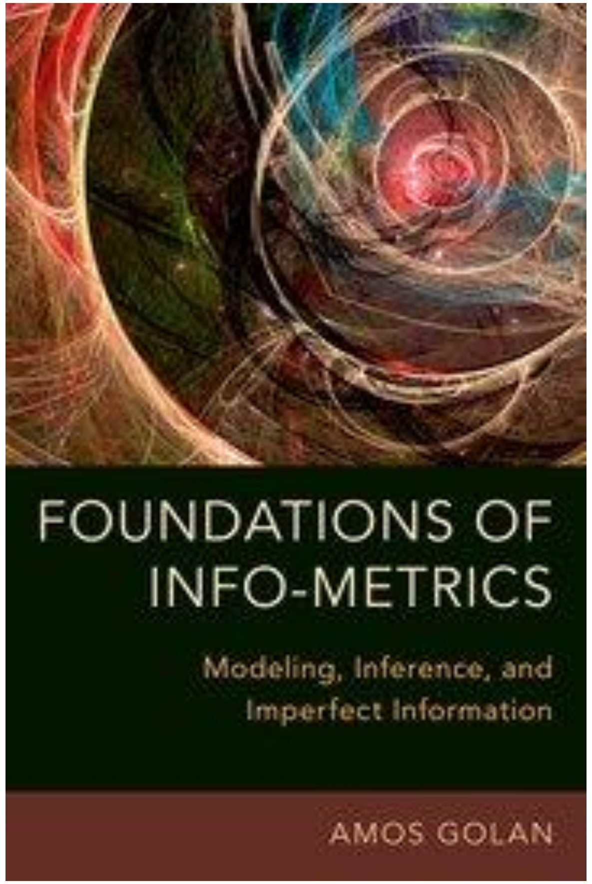


http://info-metrics.org/

AMERICAN UNIVERSITY
WASHINGTON, DC
College of Arts and Sciences
Info-Metrics Institute

Info-Metrics, Complexity and Systems Far from Steady State

Amos Golan
American University; Santa Fe Institute; Pembroke College, Oxford

(SFI, March 15-16, 2018)

A Brief Presentation:

- The Info-Metrics Framework via Four Illustrations
- The Working Group Questions – Some Thoughts

