Unsupervised Transcription of Historical Documents and Classical Music

Taylor Berg-Kirkpatrick
UC Berkeley
Joint work with Greg Durrett, Jacob Andreas, and Dan Klein

Historical Document Transcription

Historical Document

Old Bailey Court Proceedings 1775

the prifoner at the bar. Jacob Lazarus and his wife, the prifoner, were both together when I received them. I fold eleven pair of them for three guineas, and delivered the remainder back to the prifoner. I fold feven pair of filk to Mark Simper: one pair of mixed, and two pair of thread to the footman, and one pair of thread to the barber.

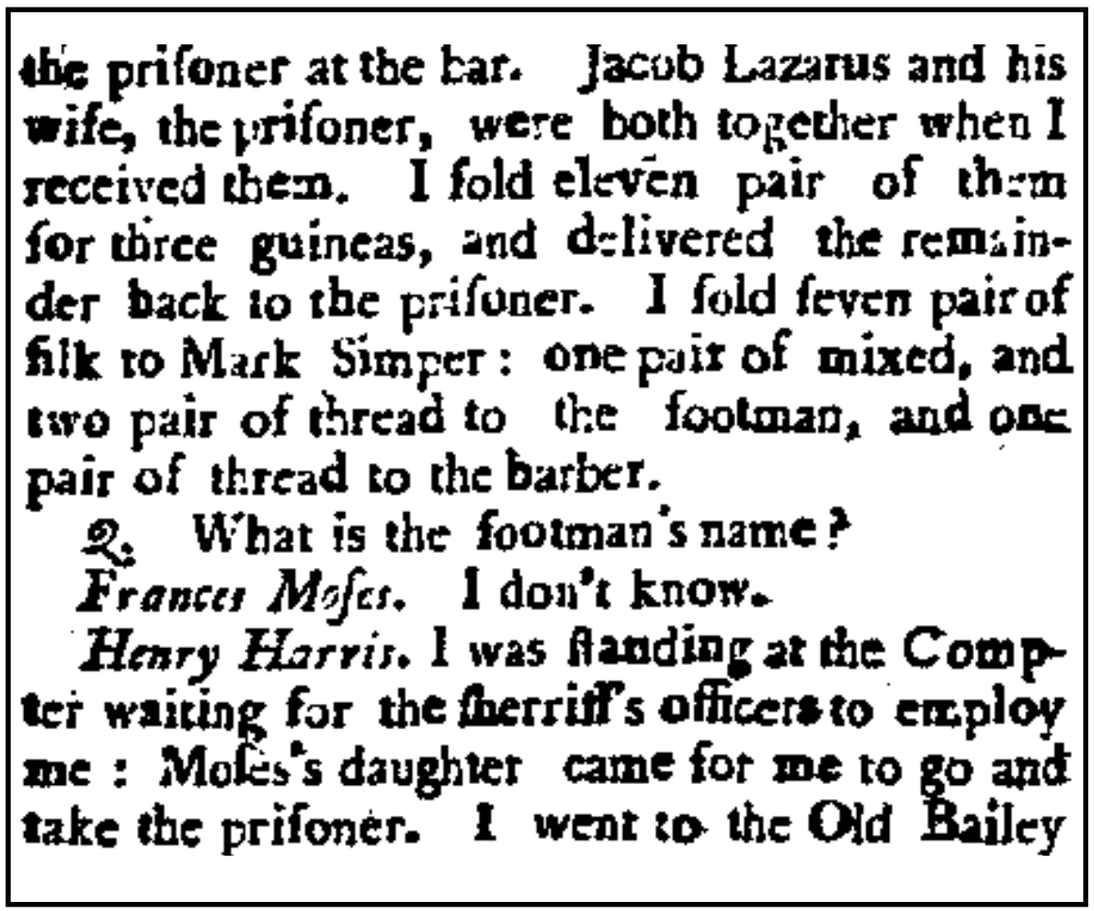
Q. What is the footman's name?

Frances Moſes. I don't know.

Henry Harris. I was ſtanding at the Compter waiting for the ſherriff's officers to employ me: Moſes's daughter came for me to go and take the prifoner. I went to the Old Bailey

Transcription

Document Image



the prifoner at the bar. Jacob Lazarus and his wife, the prifoner, were both together when I received them. I fold eleven pair of them for three guineas, and delivered the remainder back to the prifoner. I fold feven pair of filk to Mark Simper: one pair of mixed, and two pair of thread to the footman, and one pair of thread to the barber.

Q. What is the footman's name?

Frances Moſes. I don't know.

Henry Harris. I was ſtanding at the Compter waiting for the ſherriff's officers to employ me: Moſes's daughter came for me to go and take the prifoner. I went to the Old Bailey

Transcription (Google Tesseract)

and Ch': priftmer anhc bar. Jacob Lazarus and his IHP1 uh: prifoner. were both together when! rcccivcd lhczn. I fold eievén pair of than for xiirce guincas, and dclivcrcd the rcll'l:.in-d:r hack lo :11: prifuner. 1 fold ftvcn pairof filk to Mark Simpcr : nncpuir of mixcd. and. mo pair of Ifircad to lhz: foolnun, and on: pair of zhrzad to lh: barber. '

Q: What is the foolmarfs name?

Fraum Mgfzr. I dun': know.

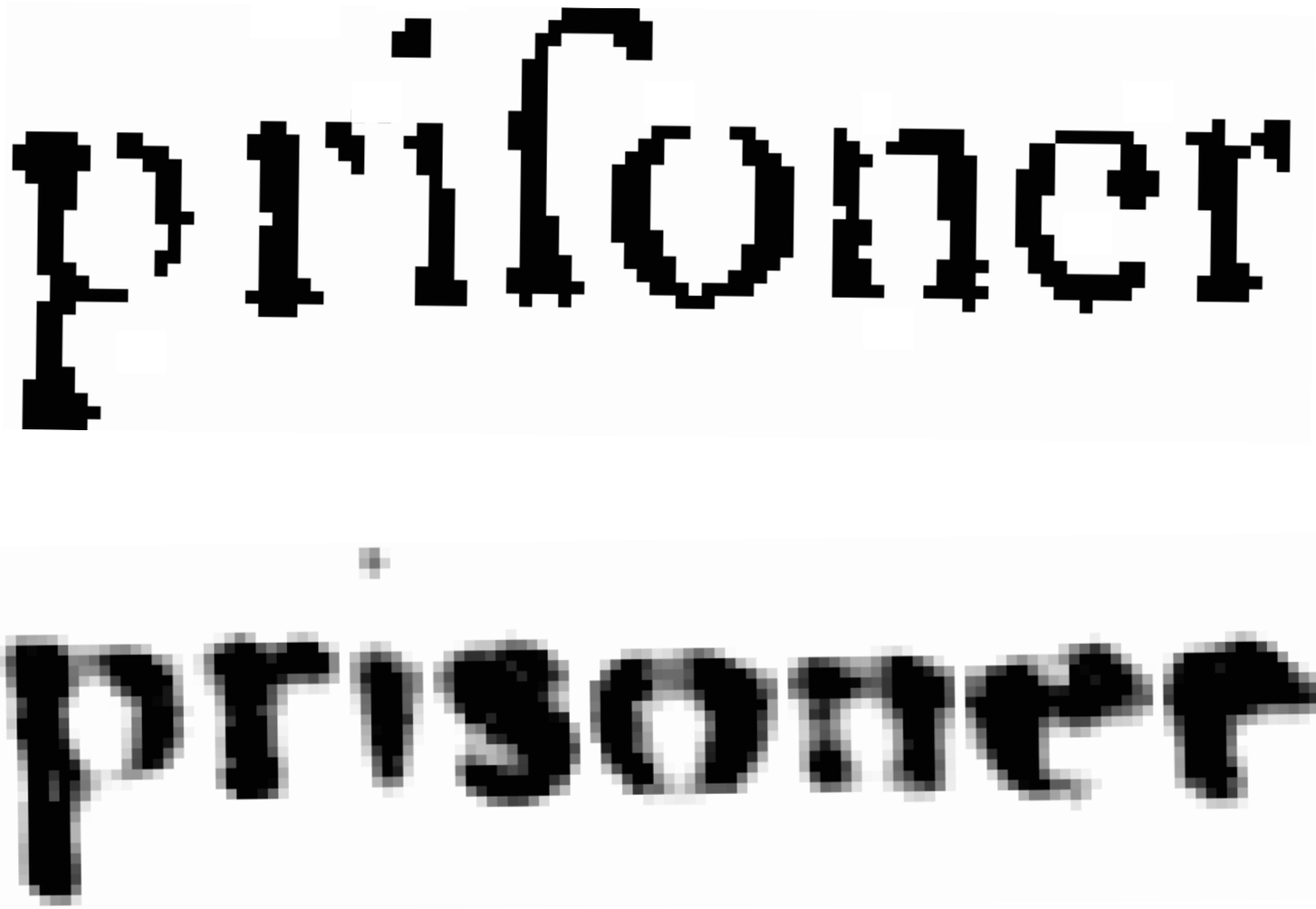
Hairy Hzrvir. l was flandingar the Camp Icr waizin far the thcrrilfs ufliceruo employ in: : Mo 3': daughter came for me to 0 am! take the prifoncr. 1 Wm! to |hc Old aailcy

Unknown Fonts

Unknown Fonts

Berkeley NLP

Unknown Fonts



Wandering Baseline

Wandering Baseline

Wandering Baseline

Uneven Inking

Uneven Inking

Various Historical Documents

1725

tle before the Prisoner came drunk to his Stand, (at Mr. Bird's Door in Castle-Court) and without any Provocation began to be very quarrelsome, swearing, calling him ill Names, and striking him two or three times; Hemmit desired him to get out of his Beat, or he'd make him forfeit Sixpence. (Such a Forfeit being customary among the Watchmen, if one comes into the other's Beat.) Mr. Bird then came to the Door, and threaten'd the Prisoner that he would charge a Constable with him, and send him to Bridewell; upon which the Prisoner was very free of his ill Language to Mr. Bird,

1875

to be conscious at the time—he was caught up and thrown out between them—he could not resist—that was the last I saw—I then went down stairs, and saw him lying on the stones below the window, and his mother came up directly after—I saw a soldier catch hold of William Bagley—all the men went down stairs directly they had thrown him out of the window—there is a court-way leading from the lodging house into the street—I cannot say what way they went, they walked away quickly—I did not see Mr. or Mrs. Rowe or Joseph do anything to promote this attack.

Cross-examined. There were six or seven men by the window at the time he was thrown out—I was standing by the fire-place—I saw him actually thrown out by the four—I could not see whether he went out head first o

1823

in the Police Office at Sydney, on the 7th of November, 1821. Mr. Norton, the plaintiff's solicitor, laid the case before the Judge in nearly the following words.—He stated that his client, in this case, was Mr. James the owner of the schooner Little Mary, a resident at Port Dalrymple, and that the defendant was Mr. Peter Dillon, commander and owner of the late East India ship Fatisalam, with which vessel he sailed from Bengal bound to these Colonies, with a valuable cargo, but

1883:

I have been told. In this case I suspect our voices startled them. We will wrap it carefully as it is."

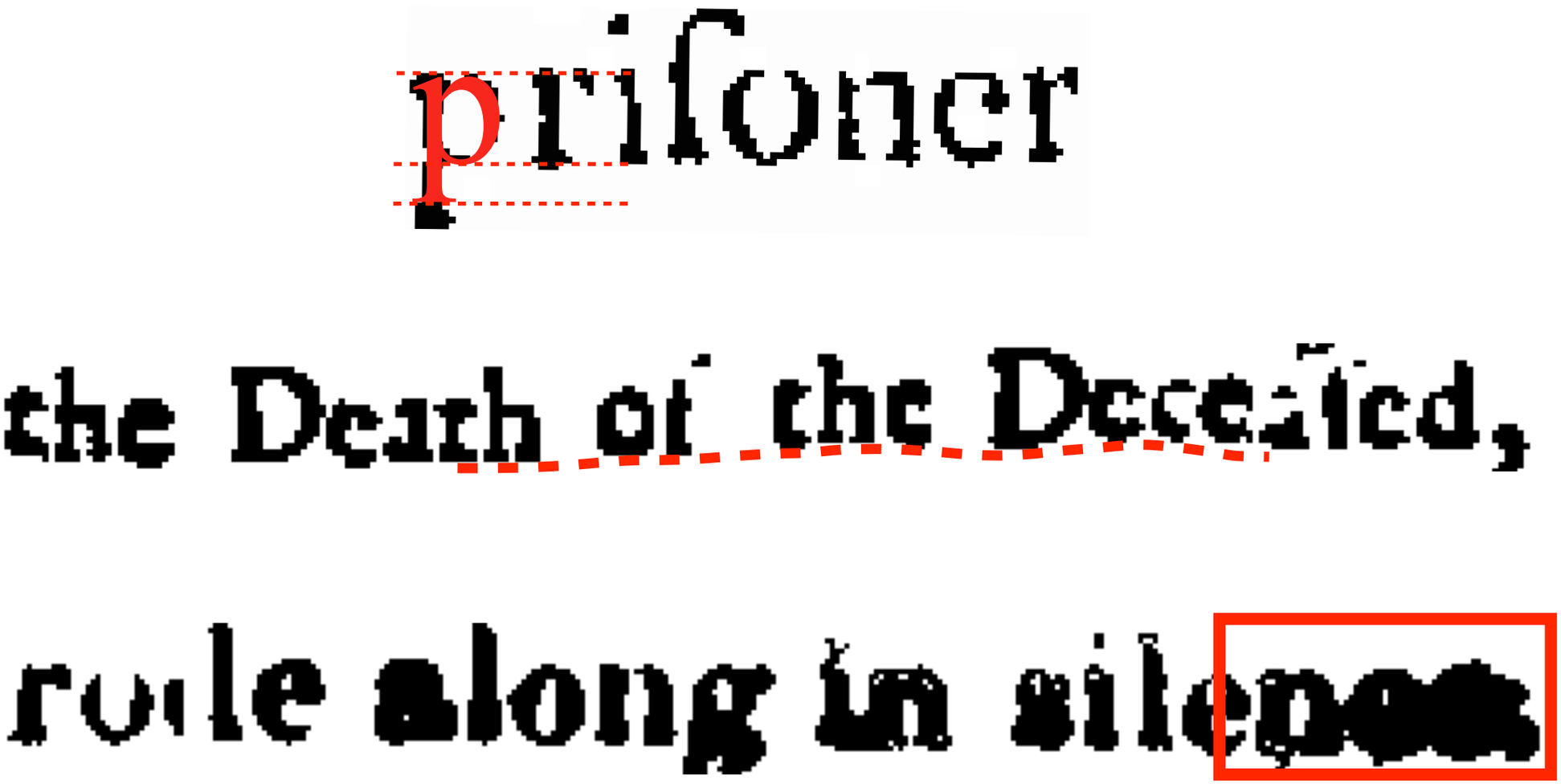
"What for?" said Boolger.

