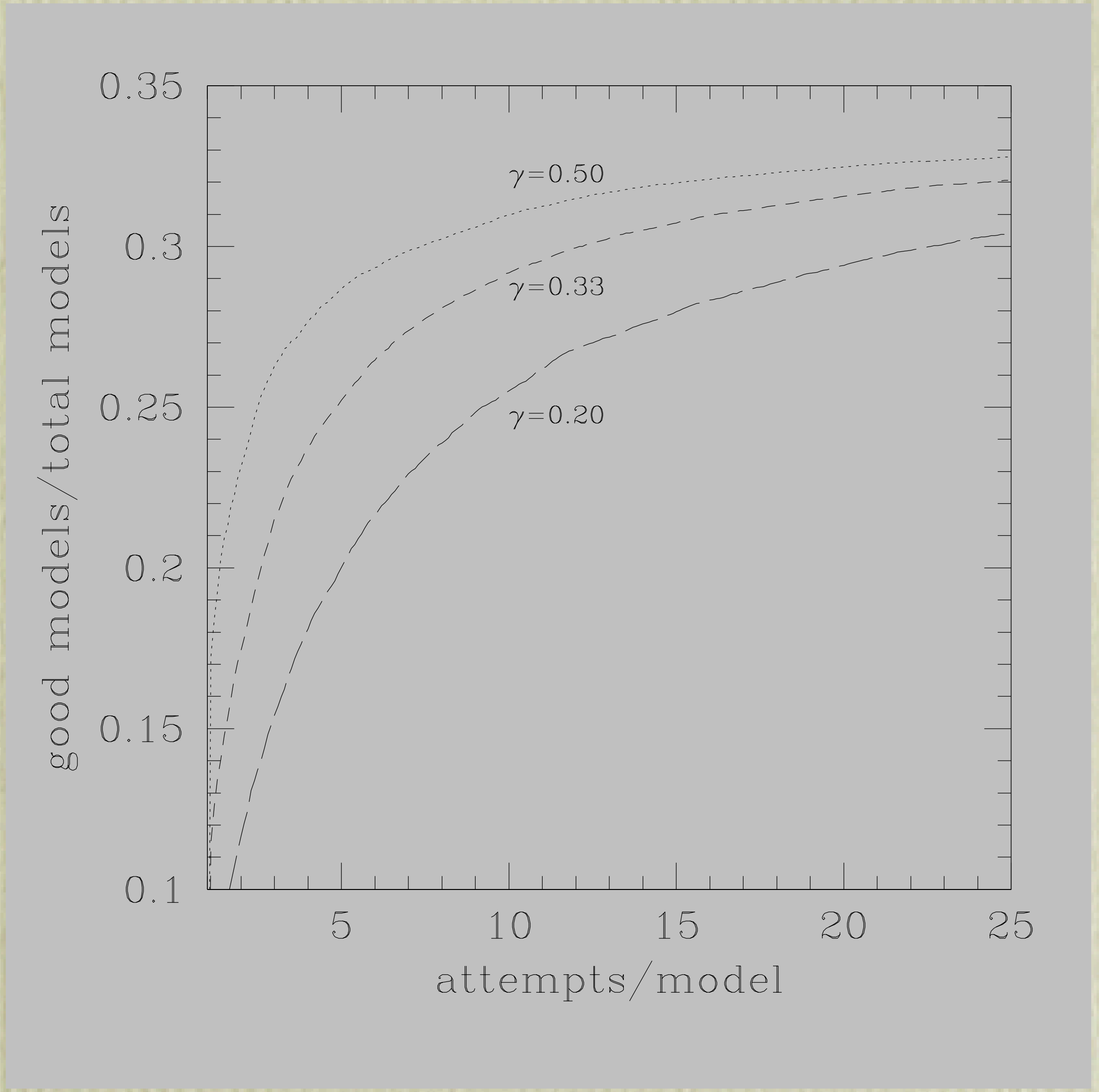
So long as groups are:



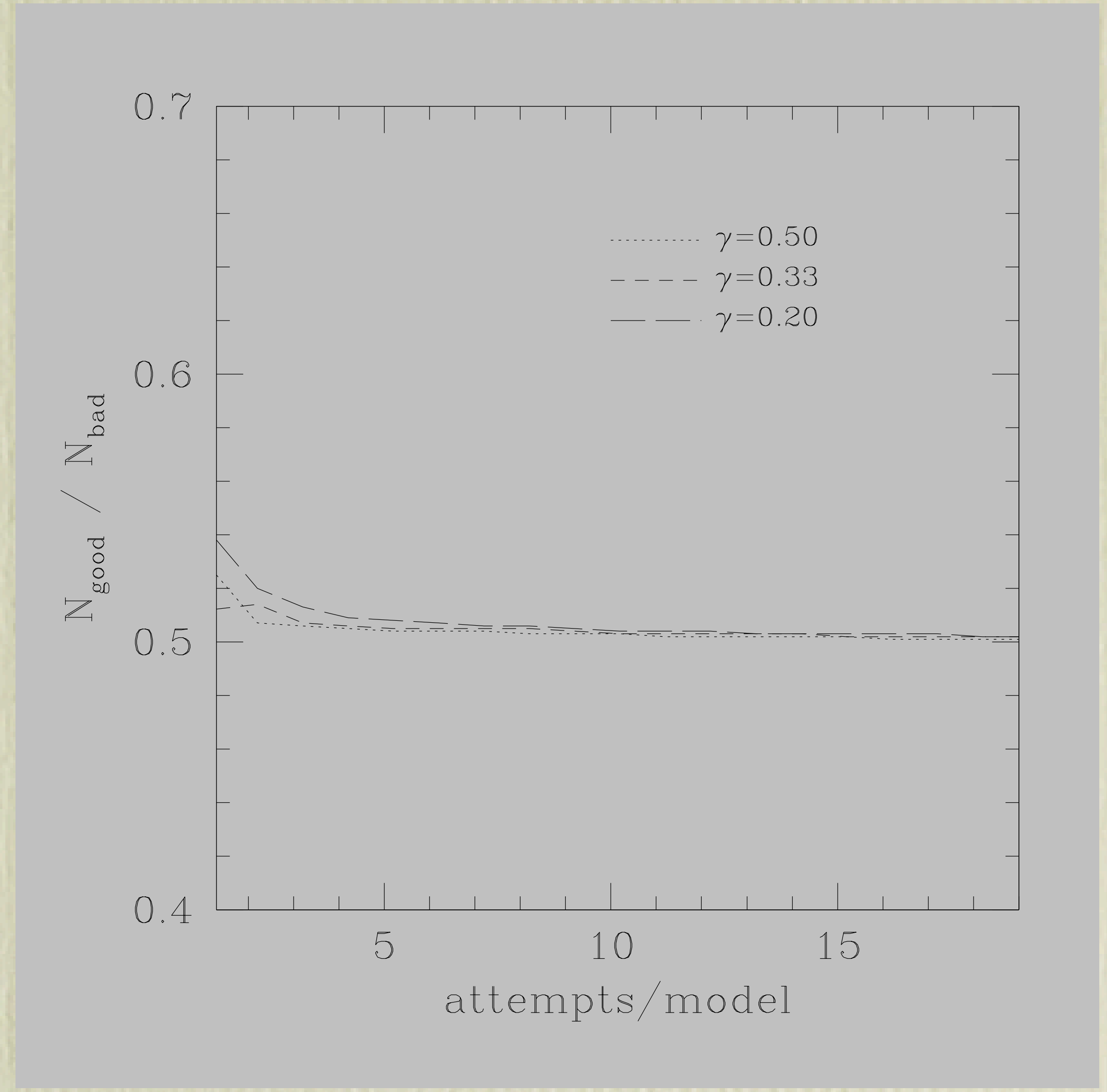
Not:



Quick Confirmation



the old method



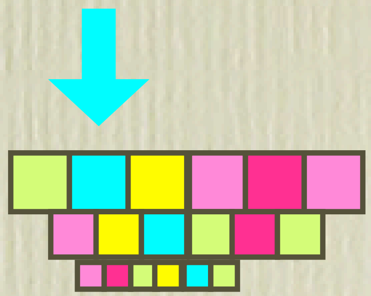
the new method

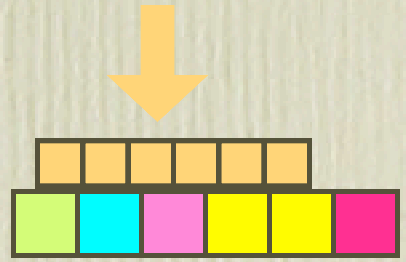
N_{good}/N_{total} = 1/3 as expected!