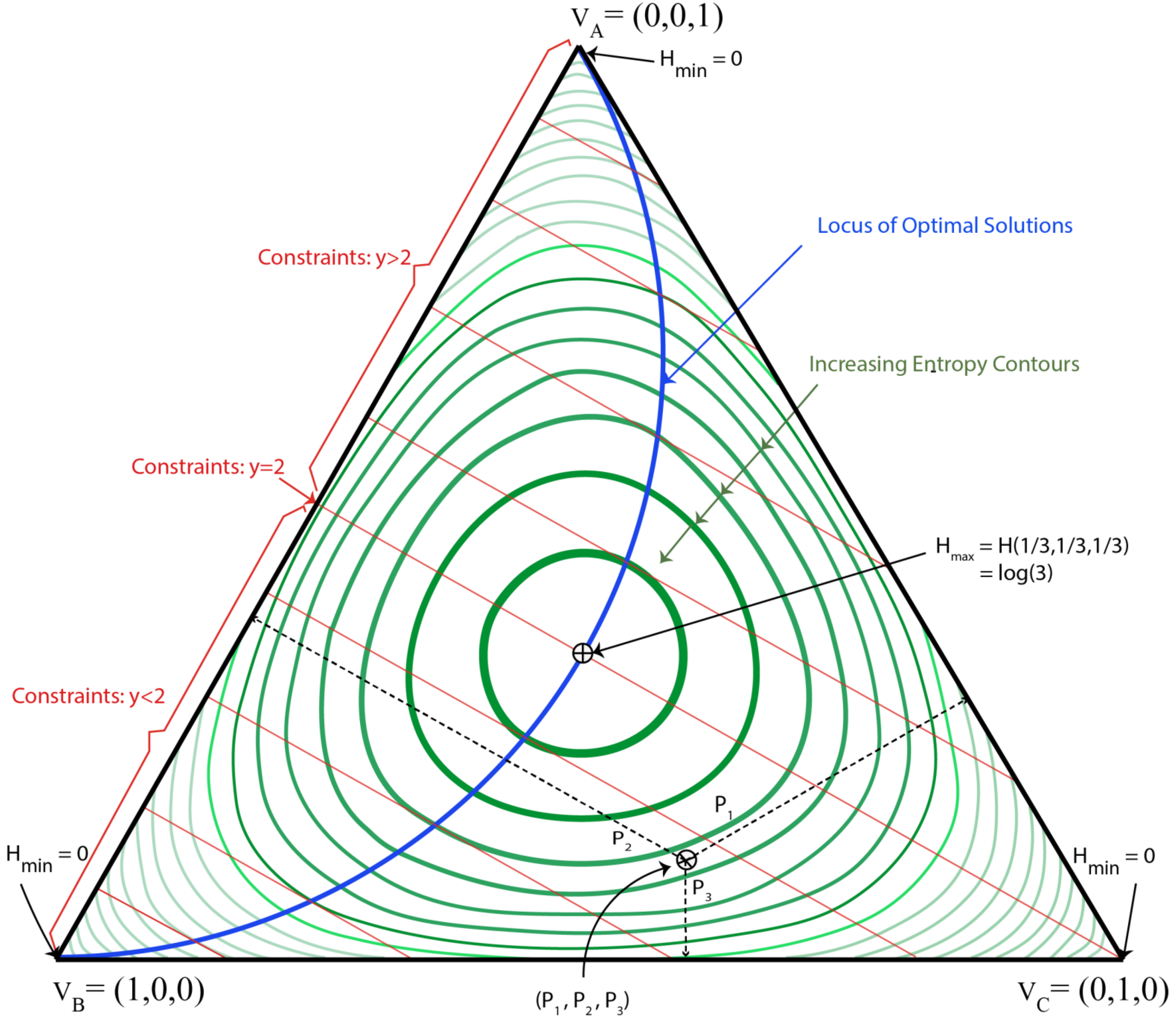
A Visual representation of the theory:

The ‘Classical’ Theory

(A Three-Sided Die)

The Simple/Basic Problem:

- We know the empirical mean value (first moment) of N tosses of the die.
- With that information we want to predict the probability that in the next throw of the die we will observe the value 1, 2 or 3.
- We also know that the probability is proper (sum of the probabilities is one):
    - There are **three** unknown values and only **two** observed values (Underdetermined problem).
- There are infinitely many probability distributions that sum up to one and satisfy the observed mean.
- Question: What would you bet on if the mean (after 100 games) is 1.5? 2.6? 3? ….



Theory: **3-Sided Die**
(with values 1, 2, 3)

Objective: Infer the probability of observing 1, 2 or 3.

Observed information:
Mean (say 2 or 2.2);

