# System Robustness

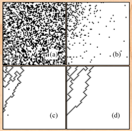
## Principles of Complex Systems CSYS/MATH 300, Fall, 2011

Prof. Peter Dodds

Department of Mathematics & Statistics | Center for Complex Systems | Vermont Advanced Computing Center | University of Vermont







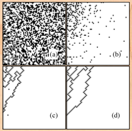
# Outline

## Robustness

- HOT theory
- Self-Organized Criticality
- COLD theory
- Network robustness

## References

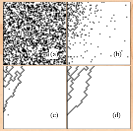




# Robustness



- Many complex systems are prone to cascading catastrophic failure: exciting!!!
  - Blackouts
  - Disease outbreaks
  - Wildfires
  - Earthquakes
- But complex systems also show persistent robustness (not as exciting but important...)
- Robustness and Failure may be a power-law story...

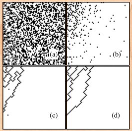


# Robustness





- System robustness may result from
  1. Evolutionary processes
  2. Engineering/Design
- Idea: Explore systems optimized to perform under uncertain conditions.
- The handle:
  'Highly Optimized Tolerance' (HOT) [4, 5, 6, 9]
- The catchphrase: Robust yet Fragile
- The people: Jean Carlson and John Doyle (⊞)

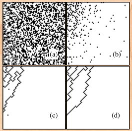


# Robustness



## Features of HOT systems: [5, 6]

- High performance and robustness
- Designed/evolved to handle known stochastic environmental variability
- Fragile in the face of unpredicted environmental signals
- Highly specialized, low entropy configurations
- Power-law distributions appear (of course...)



# Robustness

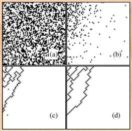




## HOT combines things we've seen:

- Variable transformation
- Constrained optimization

- Need power law transformation between variables: $(Y = X^{-\alpha})$
- Recall PLIPLO is bad...
- MIWO is good: Mild In, Wild Out
- $X$ has a characteristic size but $Y$ does not

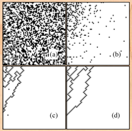


# Robustness



## Forest fire example: [5]

- Square $N \times N$ grid
- Sites contain a tree with probability $\rho$ = density
- Sites are empty with probability $1 - \rho$
- Fires start at location $(i, j)$ according to some distribution $P_{ij}$
- Fires spread from tree to tree (nearest neighbor only)
- Connected clusters of trees burn completely
- Empty sites block fire
- Best case scenario:
  Build firebreaks to maximize average # trees left intact given one spark



# Robustness

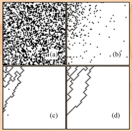


## Forest fire example: [5]

- Build a forest by adding one tree at a time
- Test $D$ ways of adding one tree
- $D$ = design parameter
- Average over $P_{ij}$ = spark probability
- $D = 1$: random addition
- $D = N^2$: test all possibilities

## Measure average area of forest left untouched

- $f(c)$ = distribution of fire sizes $c$ (= cost)
- Yield = $Y = \rho - \langle c \rangle$



# Robustness





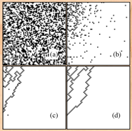
## Specifics:

- $$P_{ij} = P_{i;a_x,b_x} P_{j;a_y,b_y}$$

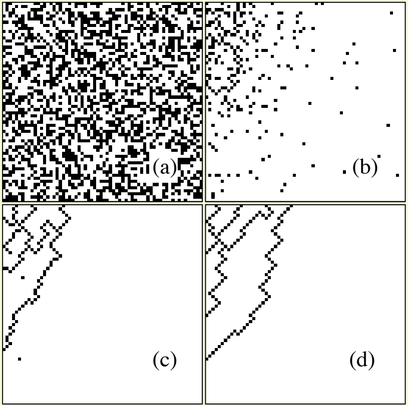
  where

  $$P_{i;a,b} \propto e^{-[(i+a)/b]^2}$$

- In the original work, $b_y > b_x$
- Distribution has more width in $y$ direction.