"The murderers must be iden

"Hardly. One tomahawk is just

"Yes, but there are marks o... which, though meaningless to us, are

Our Approach



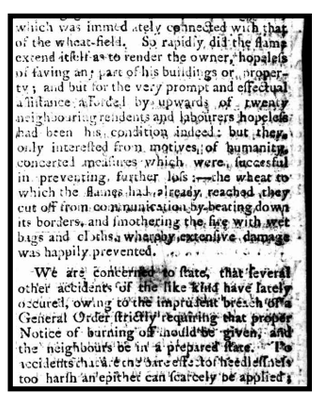
[Berg-Kirkpatrick et al. 2013]
[Berg-Kirkpatrick et al. 2014]

Unsupervised Transcription

Unsupervised Transcription

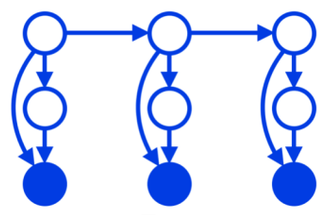
Learning

Doc Image



Generative Model

prisoner



prisoner

Font Parameters

Text

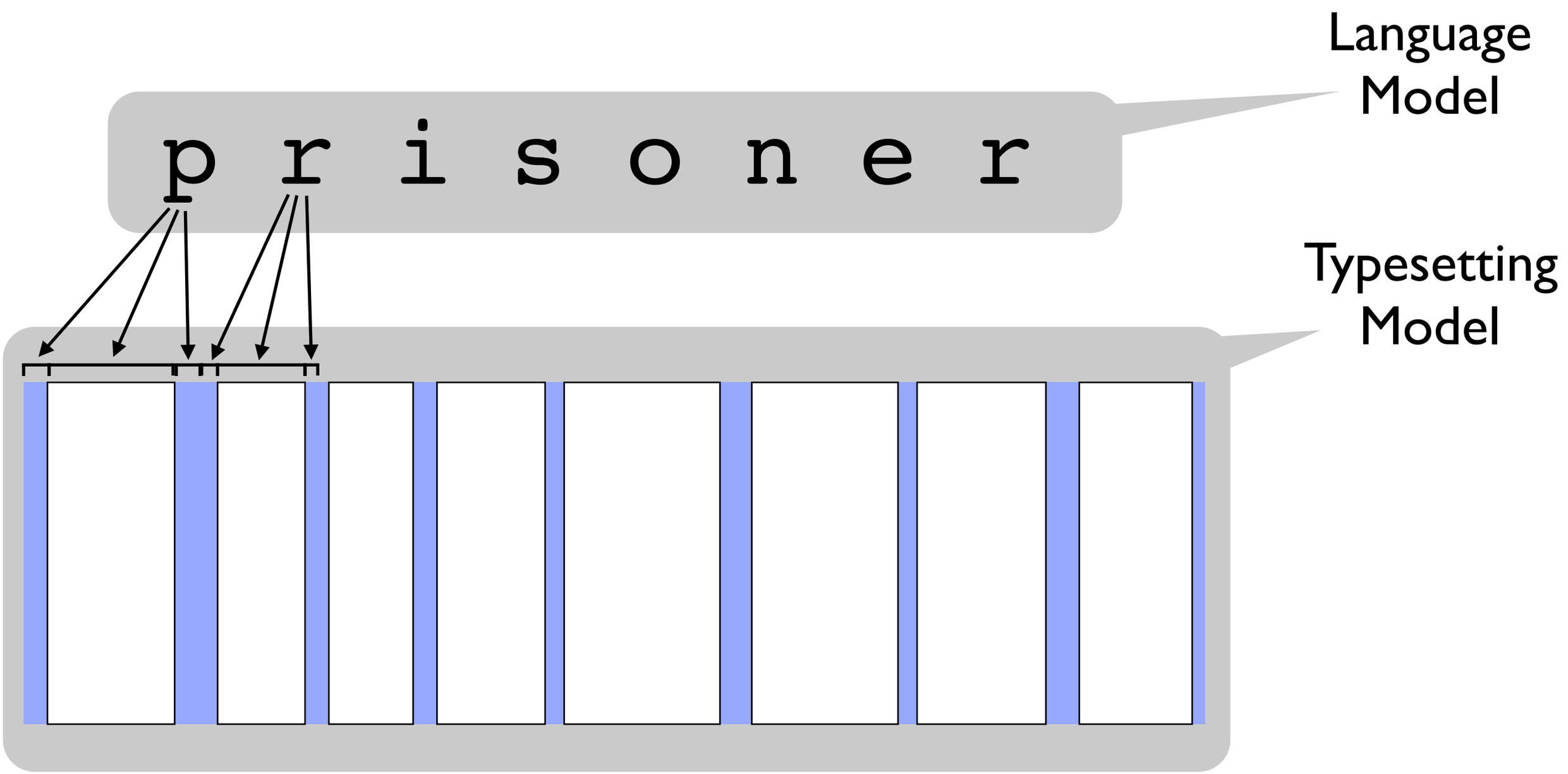
which was immediately connected with that of the wheat-field. So Pollamount came into the Kitchen. The little Girl went away, and the Prisoner and Cornish went out together to see for Peg, but while they were gone, Peg came in again by her self and asked Mrs. Provost, the Bar-keeper, where

?