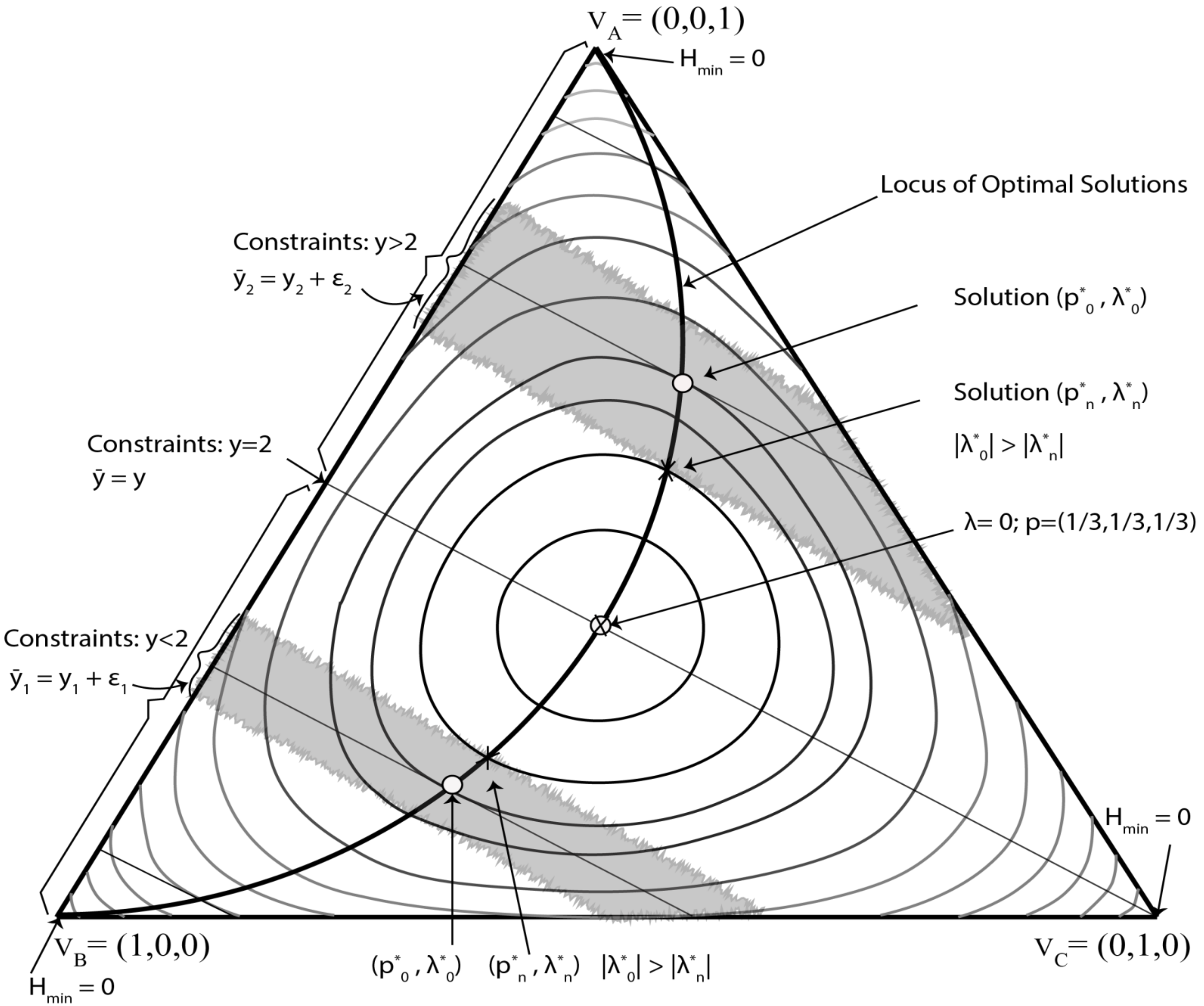
Other known info:
Sum of probabilities is 1.

This is a simplex representation of the maximum entropy problem and solution for a discrete probability distribution defined over three possible events.

A Visual representation of the theory:

The info-metrics Framework

(The Realistic but Simple Case: All Types of Uncertainties may Exist)