# HOT Forests




[5]

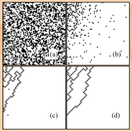
$N = 64$

(a) $D = 1$
(b) $D = 2$
(c) $D = N$
(d) $D = N^2$

$P_{ij}$ has a Gaussian decay

- Optimized forests do well on average (robustness)
- But rare extreme events occur (fragility)





# HOT Forests



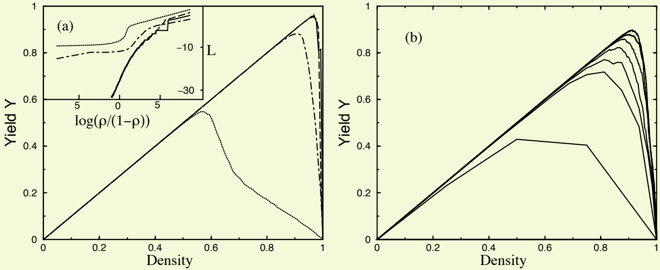
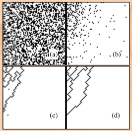



FIG. 2. Yield vs density $Y(\rho)$: (a) for design parameters $D = 1$ (dotted curve), 2 (dot-dashed), $N$ (long dashed), and $N^2$ (solid) with $N = 64$, and (b) for $D = 2$ and $N = 2, 2^2, \ldots, 2^7$ running from the bottom to top curve. The results have been averaged over 100 runs. The inset to (a) illustrates corresponding loss functions $L = \log[\langle f \rangle/(1 - \langle f \rangle)]$, on a scale which more clearly differentiates between the curves.

[5]



# HOT Forests:

- $Y$ = 'the average density of trees left unburned in a configuration after a single spark hits.' [5]

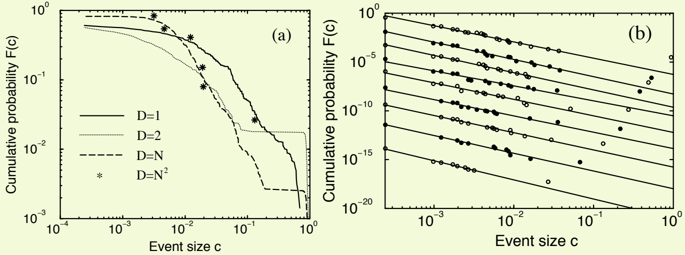
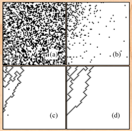



FIG. 3. Cumulative distributions of events $F(c)$: (a) at peak yield for $D = 1, 2, N$, and $N^2$ with $N = 64$, and (b) for $D = N^2$, and $N = 64$ at equal density increments of 0.1, ranging at $\rho = 0.1$ (bottom curve) to $\rho = 0.9$ (top curve).



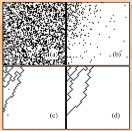


# Random Forests



## $D = 1$: Random forests = Percolation [10]

- Randomly add trees
- Below critical density $\rho_c$, no fires take off
- Above critical density $\rho_c$, percolating cluster of trees burns
- Only at $\rho_c$, the critical density, is there a power-law distribution of tree cluster sizes
- Forest is random and featureless

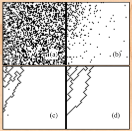


# HOT forests



## HOT forests nutshell:

- Highly structured
- Power law distribution of tree cluster sizes for $\rho > \rho_c$
- No specialness of $\rho_c$
- Forest states are tolerant
- Uncertainty is okay if well characterized
- If $P_{ij}$ is characterized poorly, failure becomes highly likely



# HOT forests—Real data: [6]



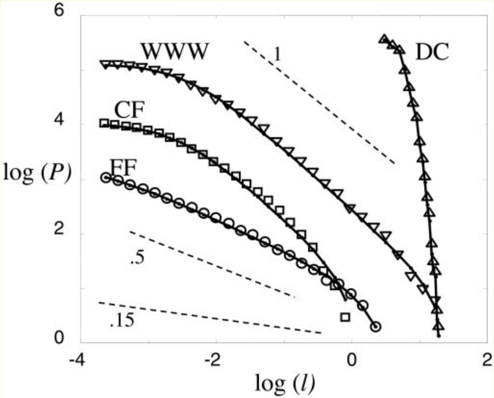
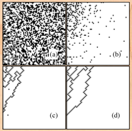



**Fig. 1.** Log–log (base 10) comparison of DC, WWW, CF, and FF data (symbols) with PLR models (solid lines) (for $\beta = 0, 0.9, 0.9, 1.85$, or $\alpha = 1/\beta = \infty, 1.1, 1.1, 0.054$, respectively) and the SOC FF model ($\alpha = 0.15$, dashed). Reference lines of $\alpha = 0.5$, 1 (dashed) are included. The cumulative distributions of frequencies $\mathcal{P}(l \geq l_i)$ vs. $l_i$ describe the areas burned in the largest 4,284 fires from 1986 to 1995 on all of the U.S. Fish and Wildlife Service Lands (FF) (17), the >10,000 largest California brushfires from 1878 to 1999 (CF) (18), 130,000 web file transfers at Boston University during 1994 and 1995 (WWW) (19), and code words from DC. The size units [1,000 $km^2$ (FF and CF), megabytes (WWW), and bytes (DC)] and the logarithmic decimation of the data are chosen for visualization.