Berkeley NLP

Generative Model

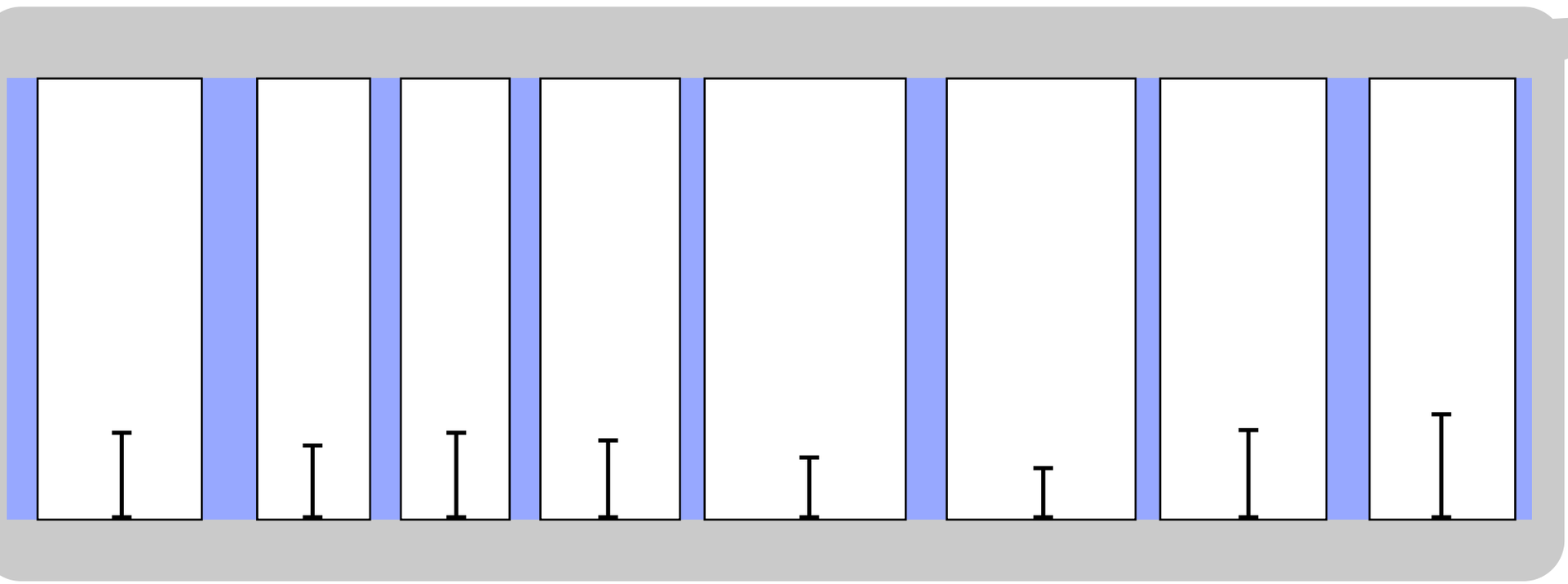
Generative Model



Generative Model

p r i s o n e r

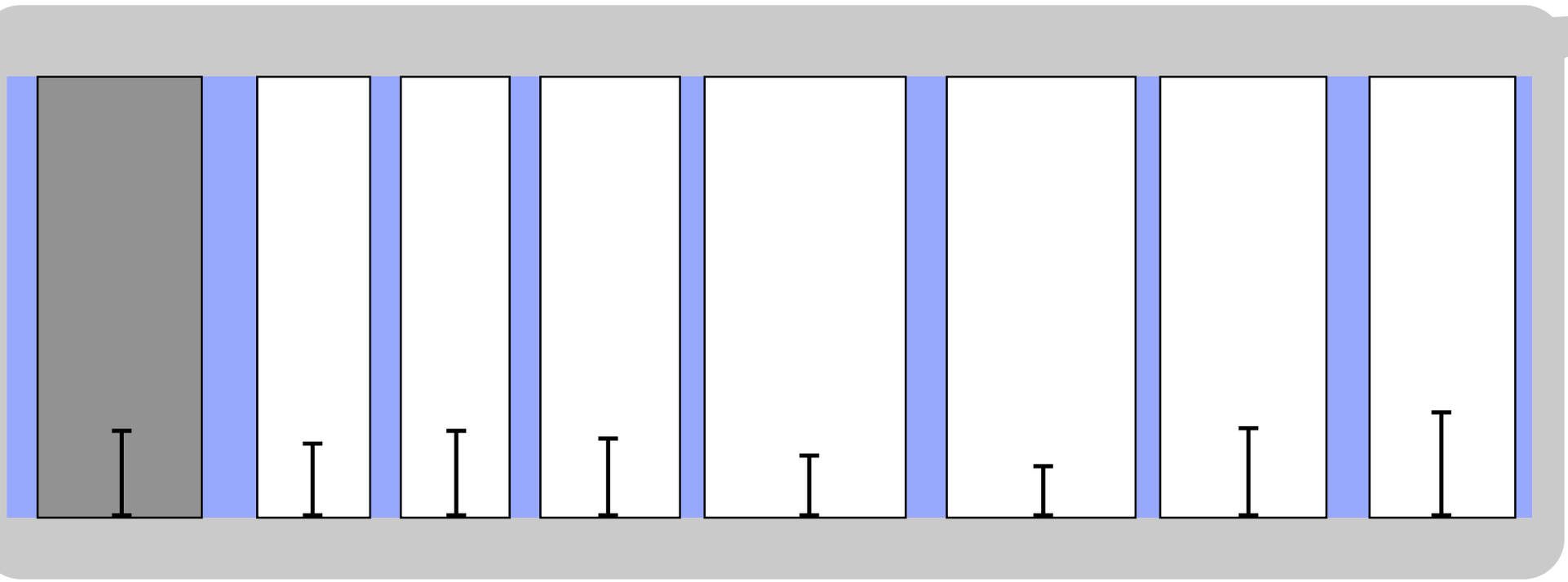
Typesetting Model



Generative Model

p r i s o n e r

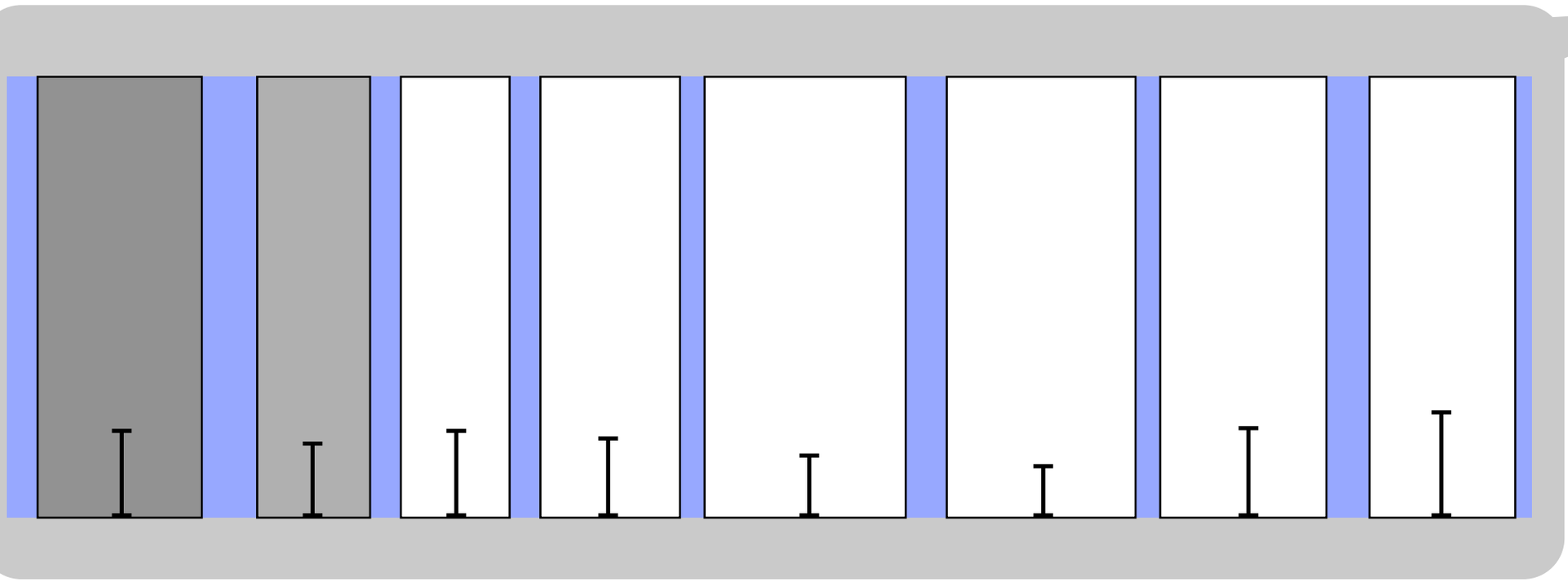
Typesetting Model



Generative Model

p r i s o n e r

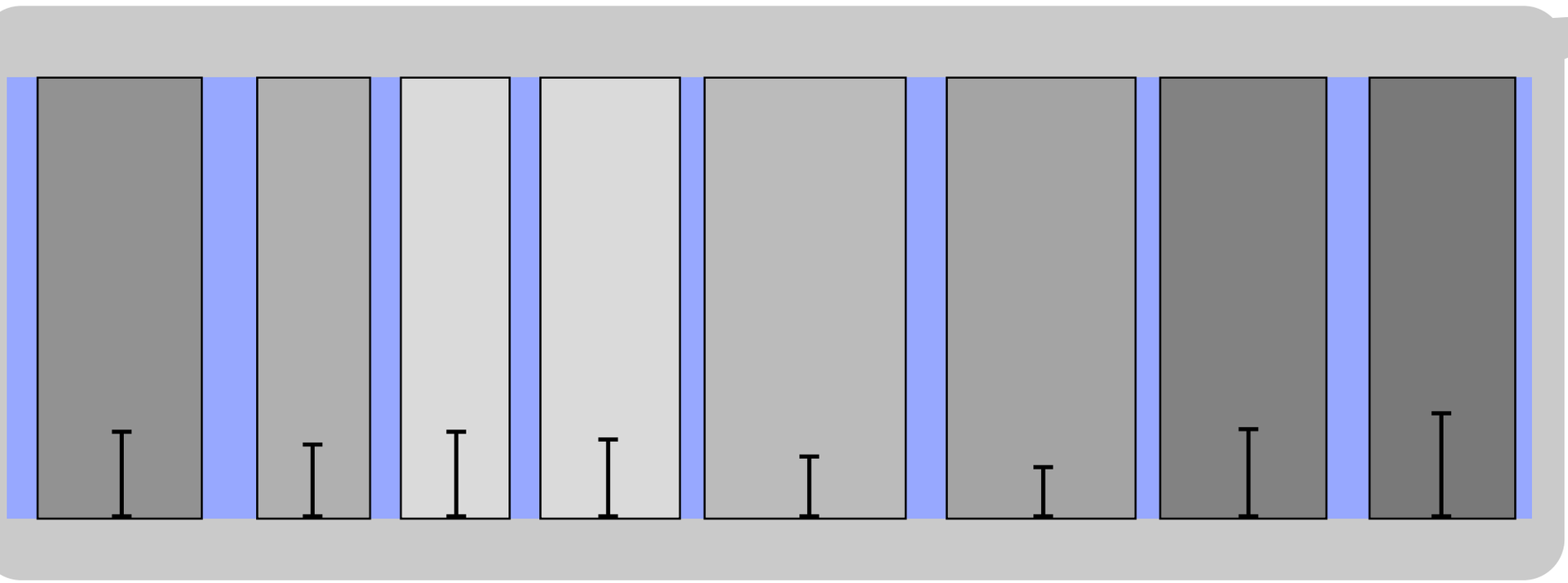
Typesetting Model



Generative Model

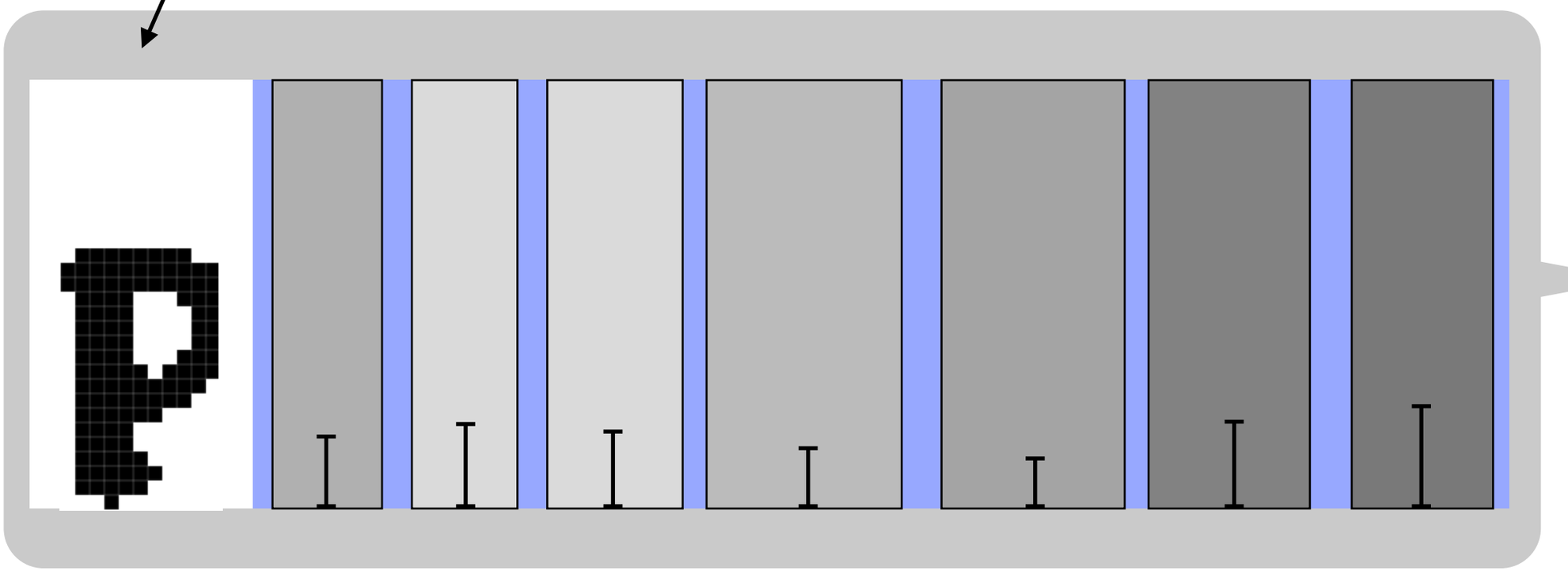
p r i s o n e r