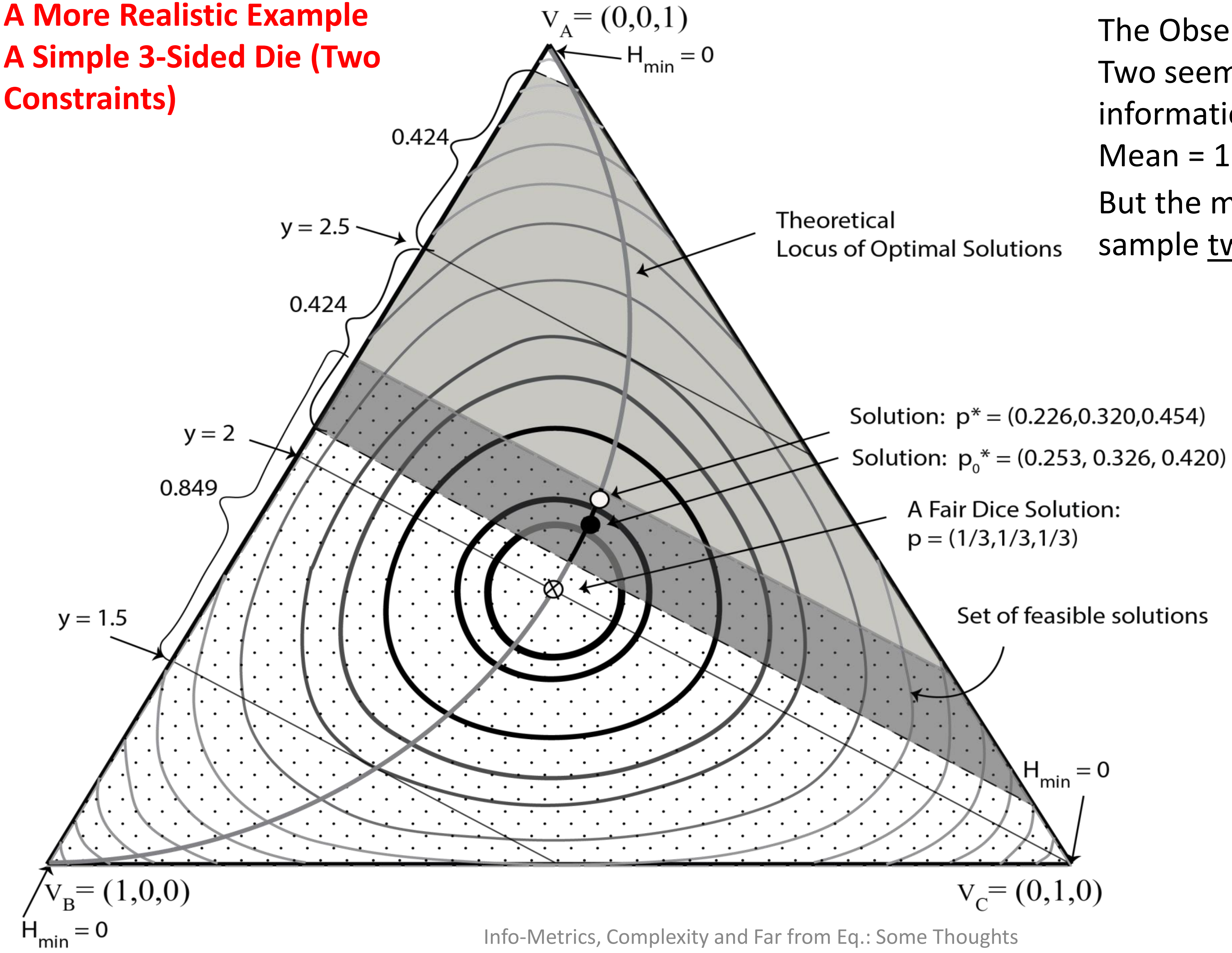


Add Noise – Uncertainty on the model (Stochastic Constraints or stochastic moments)

The Basic Structure and Solution.

Here we Maximize the joint entropy of the Probability Distribution of interest <u>and</u> that of the Noise Subject to the Stochastic Constraints and normalizations of the two sets of Prob Distributions

A More Realistic Example
A Simple 3-Sided Die (Two Constraints)