# HOT theory



## The abstract story:

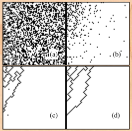
- Given $y_i = x_i^{-\alpha}$, $i = 1, \ldots, N_{\text{sites}}$
- Design system to minimize $\langle y \rangle$ subject to a constraint on the $x_i$
- Minimize cost:

$$C = \sum_{i=1}^{N_{\text{sites}}} Pr(y_i) y_i$$

Subject to $\sum_{i=1}^{N_{\text{sites}}} x_i = \text{constant}$

- Drag out the Lagrange Multipliers, battle away and find:

$$p_i \propto y_i^{-\gamma}$$



# HOT Theory—Two costs:





1. Expected size of fire:

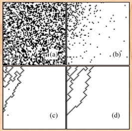
$$C_{\text{fire}} \propto \sum_{i=1}^{N_{\text{sites}}} (p_i a_i) a_i = \sum_{i=1}^{N_{\text{sites}}} p_i a_i^2$$

- $a_i$ = area of $i$th site's region
- $p_i$ = avg. prob. of fire at site in $i$th site's region
- $N_{\text{sites}}$ = total number of sites

2. Cost of building and maintaining firewalls

$$C_{\text{firewalls}} \propto \sum_{i=1}^{N_{\text{sites}}} a_i^{1/2} a_i^{-1}$$

- We are assuming isometry.
- In $d$ dimensions, 1/2 is replaced by $(d-1)/d$



# HOT theory



## Extra constraint:

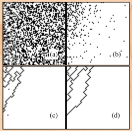
- Total area is constrained:

$$\sum_{i=1}^{N_{\text{sites}}} 1 = N^2.$$

$$\sum_{i=1}^{N_{\text{sites}}} \frac{1}{a_i} = N_{\text{regions}}$$

where $N_{\text{regions}}$ = number of cells.

- Can ignore in calculation...

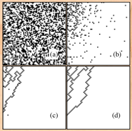


# HOT theory

- Minimize $C_{\text{fire}}$ given $C_{\text{firewalls}} = \text{constant}$.
-
$$0 = \frac{\partial}{\partial a_j}\left(C_{\text{fire}} - \lambda C_{\text{firewalls}}\right)$$
$$\propto \frac{\partial}{\partial a_j}\left(\sum_{i=1}^{N} p_i a_i^2 - \lambda' a_i^{(d-1)/d} a_i^{-1}\right)$$
-
$$p_i \propto a_i^{-\gamma} = a_i^{-(2+1/d)}$$
-
$$\text{For } d = 2, \gamma = 5/2$$







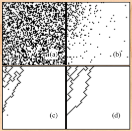
# HOT theory





## Summary of designed tolerance [6]

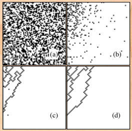
- Build more firewalls in areas where sparks are likely
- Small connected regions in high-danger areas
- Large connected regions in low-danger areas
- Routinely see many small outbreaks (robust)
- Rarely see large outbreaks (fragile)
- Sensitive to changes in the environment ($P_{ij}$)



# Avalanches of Sand and Rice...







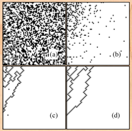
# SOC theory





## SOC = Self-Organized Criticality

- Idea: natural dissipative systems exist at ‘critical states’
- Analogy: Ising model with temperature somehow self-tuning
- Power-law distributions of sizes and frequencies arise ‘for free’
- Introduced in 1987 by Bak, Tang, and Weisenfeld [3, 2, 7]:
  “Self-organized criticality - an explanation of 1/f noise” (PRL, 1987).
- Problem: Critical state is a very specific point
- Self-tuning not always possible
- Much criticism and arguing...