Typesetting Model



Generative Model

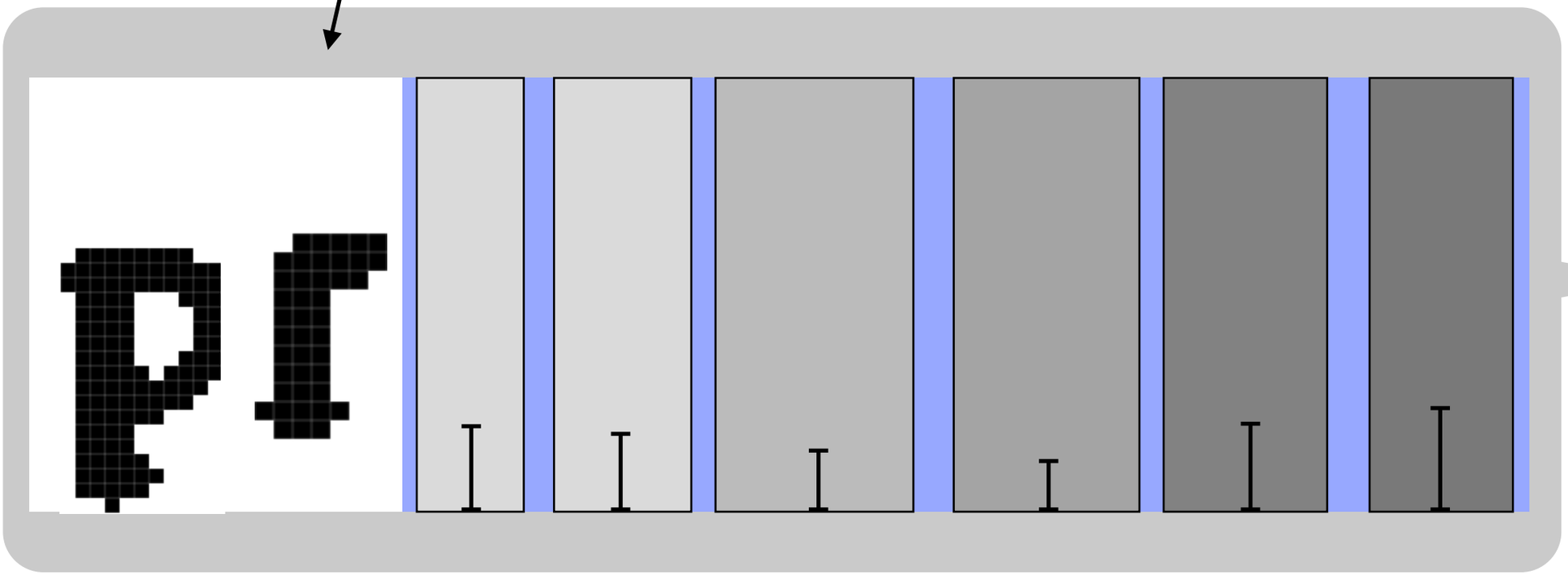
p r i s o n e r



Rendering Model

Generative Model

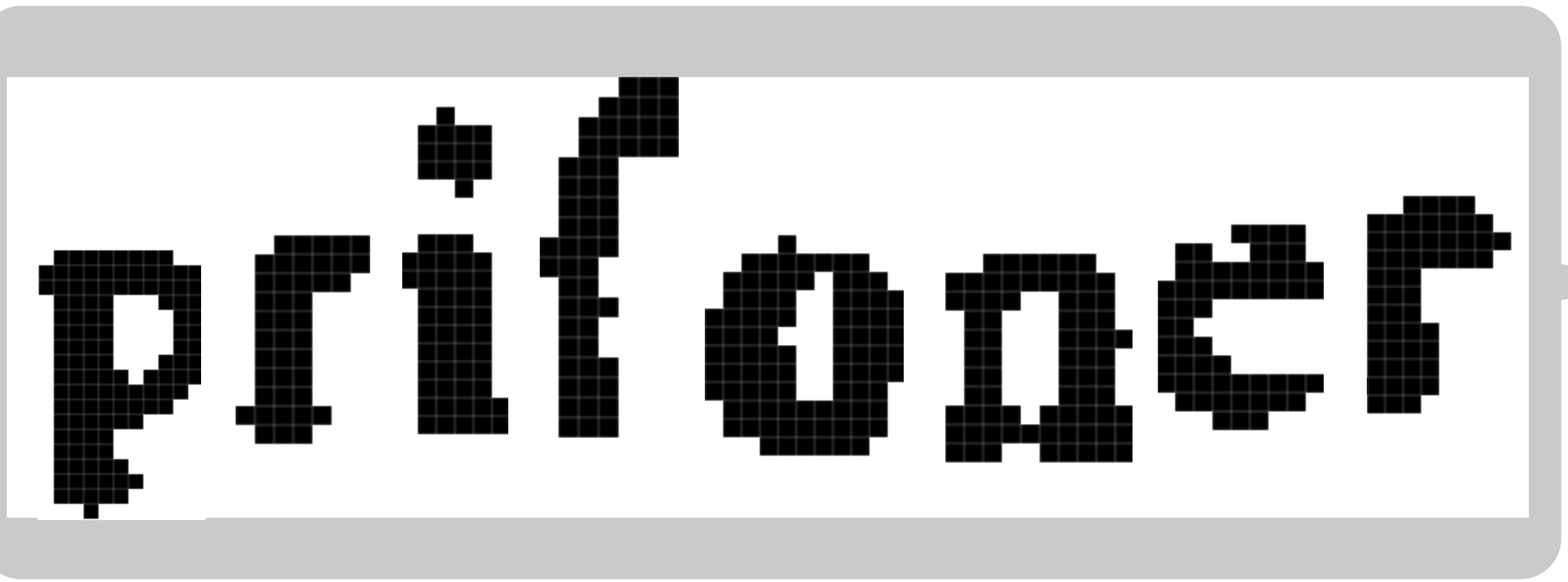
p r i s o n e r



Rendering Model

Generative Model

p r i s o n e r



Rendering Model

Generative Model

Language Model

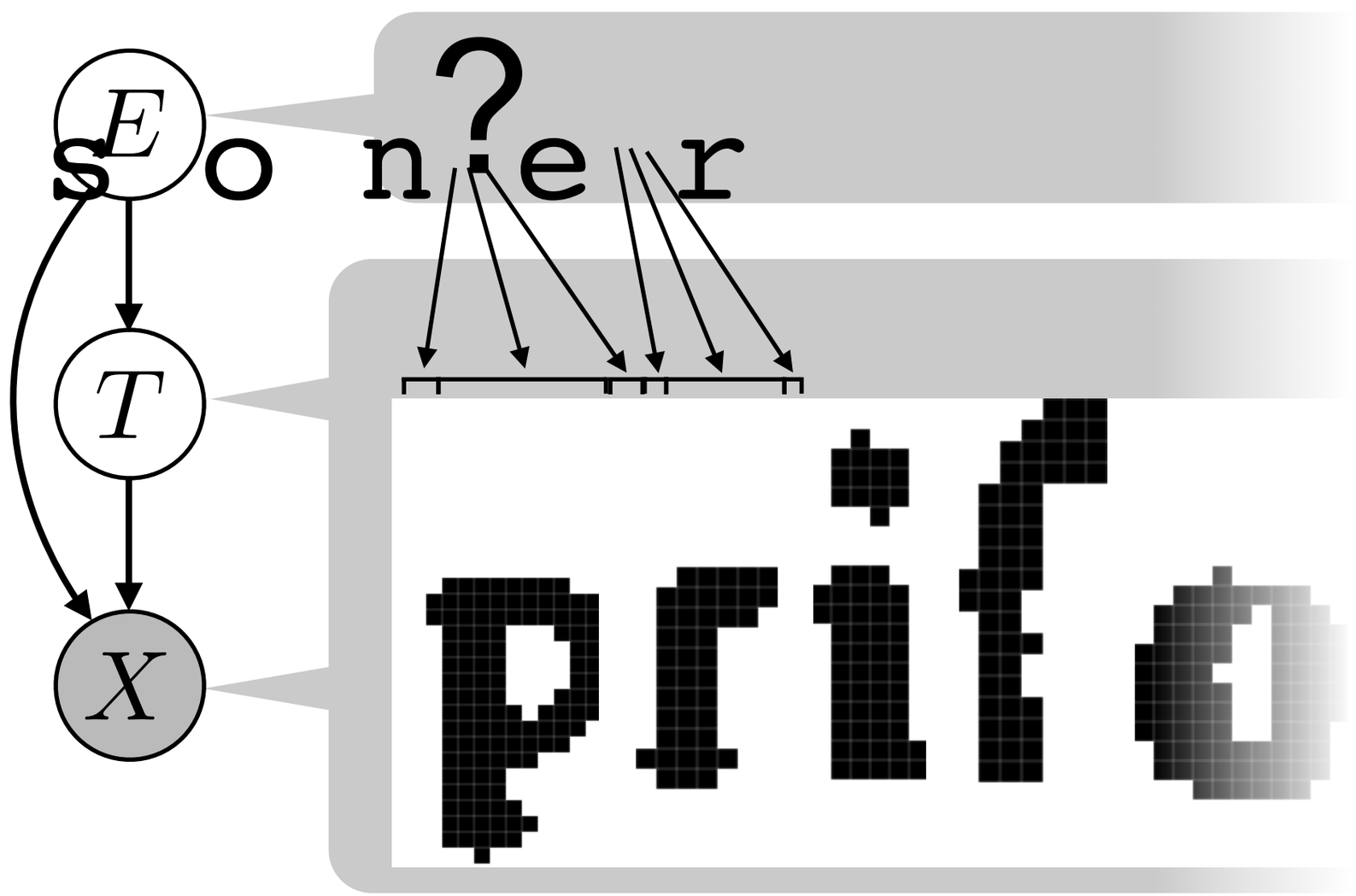
$P(E)$

Typesetting Model

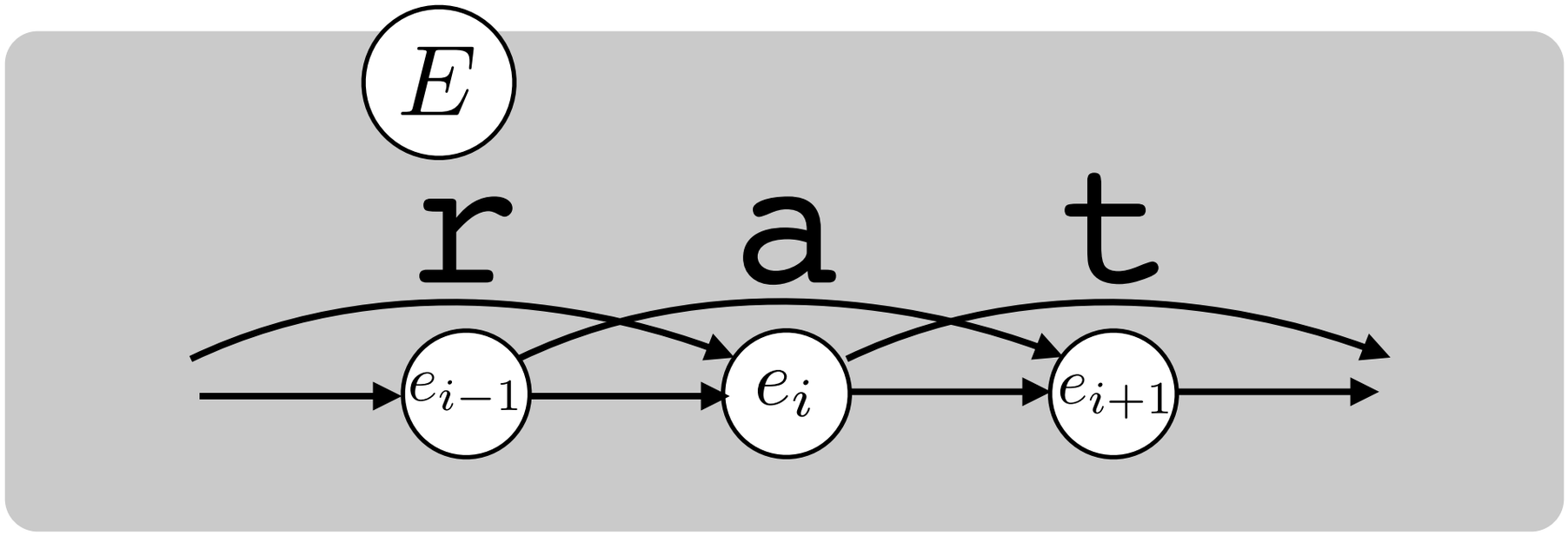
$P(T|E)$

Rendering Model

$P(X|E,T)$