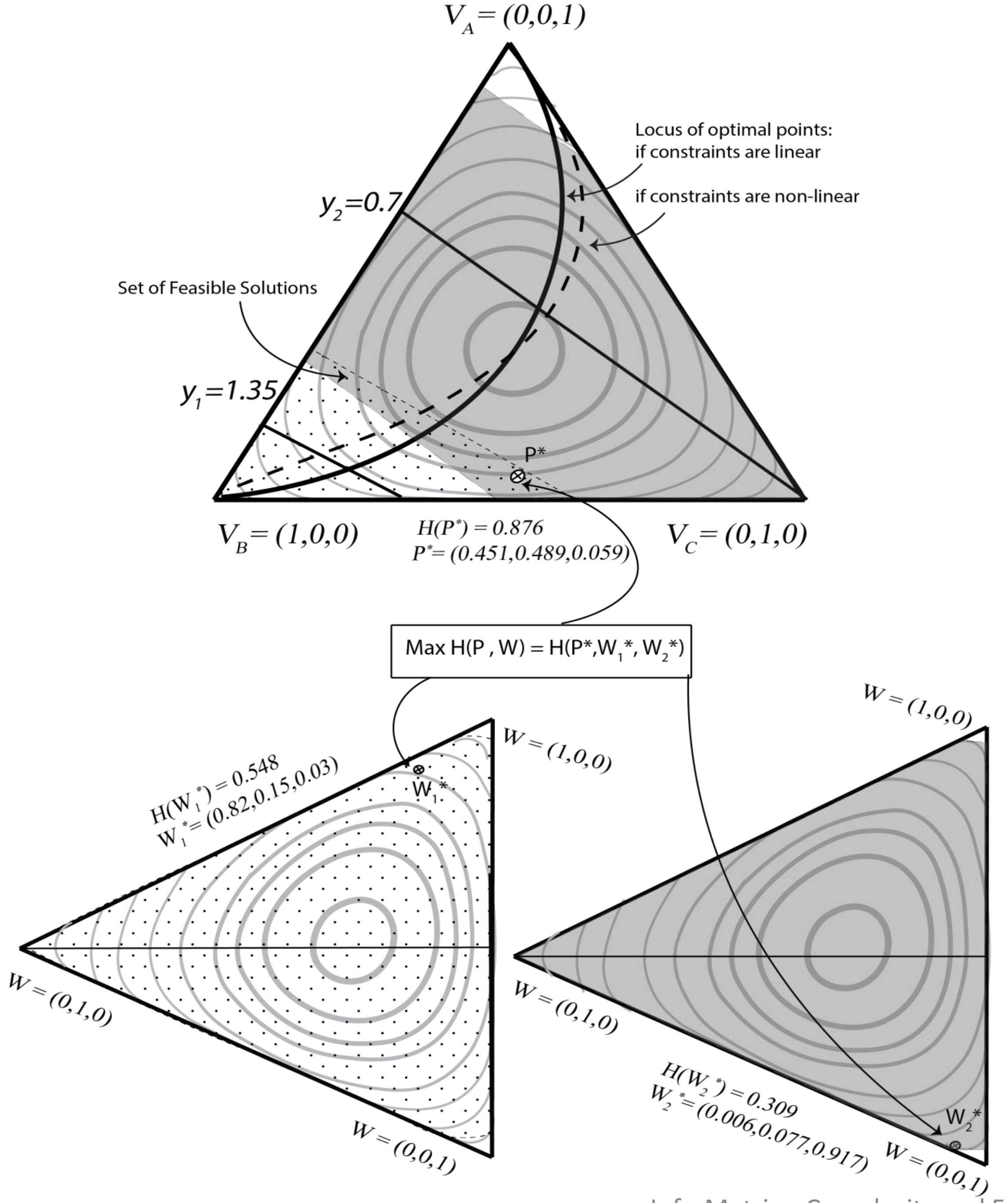


The Observed Information:

Two seemingly contradicting pieces of information:

Mean = 1.5 and Mean = 2.5

But the mean of 2.5 comes from a sample twice as large as the 1.5 mean.



A Simple Example (with nonlinearity)

The Information:
One Arithmetic Mean (Y=1.35)
One Geometric Mean (Y=0.7)

Set of Feasible Solutions: The Intersection of all uncertainties

The Solution: At Maximum H(P,W)
(The joint entropy of P and the Probability Defined over the Noise W)

Initial Thoughts On Q1 and Q2:

Q1 (hierarchical scales). I don't have anything to add on that right now. But one question: Is this related to modeling with different time scales (say micro vs macro time scales)?

Q2 (far from steady state).

- One practical solution is applying the info-metrics approach using stochastic constraints. But it may not be a good representation of a system <u>very</u> far from equilibrium (steady state)
- Other practical solution: Use info-metrics to model the dynamics (e.g., Markov process): The solution is the closest 'Markov process' that captures the dynamics and is fully consistent with the observed (time series) information.

