

## Illustration of the randomness in synthetic NIPS data

1. The synthetic NIPS data is created by removing words with counts bigger than 3,000 or smaller than 1,000 and restaurants with less than 400 of the remaining words.
2.  $J=984, N=480$  (avg)
3. Fixed Gibbs Sampling parameter:  $\gamma = 5, \alpha = 1, \lambda = 1$
4. Run Gibbs Sampling for 10,000 iterations and the likelihood converges
5. In the topics shown below, only words that appear above average counts of the appearing words are listed.

Below, we will see that the randomness in the documents in the synthetic NIPS data can harm the performance of Gibbs Sampling.

The format is : ("topic index", "number of customers", "average likelihood per customer")

**Run I: "Denoised Restaurants":**

remove the words that appear only once or twice

First 10 topics in terms of  $\frac{Likelihood}{numberofcustomers}$ :

**topic 32:** (4978) -2.6894

cell cells firing spatial cortex complex properties  
activity inputs active simple connections simulation  
average relative responses

**topic 26:** (4613) -2.8383

feature features high size level search stage dimen-  
sional general found experiments algorithms complex  
large multiple bit simple maps

**topic 34:** (4019) -2.9041

node nodes tree graph decision procedure multi large

**topic 7:** (3896) -3.0594

field receptive fields center size local approximation  
position type presented theory small present simple  
dimensional structure term

**topic 48:** (3616) -3.1882

signal filter signals detection delay gaussian desired  
adaptive line ieee experiments fig prediction decision  
low optimal process work proc

**Run II: "Original Restaurants":**

do nothing

Topics that matched those on the left using

Hungarian Matching algorithm.

**topic 16:** (6452) -4.0124

cell cells direction complex firing properties spatial  
goal step active environment cortex connections  
activity rate analysis simulation location relative  
university brain determined experimental simple  
shows center similar left circuit measure level present  
press inputs presented specific significant higher  
references long proposed respect structure average  
found temporal position dependent form paper

**topic 3:** (8587) -4.3471

feature features map maps regions dimensional  
location large representation region small process  
high vectors present size found represent complex  
multiple parallel work search spatial dimension level  
general part represented line computer form hand  
higher center patterns position analysis real distance  
true represents type code vision find important layers  
simple difficult local examples note university good  
orientation required chosen mapping task

**topic 7:** (6164) -4.1237

node nodes tree level decision structure graph  
architecture machine procedure large binary theory  
multi rate size connected algorithms net called pages  
form increase paper adaptive represents final shows  
fact top research sample equal step means internal  
inputs continuous layers work long equivalent  
conference class random left applied made efficient  
small ieee previous elements artificial

**topic 10:** (5591) -4.2426

field receptive fields size local center large  
approximation structure dimensional gaussian  
connections type present individual average term due  
consists test small rule regions part sum similar  
effects standard references simple show general  
research total real learn connection presented high  
architecture independent scale ing fig result contrast  
response properties correlation

**topic 30:** (6630) -4.6624

signal filter signals optimal fig detection gain  
gaussian desired real samples nonlinear line rate  
estimation term ieee analysis adaptive form random  
response level theory speech parameter experiments  
present correlation delay high general chosen multiple  
proc independent solution make obtain design sample  
structure decision study equal considered defined  
maximum complex domain shows paper section  
presented similar result important vol stage work  
process prediction low conditions outputs continuous  
methods problems average required represent correct  
compared train terms change distributed resulting  
provide representation applied find series

**topic 33:** (4610) -3.2779

stimulus response stimuli responses visual patterns  
activity cortex presented theory left cortical show  
properties effect log type effects standard multiple  
current shows

**topic 43:** (3770) -3.2973

fig phase range shows patterns simulations correlation  
simulation complex behavior computer high param-  
eter form show research gain connected active left  
dimensional center multiple

**topic 50:** (3750) -3.316

motion direction visual speed component rate com-  
putation location stimuli estimate left spatial local  
research contrast computed points vision parallel  
random similar fig field trained global

**topic 46:** (7028) -3.3424

classification class classifier classifiers classes decision  
test rate patterns problems trained rates experiments  
table high accuracy regions study algorithms good  
gaussian stage basis maximum original train per-  
formed

**topic 23:** (4411) -3.3798

spike rate firing neuron train rates fig average tem-  
poral code real times constant inputs threshold eq  
dependent stimulus note

**topic 44:** (5732) -5.0545

visual response activity task human stimulus  
responses brain tasks patterns stimuli study  
experimental stage target activation effect standard  
computational related detection delay theory decision  
correct presented effects computation cortex  
conditions experiments science specific analysis  
modeling experiment rate level university threshold  
research behavior multiple determined individual  
difference average result active significant field source  
test tion computed times prior shows due provide top  
control press make observed made relative location  
gain present simulation procedure statistical contrast  
perform framework trained parameter important  
neuron signal signals representations

**topic 46:** (5209) -4.8286

phase correlation frequency patterns architecture fig  
high low range shows behavior connected complex  
center large due simulations computer form delay  
parameter show difference representation activity  
research threshold temporal analog solution feature  
work positive tion active applied region length defined  
important dynamic simulation simple correct process  
computation binary fact potential represents addition  
gaussian small continuous negative dimensional cross  
free generated determined problems standard factor  
result maximum abstract errors representations scale  
resulting sum effect make determine

**topic 26:** (7546) -4.5081

motion direction visual eye position stimulus location  
stimuli speed spatial simple activity response vision  
left range signals cortex motor map computed scheme  
responses component field similar signal temporal  
trajectory analysis computation objects  
representation computational rate object relative  
cells human layers generated equation parallel  
research initial contrast detection science gaussian  
inputs architecture correct local process sequences  
increase activation control step points basis

**topic 12:** (11062) -4.4545

classification classifier class classifiers classes decision  
patterns test rate feature problems regions trained  
high multi samples rates gaussian train maximum  
form probabilities complexity techniques large  
algorithms experiments work rule statistical basis low  
applications performed methods good research design  
inputs mixture consists region determine ieee vol  
binary study application real presented required  
dimensional task tasks simple size outputs dimension  
speech sample vectors back provide accuracy shows  
small correct desired determined

**topic 22:** (8332) -4.6578

spike firing rate neuron rates train current voltage  
threshold fig activity synaptic potential stimulus  
constant temporal average code inputs action high  
dependent times stochastic change low range eq  
response term note small effect increase and 3rd  
synapses real press due simple dynamics estimate  
difference lower assume simulations cortical  
parameter total large relative line science properties  
higher curve positive sum study individual level