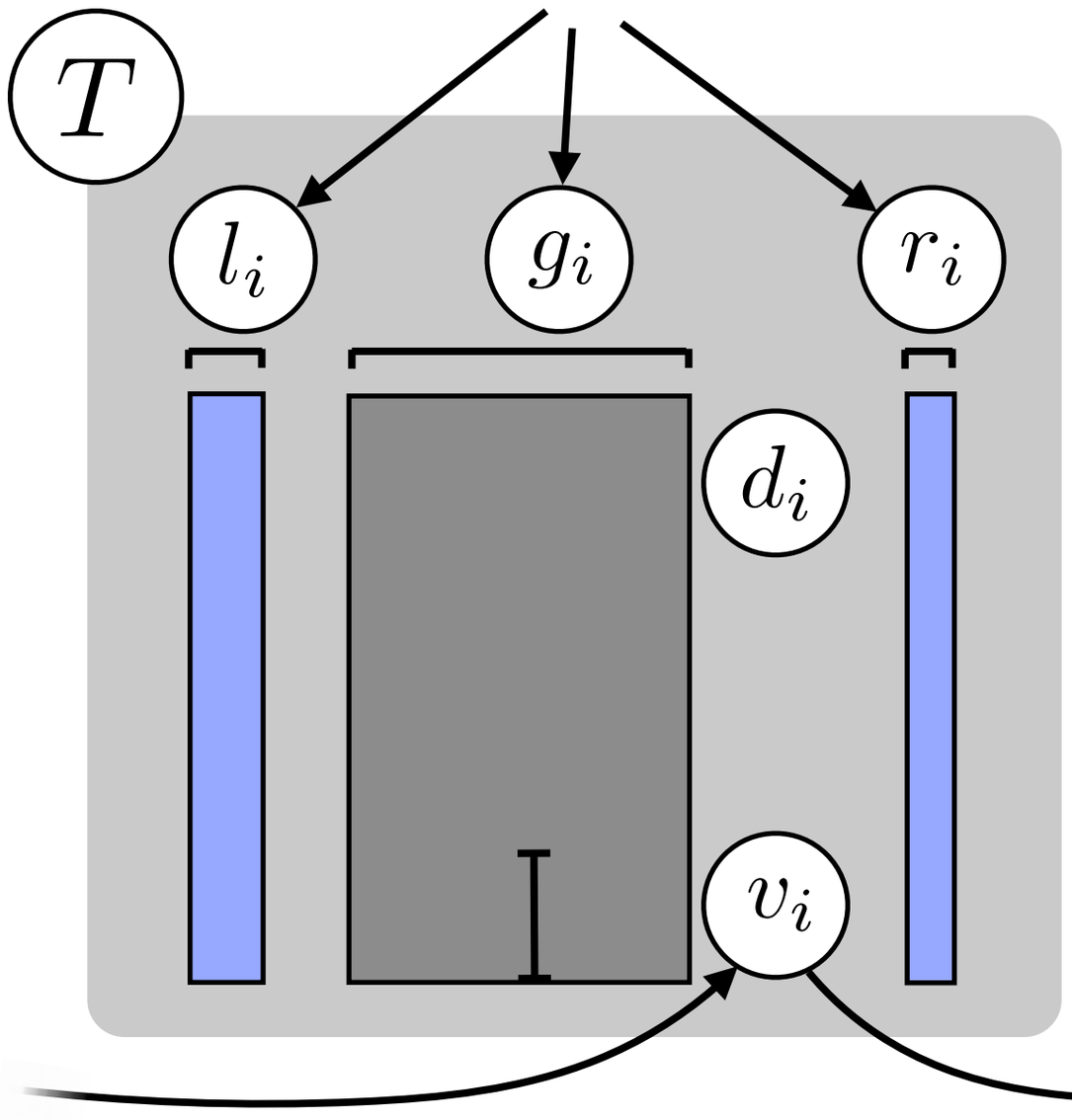


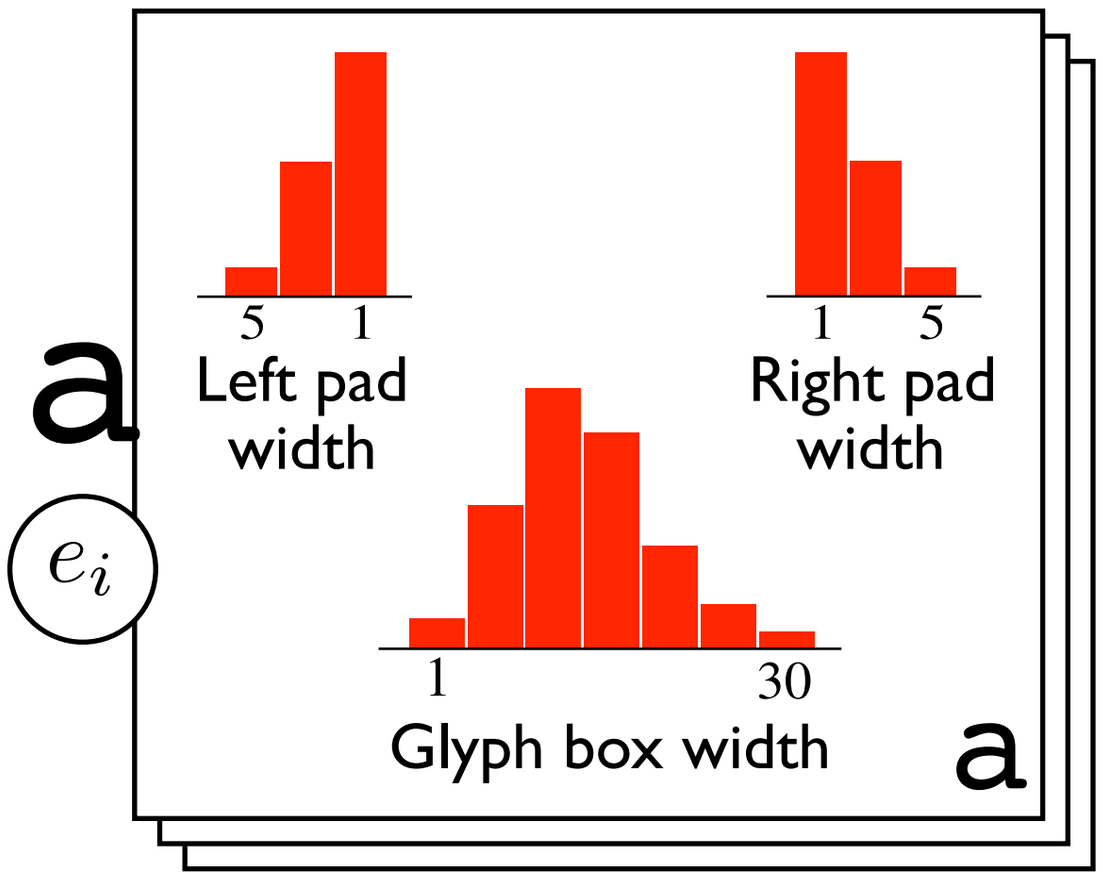
Language Model

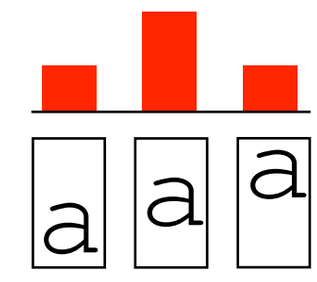


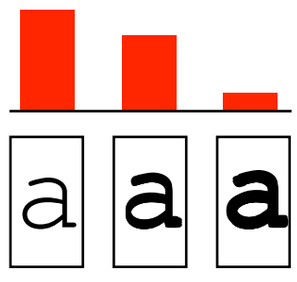
Kneser-Ney smoothed character 6-gram

Typesetting Model

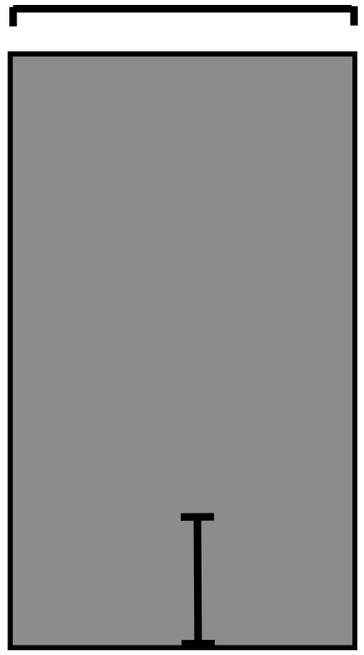




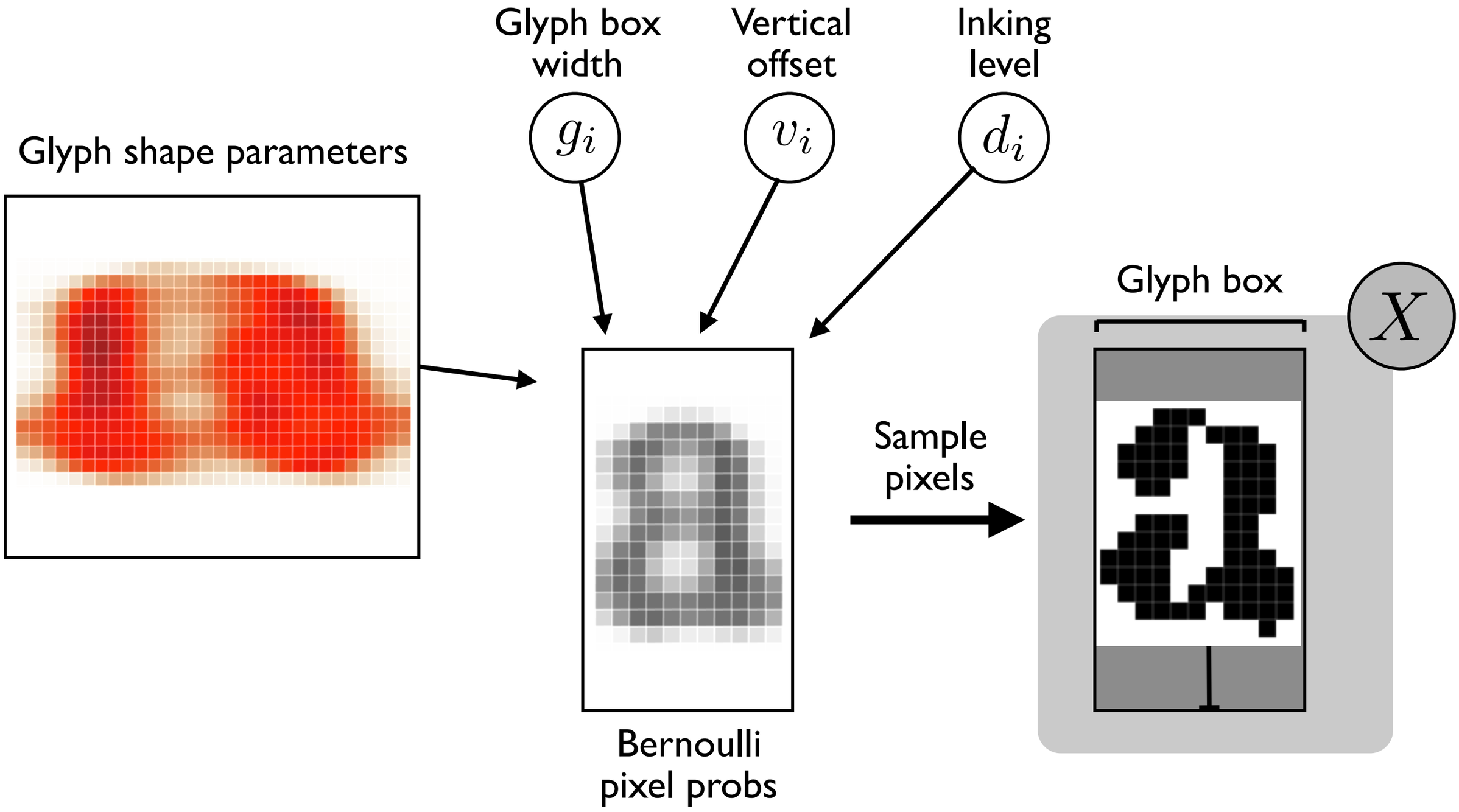




Rendering Model

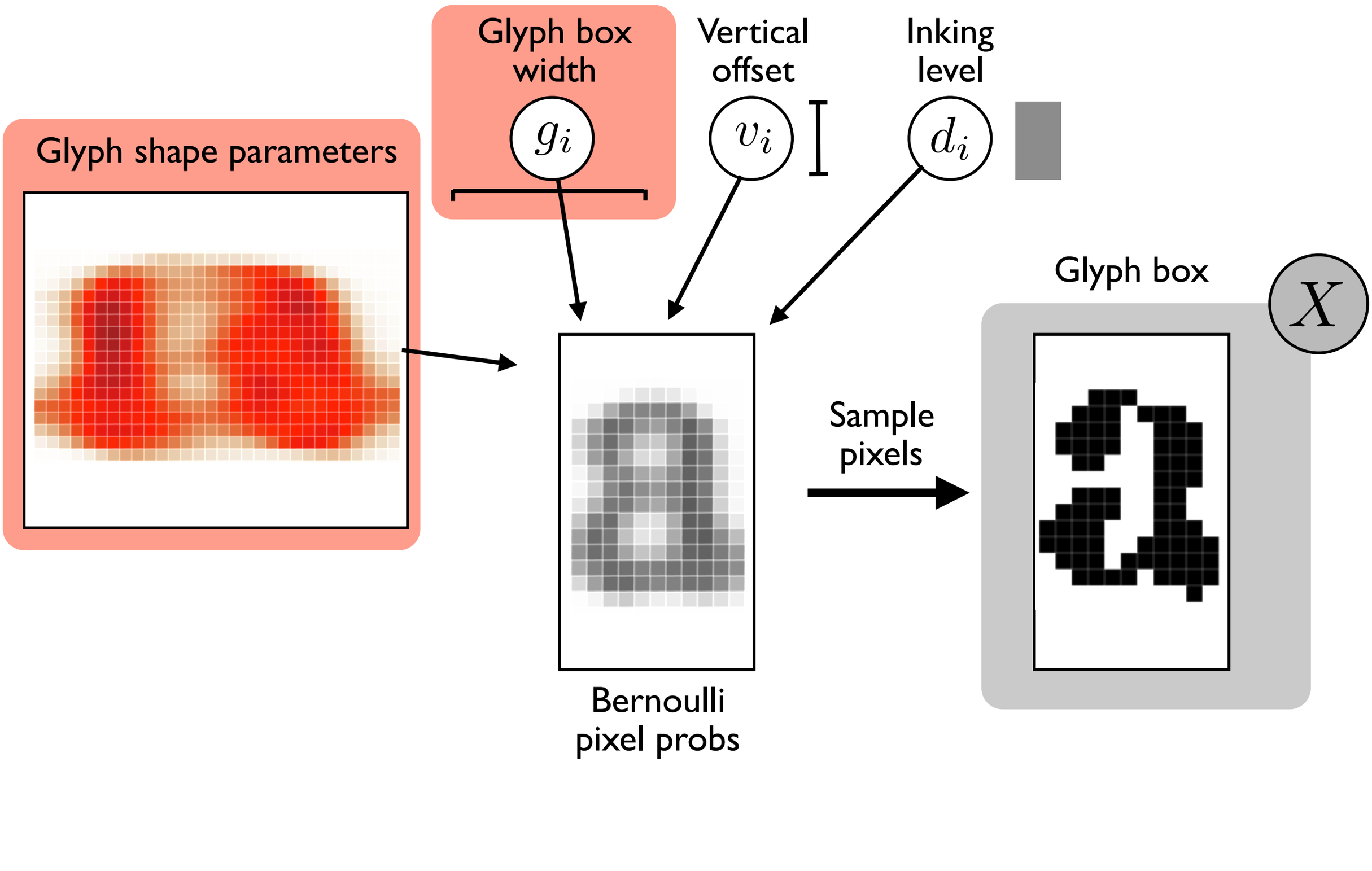


Rendering Model

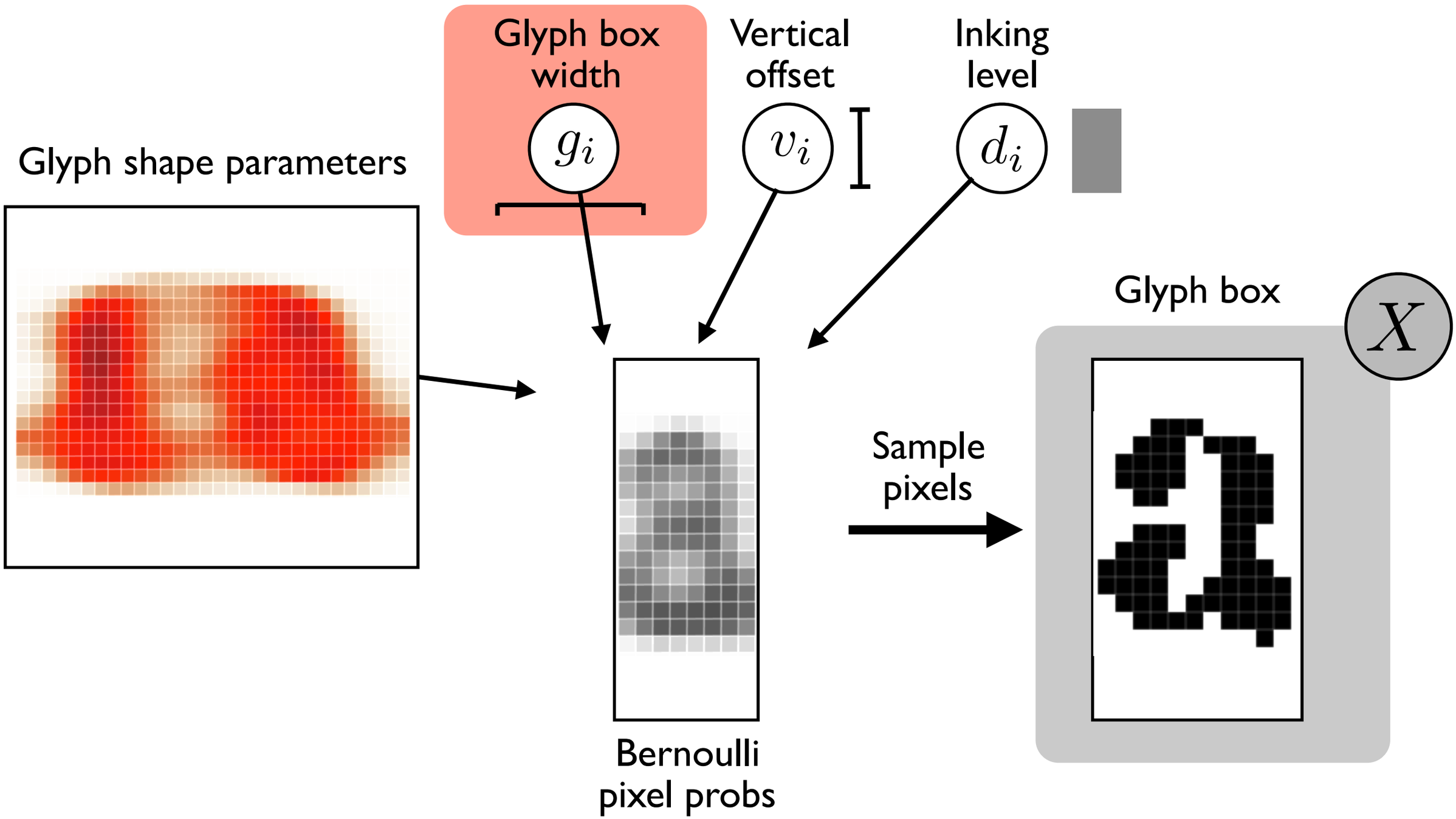


Berkeley N L P