But what if we don't know the (approximate) dynamics?

Use the framework to figure is out by studying the constraints.

Thank You!!

The Very Basic Idea:

The available information (for modeling, inference and decision making) is most often insufficient to provide a unique answer or solution for most interesting decisions or inferences for problems across all disciplines.

Implication:

There is a continuum of solutions (models, inferences, decisions) that are consistent with the information (most often ‘circumstantial’ evidence/data) we have.

Therefore:

We can impose many different requirements on the data. These requirements/information can be correct or not. Regardless, they force the data to ‘confess’ too fast. Different requirements (‘tortures’) will result in different solutions (confessions), most are untrue.

Classical methods:

Impose too much structure. They torture the data to confess too fast.

Info-Metrics:

Uses minimally needed information (minimal/no torture)

Why we cannot come up with the same conclusion (solution, decision, inference or model)?

- It is not because of the evidence/data!
- It is because of the soft information (assumptions, structures, intuition, values…) imposed in the inferential method.
- Each person uses a different 'input information.' The evidence (data, hard information or circumstantial evidence) is just one component of the input information used to solve a problem.

Modeling and ‘Confessions’ of Data – What are the most common problems?

- We don’t know the statistical properties of the data.
- We don’t know the functional/structural form of the system we try to model.
- We don’t know all of the information needed to construct a theory (or model).
- The systems we model (or develop theories to) are constantly evolving.
- The information often comes from different sources; it can even be somewhat contradictory, and of different structures.

In all of the above, we can practically ‘squeeze’ our evidence or ‘data’ (information) to practically tell us anything – the more correct story and/or the ‘tortured/false’ story.

The Problem:

All of these problems (modeling, inferential and decision making) are **under-determined**.
Unless we use additional ‘soft’ information we do not get a unique solution.

(This is especially the case for complex, small and collinear data, or data with extreme events.)

Info-Metrics: *Definition*

info-metrics is the science of modeling, inference, and reasoning under conditions of noisy and insufficient information.

- It deals with the question of how to model effectively, draw appropriate inferences, and make informed decisions when dealing with inadequate or incomplete information.
- It also deals with the complementary question of how to process the available information while imposing minimal assumptions (or structures) that cannot be validated.

The Suggested Solution:

A Constrained Optimization Framework

- Information as constraints – or stochastic constraints;
- Objective function defined simultaneously on the entities of interest and the uncertainty surrounded the information/model.

The info-metrics framework combines the tools of information theory and statistical inference within a constraints optimization framework

More precisely:

All inferential problems can be converted to optimization problems:

Optimize a Certain Decision Function

Subject to

Constraints (Information; ‘conservation laws’)

(normalization)

The issues:

- What certain function to optimize? (Entropy)
- Constraints (what constraints? How to specify? Validate? Uncertainty?)
- Priors (from where?)
- Soft information (validate? Quantify? Measure?)

Example (Cont.): Six-Sided Die

The Optimization Problem

Assume the observed mean value, after *n* tosses is y.

