**Lachlan J. Gunn**[1] François Chapeau-Blondeau[2]
Andrew Allison[1] Derek Abbott[1]

[1]School of Electrical and Electronic Engineering
The University of Adelaide

[2]Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS)
University of Angers

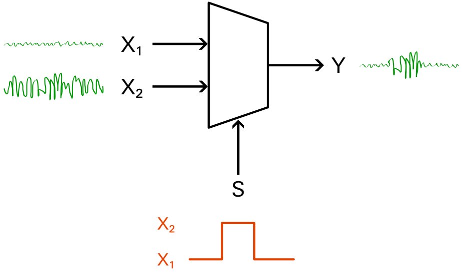
THE UNIVERSITY *of* ADELAIDE

# Towards an information-theoretic model of the Allison mixture

A canonical measure of dependence for a special mixture distribution

# Mixture processes

- Many systems can be split into independent multiplexed subsystems.




- These situations occur quite often.

# Applications of mixture processes

- Radar—sea clutter can be modelled with a KK-distribution

$$p_{KK}(x) = (1-k)\, p_K(x;\, \nu_1,\, \sigma_1) + k\, p_K(x;\, \nu_2,\, \sigma_2)$$

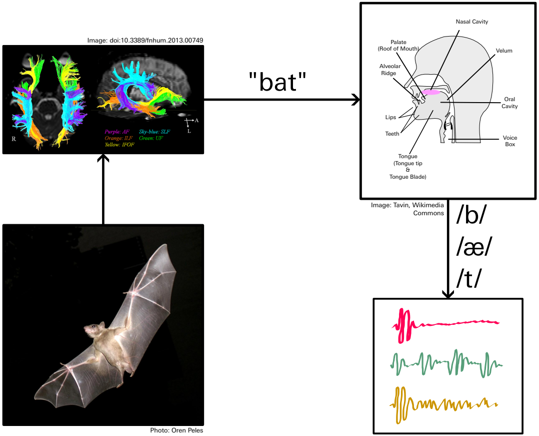

Rolling sea (1 – k) of the time…


…and an occasional spike in the return.

Yunhan Dong, DSTO-RR-0316 “Distribution of X-Band High Resolution and High Grazing Angle Sea Clutter” (2006)

# Applications of mixture processes

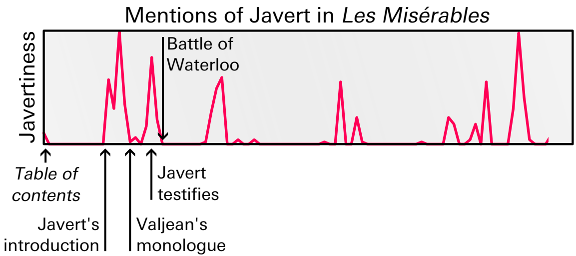
- Speech—random signals corresponding to phenomes are joined to form the sound of words.




Gales, Young, "The Application of Hidden Markov Models in Speech Recognition,', Foundations and Trends in Signal Processing, **1** (3), 2007.

# Applications of mixture processes

- Natural language processing—word frequency distributions can be modelled by a mixture of Poisson distributions.



# The Allison Mixture

- Which input process is visible at the output?
- One option is to choose independently each time:

$$p_Y(y) = kp_0(y) + (1-k)p_1(y)$$

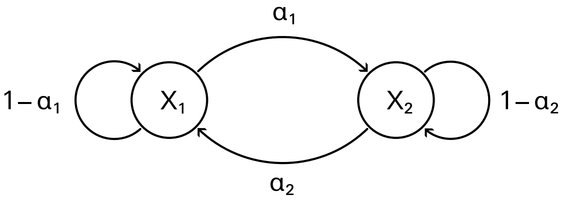
$\Longrightarrow$ Independent inputs yield independent outputs.

- This is too restrictive, often the choices of sample are dependent.

# The Allison Mixture



- Instead, let's—sometimes—switch from one input to the other.