Rendering Model

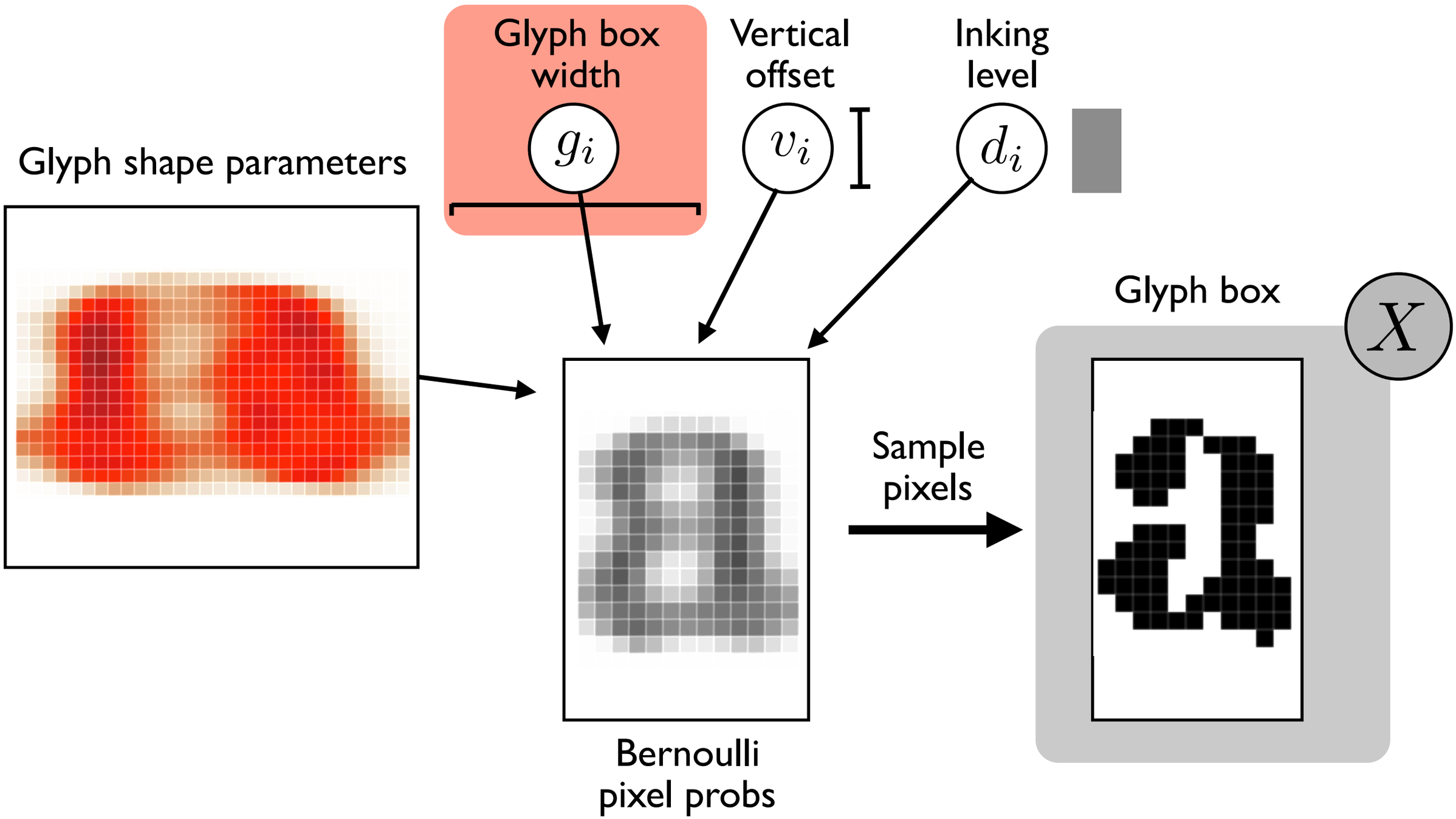


Rendering Model



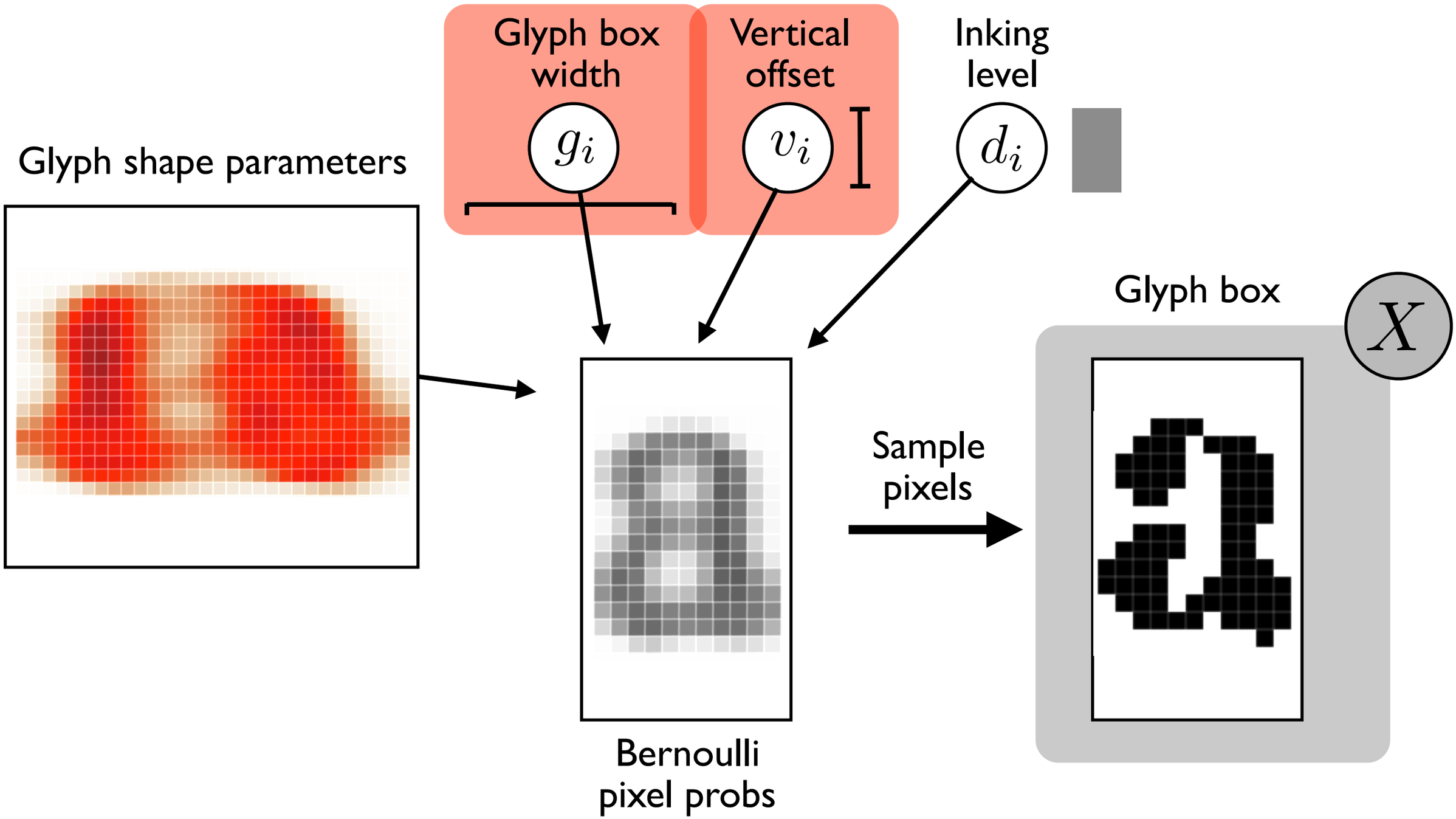
Berkeley N L P

Rendering Model

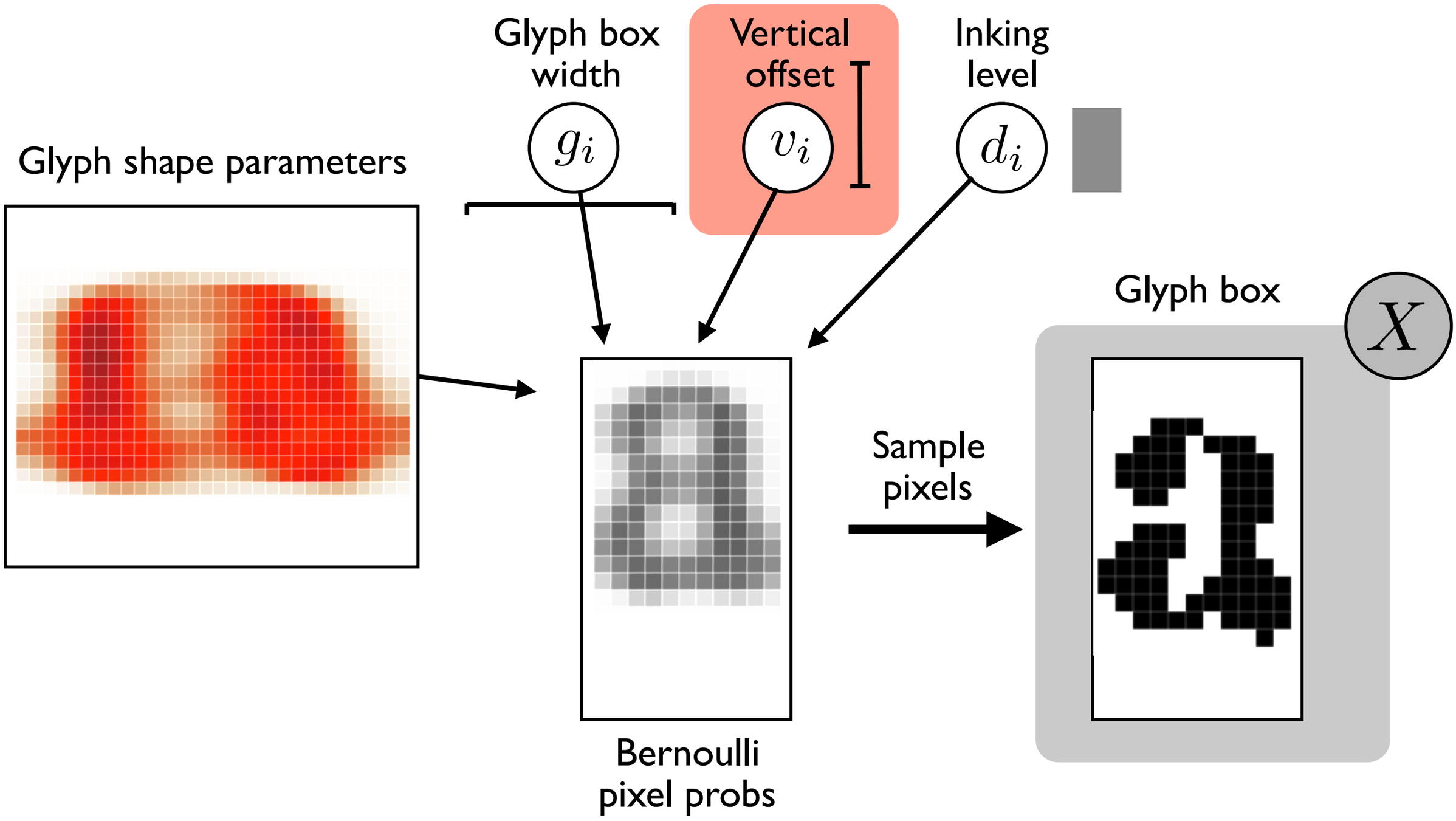


Berkeley N L P

Rendering Model

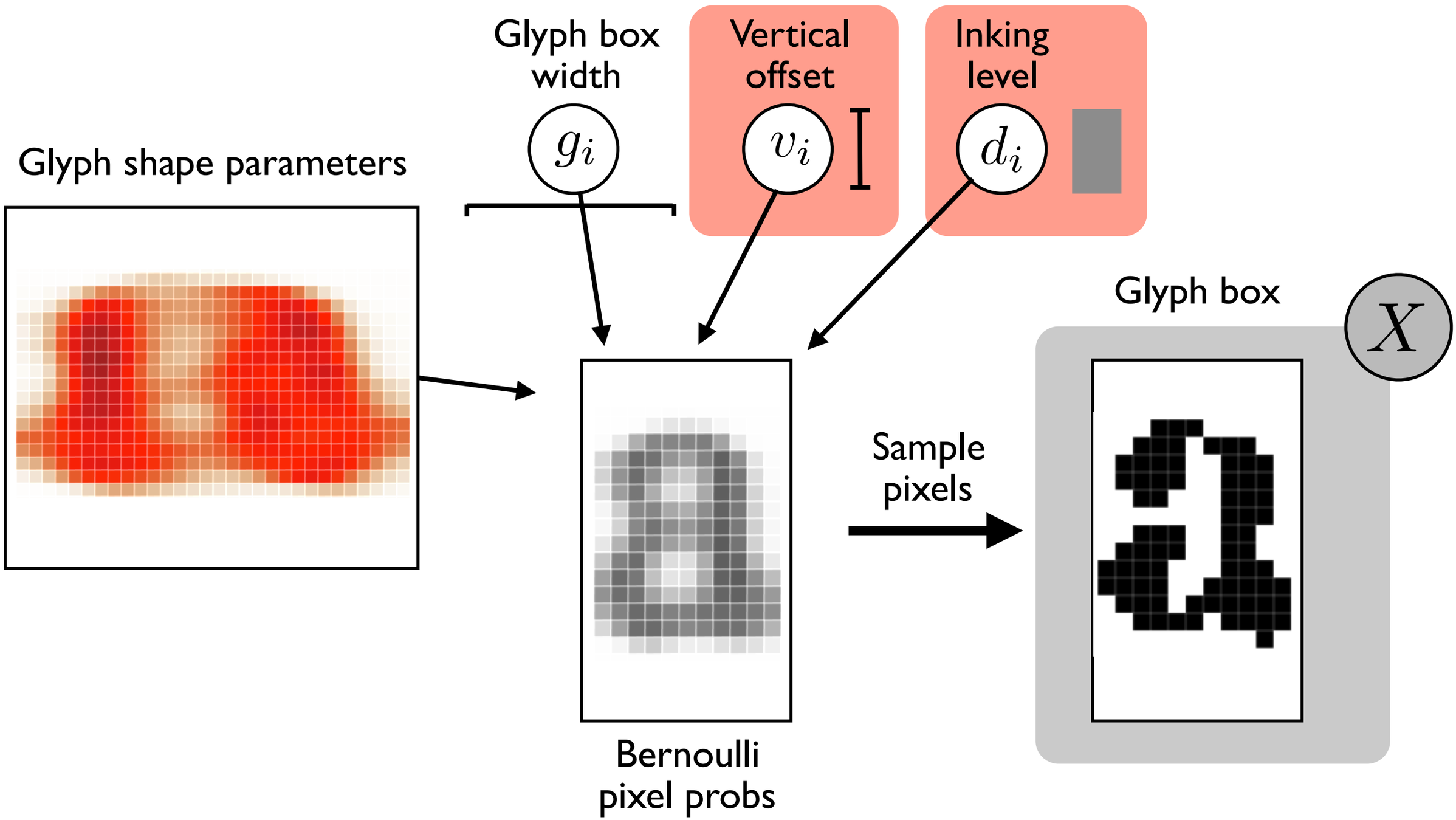


Rendering Model



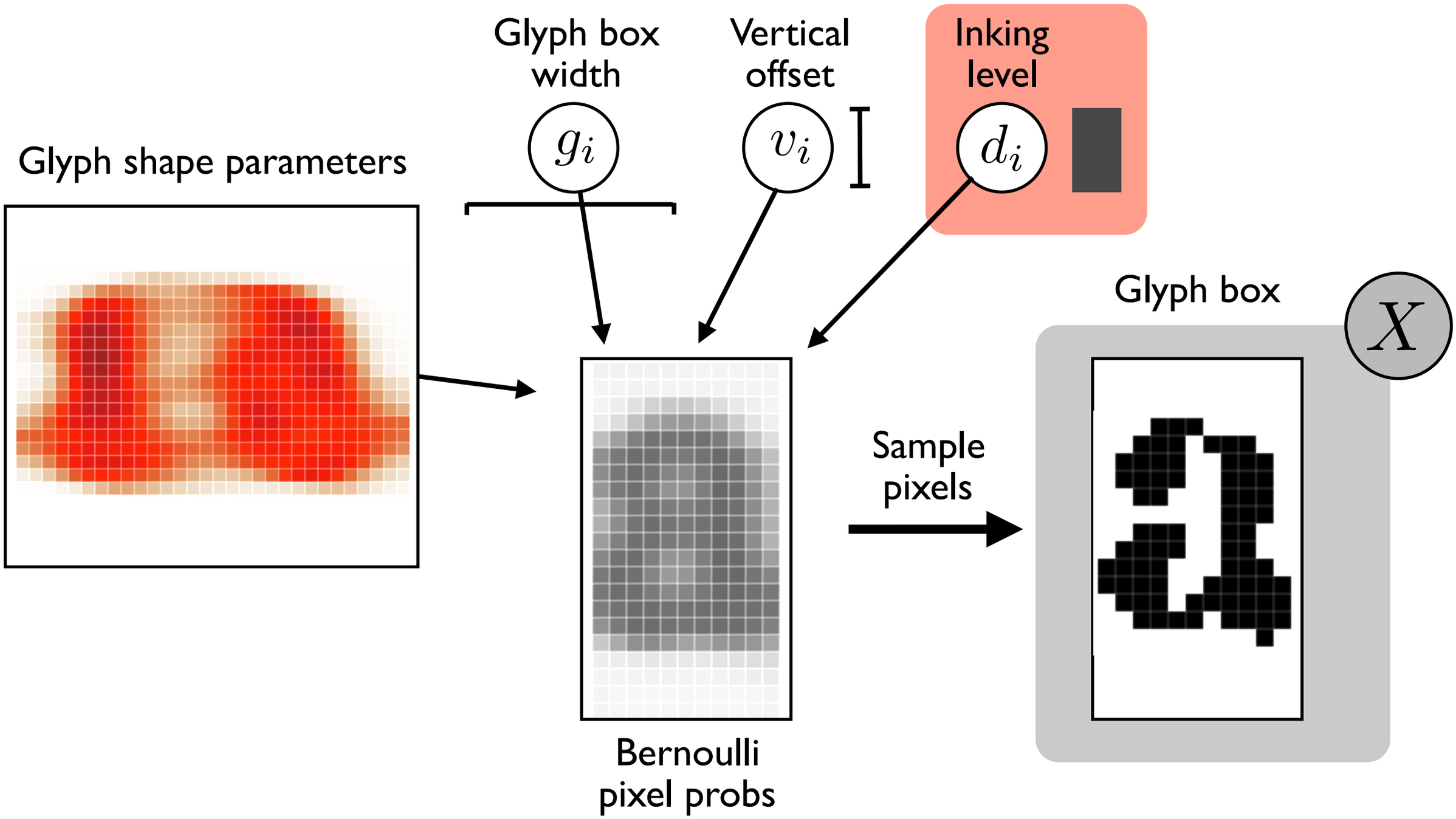
Berkeley N L P

Rendering Model



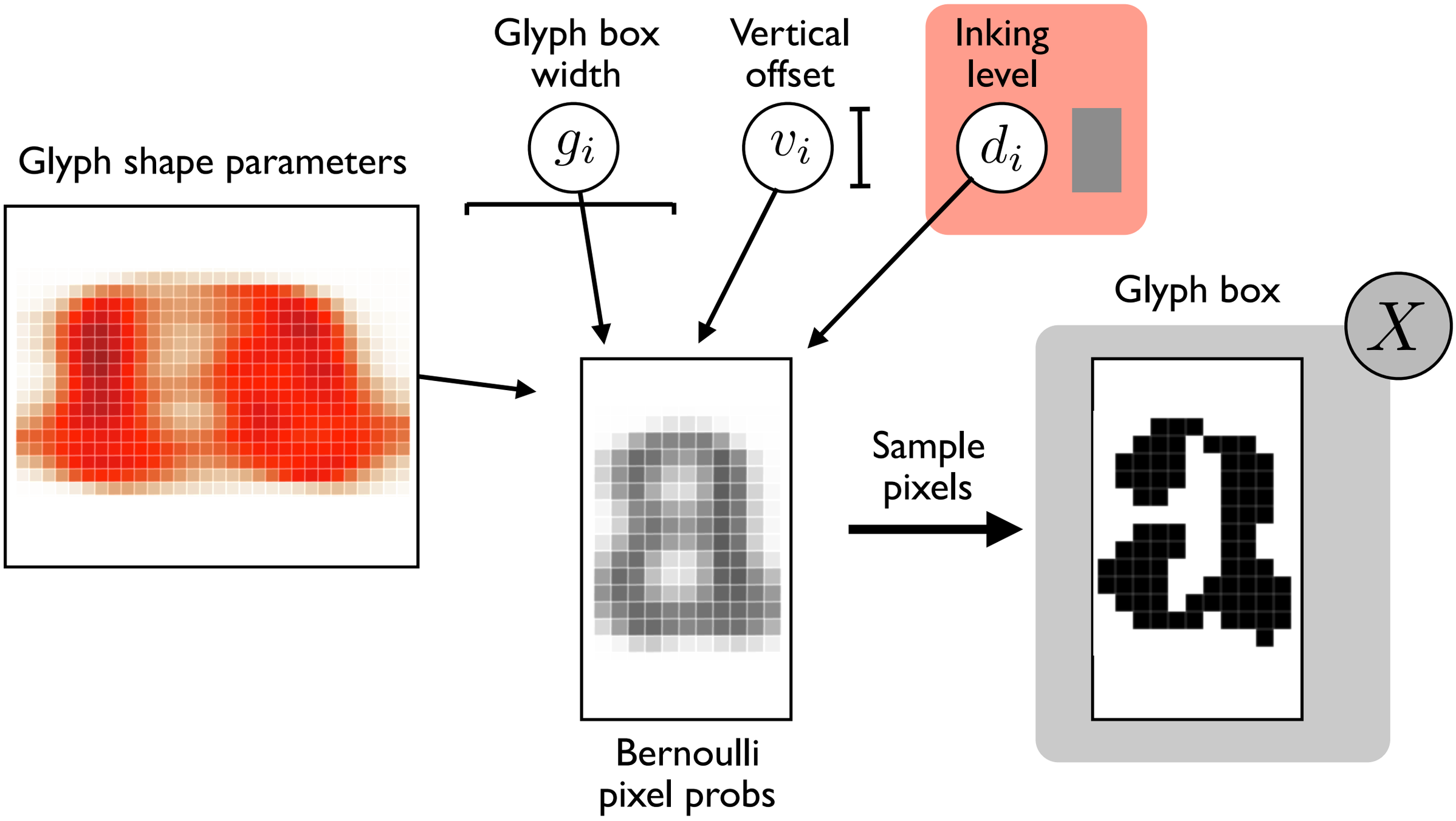
Berkeley N L P

Rendering Model