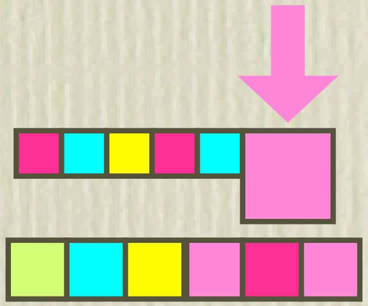
Limitations of new method

Need to divide up groups based on space-time properties

Case 1	property of interest spread uniformly amongst different biases	correlation stays constant even in old analysis method
Case 2	May only get one bias population with the property of interest	bias of generation method overcome using new method of analysis
Case 3	property of interest somewhat randomly spread amongst different bias populations	correlation will float even using new analysis method







Example 1: Just how common is SUSY anyway?

Class of Model	Percentage of Model Space
N=0	71.09%
N=1	28.36%
N=2	0.54%
N=4	0.0047%

37.86% tachyonic

62.14% tachyon-free

- Less than 30% of model space has space-time SUSY
- But, only 1/4 of model space has tachyons at tree level
- Very rare to find more than N=1 SUSY

ALL of these results are **stable**.

Example 2: Number of Unique Gauge Groups

Class of Model	Number of Unique Gauge Groups
N=0	107.00
N=1	42.01
N=2	1.00
N=4	.0087

$$\text{Table Entries}: \frac{\#\text{ of unique gauge groups in this class of models}}{\#\text{ of unique gauge groups for N = 2 models}}$$

Example 3: Effects of Float can be important!

$$\text{Table Entries}: \frac{\#\text{ of models with gauge group containing given factor}}{\#\text{ of models with gauge group containing SU}_3}$$

Group:	$N=0$	$N=1$	$N=2$	$N=4$
SU_3	1	1	1	1
SU_4	2.74	1.07	1.04	1.26
SU_5	0.14	0.15	0.34	0.69
$SU_{>5}$	0.21	0.14	0.31	1.57
SO_8	1.01	0.29	0.33	0.58
SO_{10}	0.32	0.10	0.13	0.45
$SO_{>10}$	0.19	0.06	0.08	0.58
$E_{6,7,8}$	0.01	0.01	0.02	0.47
$3\times2\times1$	0.97	0.97	0.97	0.70
$4\times2\times2$	2.61	0.98	0.92	0.56

Results from earlier sample

Group:	$N=0$	$N=1$	$N=2$	$N=4$
SU_3	1	1		1
SU_4	4.07	10.37		1.13
SU_5	2.01	15.25		0.70
$SU_{>5}$	51.1	72.1		1.36
SO_8	10.4	34.6	in progress	.50
SO_{10}	33.8	98.6		0.41
$SO_{>10}$	57.6	178.0		0.47
$E_{6,7,8}$	4.5	9.34		1.35
$3\times2\times1$	0.96	.978		.71
$4\times2\times2$	3.97	10.2		.51

After accounting for bias

Others float!

Some correlations stay the same

Final Gauge Group Populations

Group:	$N=0$	$N=1$	$N=2$	$N=4$
SU_3	1	1		1
SU_4	4.07	10.37		1.13
SU_5	2.01	15.25		0.70
$SU_{>5}$	51.1	72.1	in progress	1.36
SO_8	10.4	34.6		.50
SO_{10}	33.8	98.6		0.41
$SO_{>10}$	57.6	178.0		0.47
$E_{6,7,8}$	4.5	9.34		1.35
$3\times 2\times 1$	0.96	.978		.71
$4\times 2\times 2$	3.97	10.2		.51

Large groups most common

SM only limited by SU(3)

Pati-Salam, GUTs favored over SM for such strings.

Conclusions/Future Work

- Using probability analysis, random model generation biases can be overcome
- Refine understanding of gauge group probabilities in progress
- Look at massless particle spectrum to determine probability of realizing Standard Model in progress
- Use different search techniques to explore other model spaces

The End