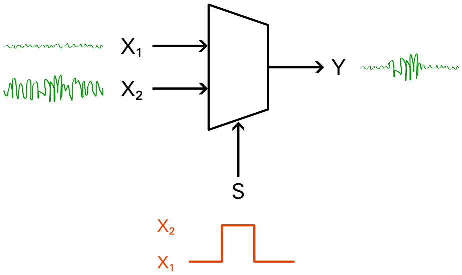



- This forms a Markov chain.

# Independence of the Allison mixture



- This makes output samples dependent, even if the inputs are not.



# Autocovariance

THE UNIVERSITY of ADELAIDE

- It is known that the lag-one autocovariance is given by

$$R_{YY}[1] = \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2}(1 - \alpha_1 - \alpha_2)(\mu_1 - \mu_2)^2,$$

where $\mu_1$ and $\mu_2$ are the input means.

- If $X_1$ and $X_2$ are $\mathcal{N}(\mu_i, \sigma^2)$ then this is just a noisy version of S.

Gunn, Allison, Abbott, "Allison mixtures: where random digits obey thermodynamic principles", International Journal of Modern Physics: Conference Series **33** (2014).

# Uncorrelatedness

- This gives us the conditions for a correlated output:

$$
\begin{aligned}
\alpha_1 &\neq 0 \\
\alpha_2 &\neq 0 \\
\alpha_1 + \alpha_2 &\neq 1 \\
\mu_1 &\neq \mu_2
\end{aligned}
$$

- If any of these are violated, consecutive samples are uncorrelated.

# Moving beyond a single step



- We previously demonstrated only the appearance of correlation in the Allison mixture. Let's fill in the remaining details!

- The state-transition matrix of the Markov chain S is

$$P = \begin{bmatrix} 1 - \alpha_1 & \alpha_2 \\ \alpha_1 & 1 - \alpha_2 \end{bmatrix}.$$

- Taking every k-th sample of an Allison mixture yields another Allison mixture—the choice of input is still Markovian.

Gunn, Chapeau-Blondeau, Allison, Abbott, Unsolved Problems of Noise (2015).

# Moving beyond a single step

- By taking $P^k$ and reading off the minor diagonal, we find the k-step transition probabilities

$$
\begin{aligned}
\alpha_1[k] &= \frac{\alpha_2}{\alpha_1 + \alpha_2}\left[1 - (1 - \alpha_1 - \alpha_2)^k\right] \\
\alpha_2[k] &= \frac{\alpha_1}{\alpha_1 + \alpha_2}\left[1 - (1 - \alpha_1 - \alpha_2)^k\right].
\end{aligned}
$$

- The initial coefficients are the stationary probabilities $\pi_1$ and $\pi_2$ respectively.

# Multi-step autocovariance

- We substitute these back into the autocovariance formula, yielding

$$R_{YY}[k] = \frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}(\mu_1-\mu_2)^2(1-\alpha_1-\alpha_2)^k$$
$$= R_{YY}[1](1-\alpha_1-\alpha_2)^k.$$

- The autocovariance thus decays exponentially with time.

# Autoinformation



- Autoinformation provides a more-easily-computed alternative to entropy rate.
- Information theory lets us capture dependence that does not induce correlation.

### Definition (Autoinformation)

We define the autoinformation $I_{XX}[n,k]$

$$I_{XX}[n,k] = I(X[n]; X[n-k]) \,,$$

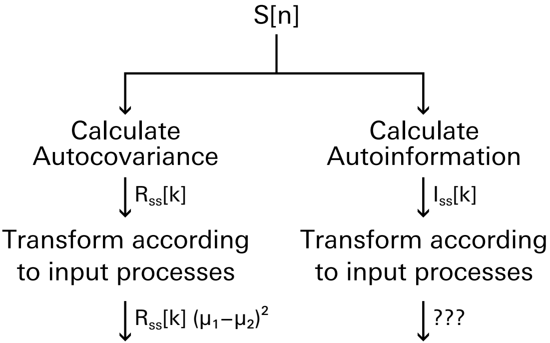
which simplifies, for a stationary process, to

$$I_{XX}[k] = H(X[n], X[n-k]) - 2H(X[n]).$$

# A path towards Allison mixture autoinformation



- Can we take the same approach as with autocovariance?