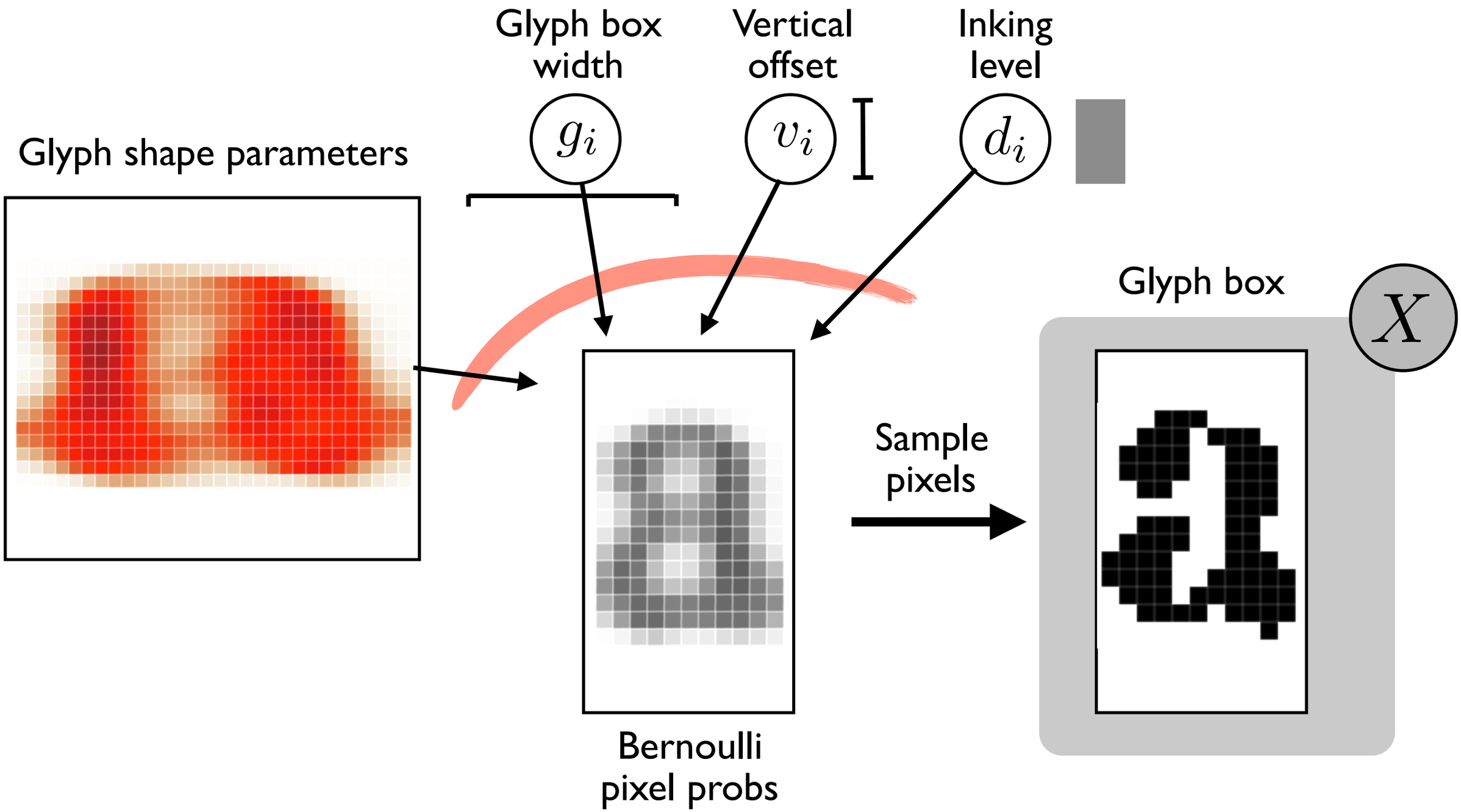


Berkeley N L P

Rendering Model

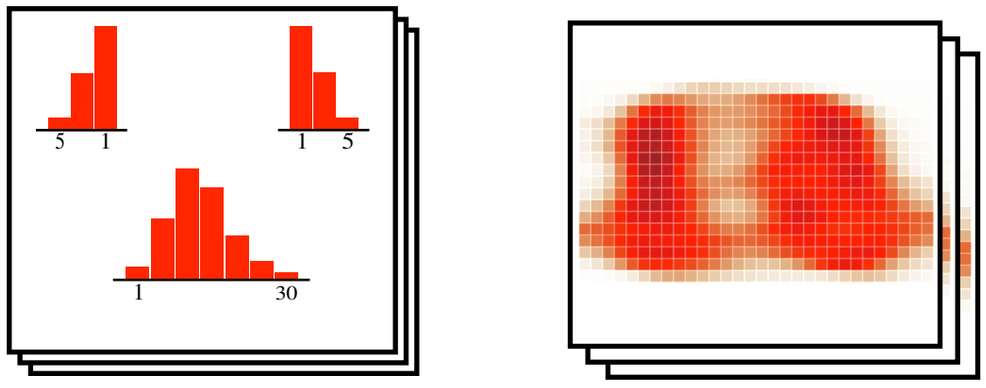


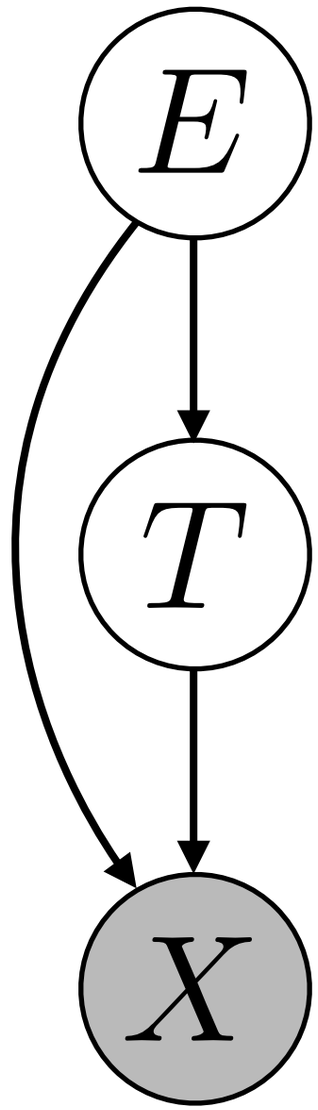
Rendering Model



Learning and Inference

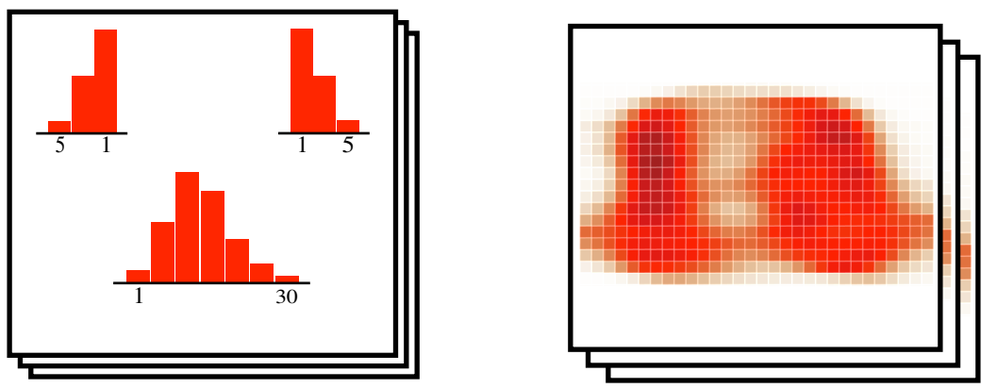
- Learn font parameters using EM



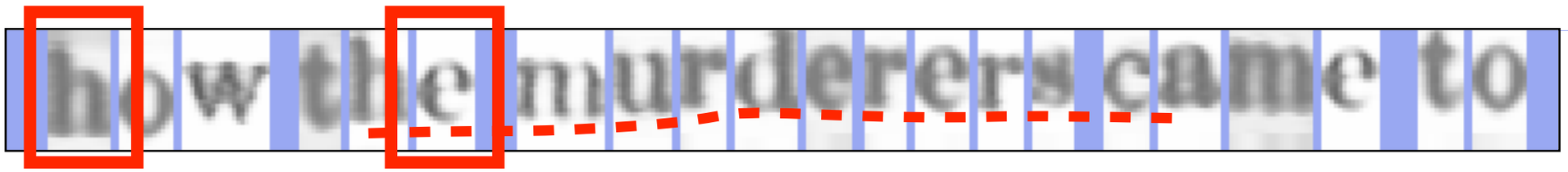


- Initialize font parameters with mixtures of modern fonts
- Semi-Markov DP to compute expectations
- Efficient inference using a coarse-to-fine approach