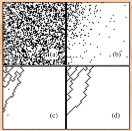
# Robustness





## HOT versus SOC

- Both produce power laws
- Optimization versus self-tuning
- HOT systems viable over a wide range of high densities
- SOC systems have one special density
- HOT systems produce specialized structures
- SOC systems produce generic structures

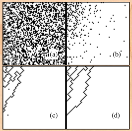


# HOT theory—Summary of designed tolerance [6]



**Table 1. Characteristics of SOC, HOT, and data**

| | Property | SOC | HOT and Data |
|---|---|---|---|
| 1 | Internal configuration | Generic, homogeneous, self-similar | Structured, heterogeneous, self-dissimilar |
| 2 | Robustness | Generic | Robust, yet fragile |
| 3 | Density and yield | Low | High |
| 4 | Max event size | Infinitesimal | Large |
| 5 | Large event shape | Fractal | Compact |
| 6 | Mechanism for power laws | Critical internal fluctuations | Robust performance |
| 7 | Exponent $\alpha$ | Small | Large |
| 8 | $\alpha$ vs. dimension $d$ | $\alpha \approx (d-1)/10$ | $\alpha \approx 1/d$ |
| 9 | DDOFs | Small (1) | Large ($\infty$) |
| 10 | Increase model resolution | No change | New structures, new sensitivities |
| 11 | Response to forcing | Homogeneous | Variable |

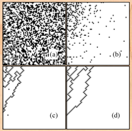


# COLD forests



## Avoidance of large-scale failures

- Constrained Optimization with Limited Deviations [8]
- Weight cost of larges losses more strongly
- Increases average cluster size of burned trees...
- ... but reduces chances of catastrophe
- Power law distribution of fire sizes is truncated

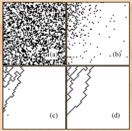


# Cutoffs



## Aside:

- Power law distributions often have an exponential cutoff
$$P(x) \sim x^{-\gamma} e^{-x/x_c}$$
where $x_c$ is the approximate cutoff scale.
- May be Weibull distributions:
$$P(x) \sim x^{-\gamma} e^{-ax^{-\gamma+1}}$$



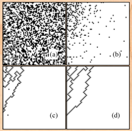

The UNIVERSITY of VERMONT

# Robustness



We'll return to this later on:

- network robustness.
- Albert et al., Nature, 2000:
  "Error and attack tolerance of complex networks" [1]
- Similar robust-yet-fragile story...
- See Networks Overview, Frame 67ish (⊞)




The UNIVERSITY of VERMONT

# References I

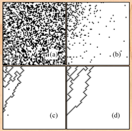

[1] R. Albert, H. Jeong, and A.-L. Barabási.
Error and attack tolerance of complex networks.
Nature, 406:378–382, 2000. pdf (⊞)

[2] P. Bak.
How Nature Works: the Science of Self-Organized Criticality.
Springer-Verlag, New York, 1996. pdf (⊞)

[3] P. Bak, C. Tang, and K. Wiesenfeld.
Self-organized criticality - an explanation of 1/f noise.
Phys. Rev. Lett., 59(4):381–384, 1987. pdf (⊞)

[4] J. M. Carlson and J. Doyle.
Highly optimized tolerance: A mechanism for power laws in design systems.
Phys. Rev. E, 60(2):1412–1427, 1999. pdf (⊞)



# References II

[5] J. M. Carlson and J. Doyle.
Highly optimized tolerance: Robustness and design in complex systems.
Phys. Rev. Lett., 84(11):2529–2532, 2000. pdf (⊞)

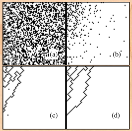
[6] J. M. Carlson and J. Doyle.
Complexity and robustness.
Proc. Natl. Acad. Sci., 99:2538–2545, 2002. pdf (⊞)

[7] H. J. Jensen.
Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems.
Cambridge Lecture Notes in Physics. Cambridge University Press, Cambridge, UK, 1998.

[8] M. E. J. Newman, M. Girvan, and J. D. Farmer.
Optimal design, robustness, and risk aversion.
Phys. Rev. Lett., 89:028301, 2002.







# References III



[9] D. Sornette.
Critical Phenomena in Natural Sciences.
Springer-Verlag, Berlin, 2nd edition, 2003.

[10] D. Stauffer and A. Aharony.
Introduction to Percolation Theory.
Taylor & Francis, Washington, D.C., Second edition, 1992.