# Sampling process autoinformation

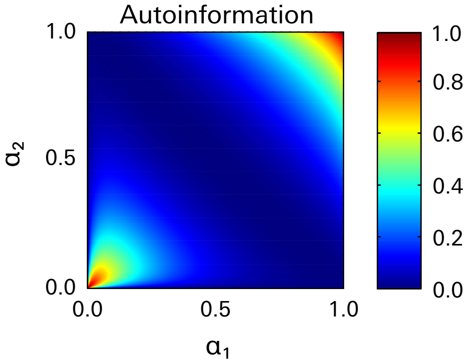


- We compute the autoinformation from the stationary and transition probabilities using $I(X;Y) = H(X) - H(X|Y)$.

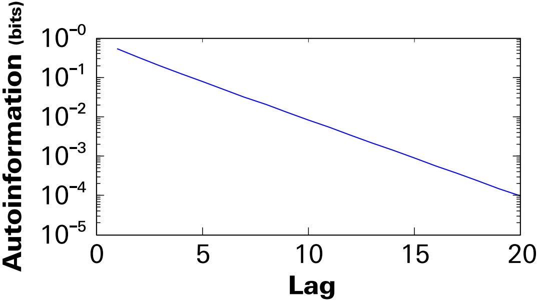
$$
\begin{aligned}
I_{xx}[1] = & \frac{\alpha_2(1-\alpha_1)\log_2 \frac{1-\alpha_1}{\alpha_2}}{\alpha_1+\alpha_2} \\
& + \frac{\alpha_1(1-\alpha_2)\log_2 \frac{1-\alpha_2}{\alpha_1}}{\alpha_1+\alpha_2} \\
& + \log_2(\alpha_1+\alpha_2),
\end{aligned}
\tag{1}
$$

# Sampling process autoinformation





# Sampling process autoinformation



# Allison mixture autoinformation



- How do we apply this to the Allison mixture?

- Binary-valued outputs: $X_1[k], X_2[k] \in 0, 1$
  - Use Bayes law to find the probability of each state:

$$P\left[S[k] = s | Y[k]\right] = \frac{P[Y[k]|S[k] = s]\pi_s}{\sum_q P[Y[k]|S[k] = q]\pi_q}$$

  - We now know enough to find the transition probabilities for Y:

$$Y[k] \longrightarrow S[k] \longrightarrow S[k+1] \longrightarrow Y[k+1]$$

# Allison mixture autoinformation



- It turns out that the previous autoinformation formula works here with transition probabilities

$$\alpha_0' = \frac{\alpha_1(1-p_0)\left[p_0(1-\alpha_0)+p_1\alpha_0\right]+\alpha_0(1-p_1)\left[p_0\alpha_1+p_1(1-\alpha_1)\right]}{\alpha_0(1-p_1)+\alpha_1(1-p_0)}$$

$$\alpha_1' = \frac{\alpha_1 p_0\left[(1-p_0)(1-\alpha_0)+(1-p_1)\alpha_0\right]+\alpha_0 p_1\left[(1-p_0)\alpha_1+(1-p_1)(1-\alpha_1)\right]}{\alpha_0 p_1+\alpha_1 p_0}.$$

- A formula of this complexity that only works for binary processes is not the end of the road.

# Open problems

- Can a similar technique be applied to more general input processes?
  - Continuous distributions are important.
- Could this system be useful for studying transfer entropy?
  - Transfer entropy is the “information transfer” between two systems.
  - Previous studies have revolved around chaotic systems.