The Classical Maximum Entropy formulation is

$$\underset{\{p\}}{Max}\ H(\mathbf{p}) = -\sum_{k=1}^{6} p_k \log_2 p_k$$

$$s.t$$

$$\sum_k p_k x_k = y \text{ and } \sum_k p_k = 1$$

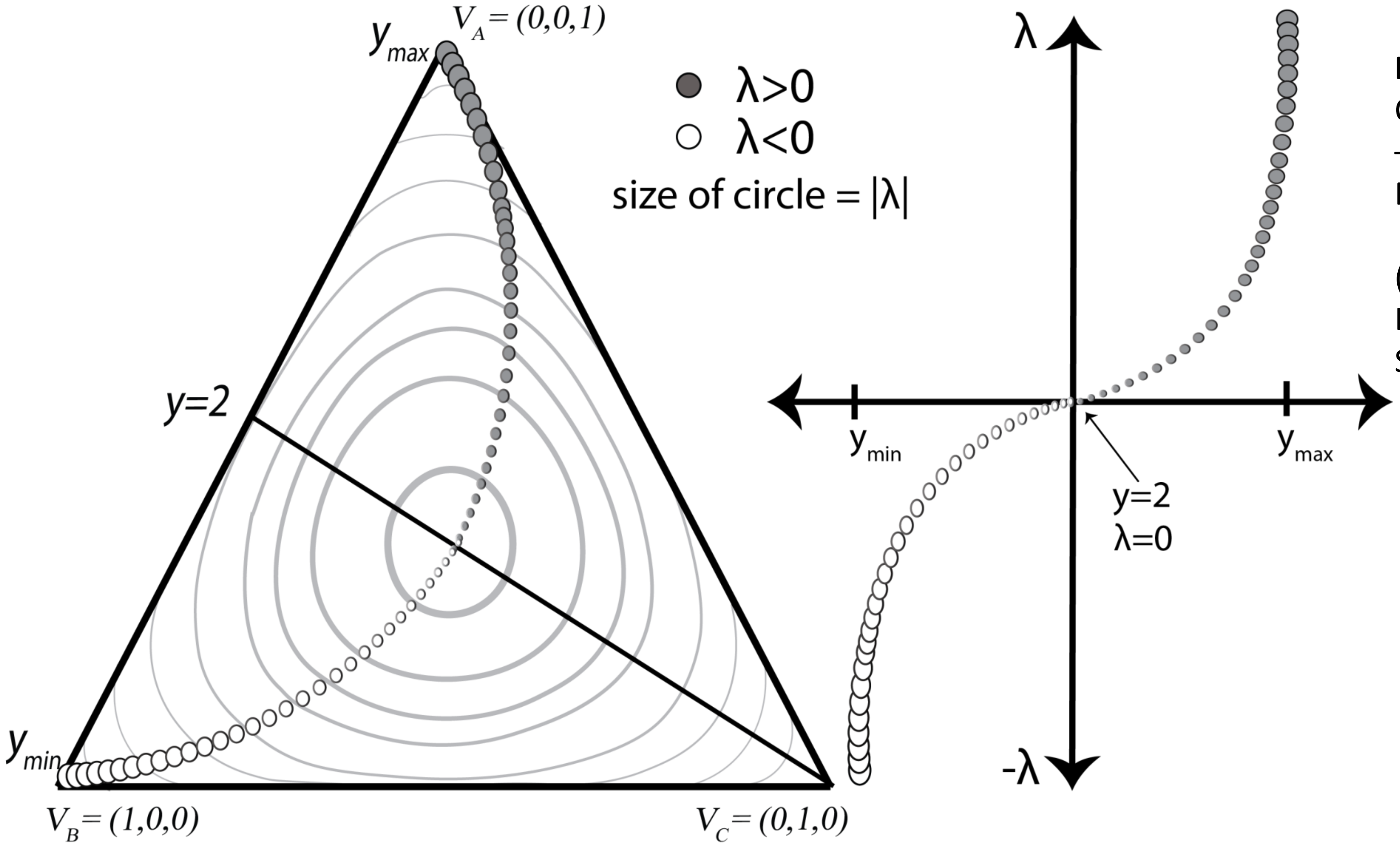
for x_k=1,…,6 for k=1,…,6.

Example (Cont.): Six-Sided Die

The Solution:

$$\hat{p}_k = \frac{2^{-\hat{\lambda} x_k}}{\sum_{k=1}^{6} 2^{-\hat{\lambda} x_k}} \equiv \frac{2^{-\hat{\lambda} x_k}}{\Omega}$$

For y=5 (which is almost as large as it can be), the solution $\hat{\mathbf{p}}$ is 0.021, 0.038, 0.072, 0.136, 0. 255 and 0.478 respectively for *k*=1,…, 6, and H()=1.973 (very small relative to the entropy of the uniform distribution: H_{Max} = 2.585.)



Primal (Constrained Optimization) Model – **Dual** (Concentrated) Model

(Or Prob Space vs. Lagrange Multiplier Space)