System Output Example

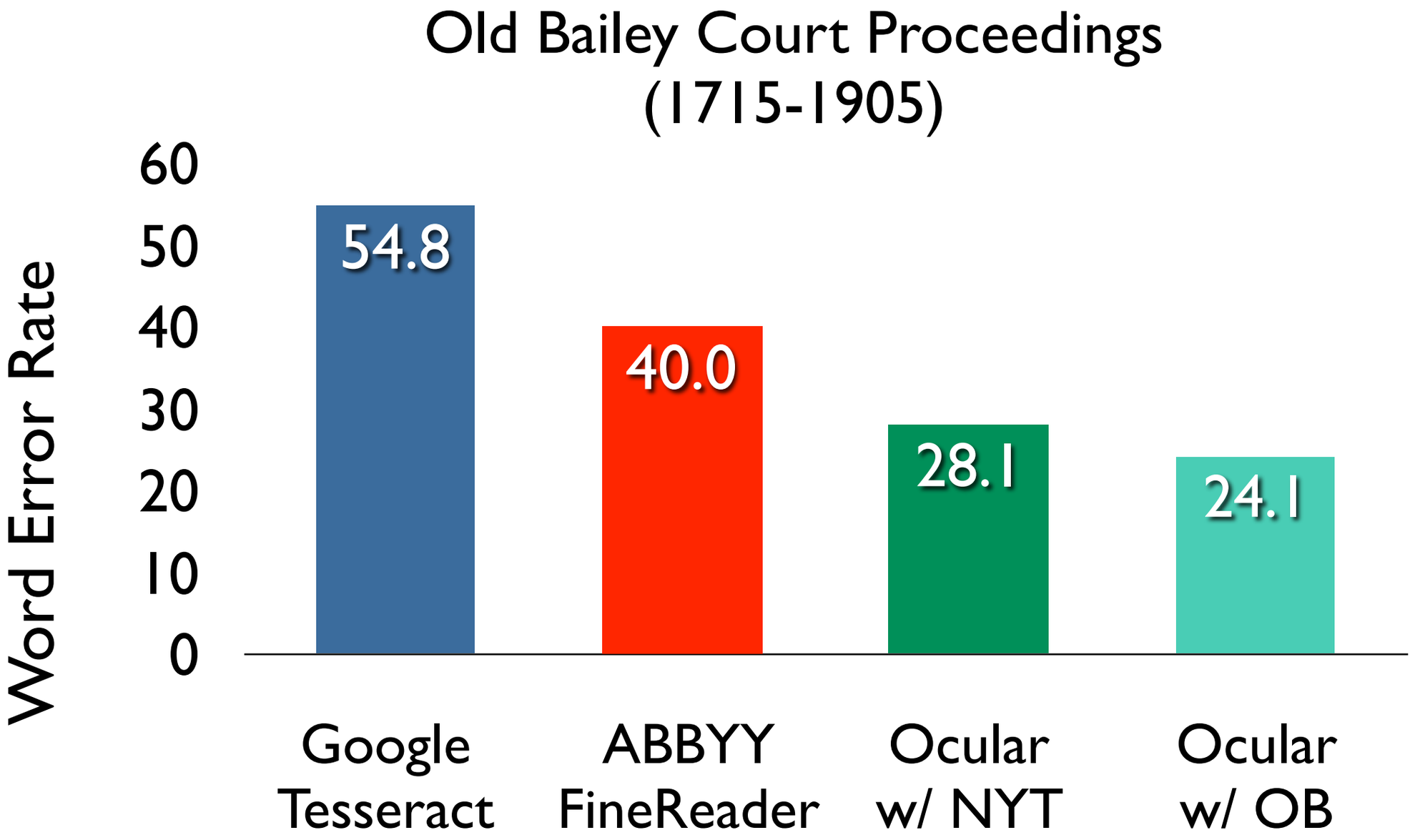


how the murderers came to



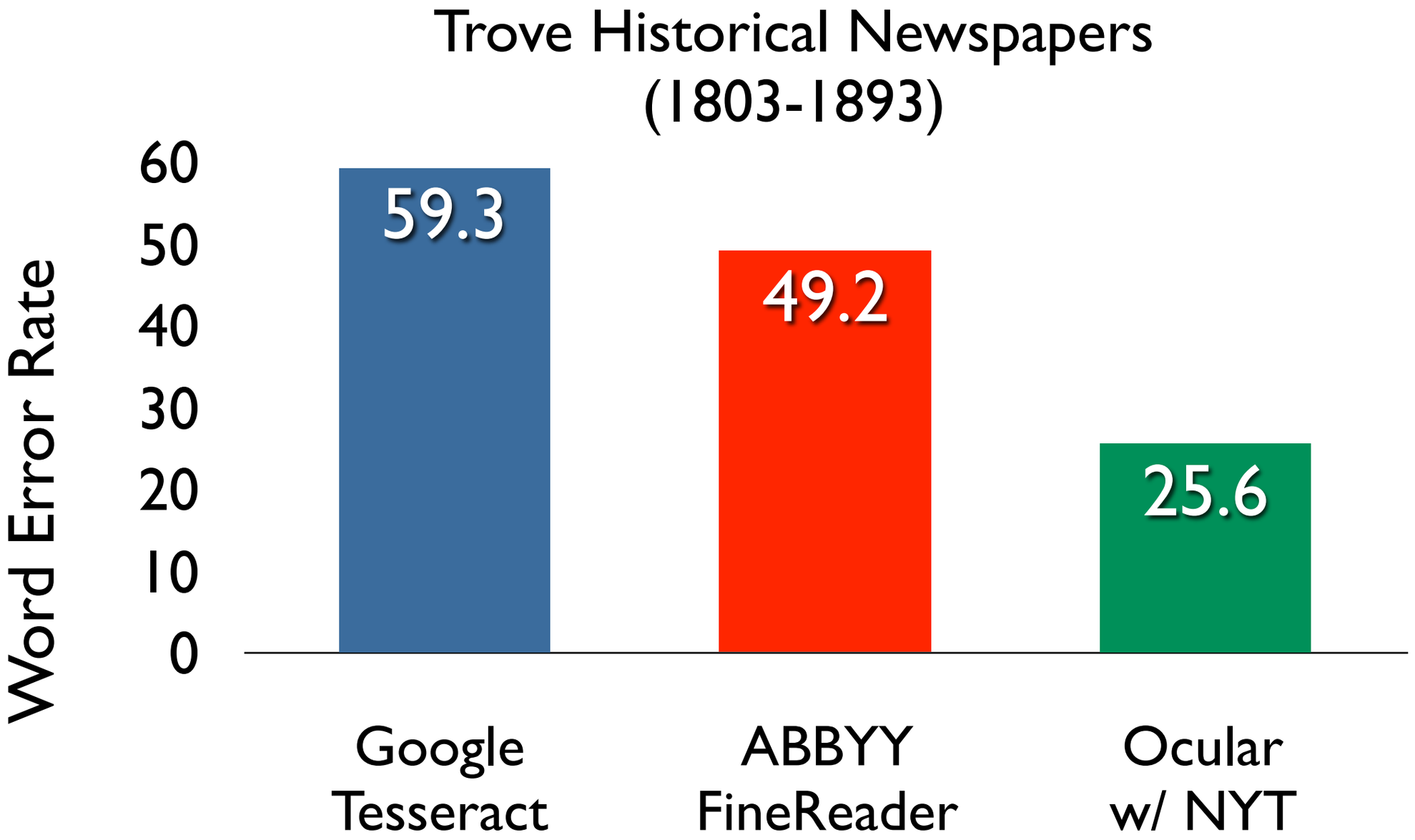
Experiments

Results


Old Bailey Court Proceedings
(1715-1905)
Word Error Rate
60
50
40
30
20
10
0
54.8
40.0
28.1
24.1
Google Tesseract
ABBYY FineReader
Ocular w/ NYT
Ocular w/ OB

Berkeley NLP

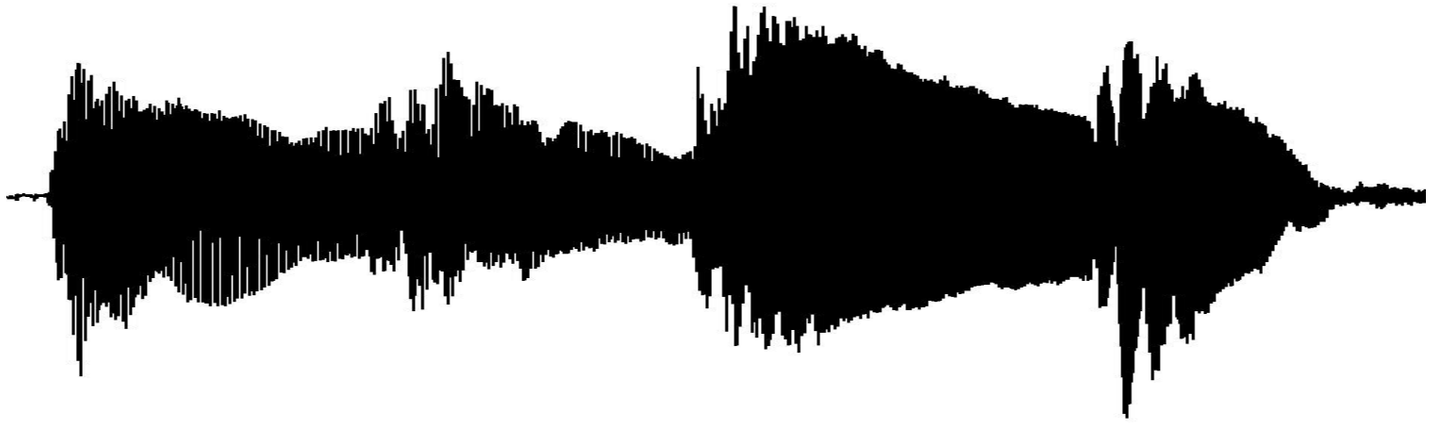
Results

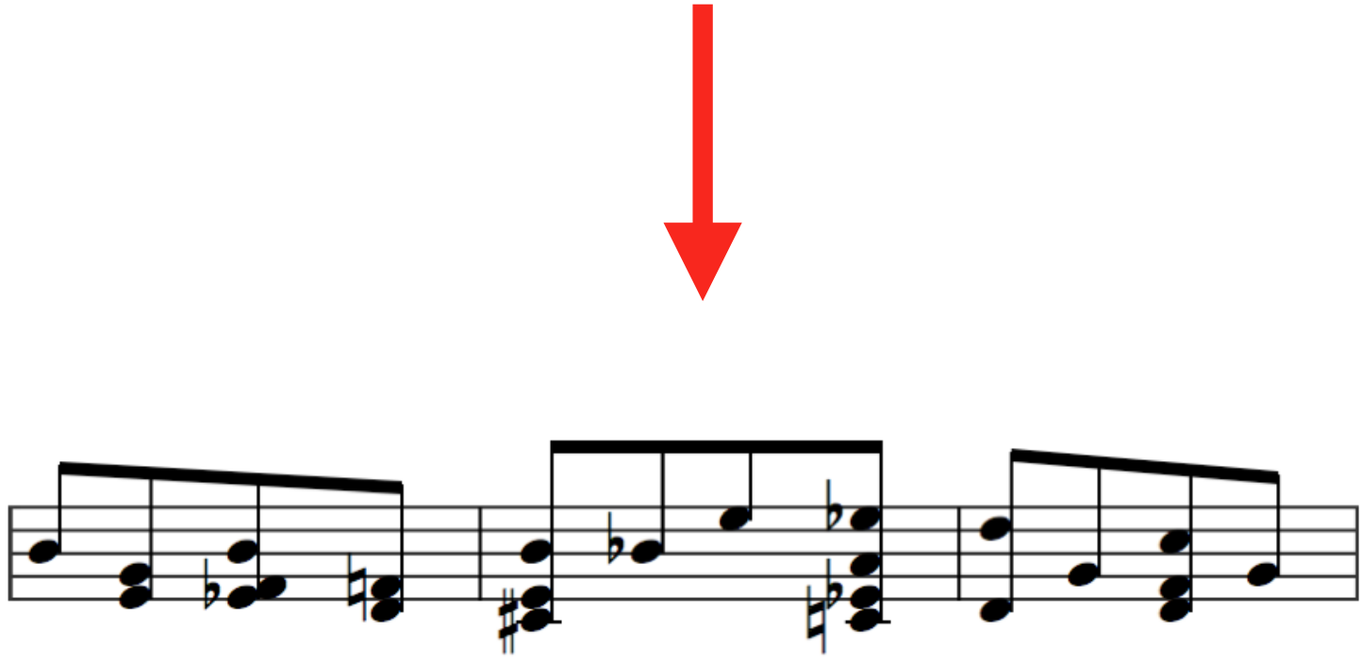


Demo!

Piano Music Transcription

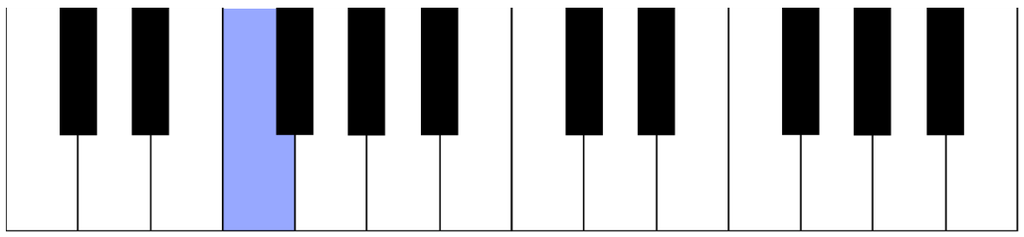
Piano Music Transcription

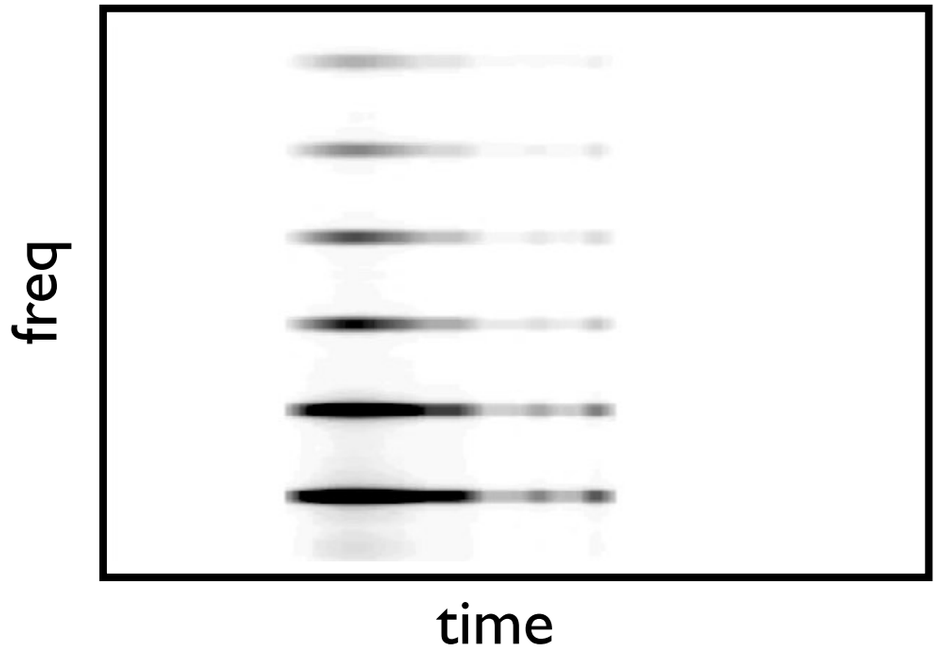




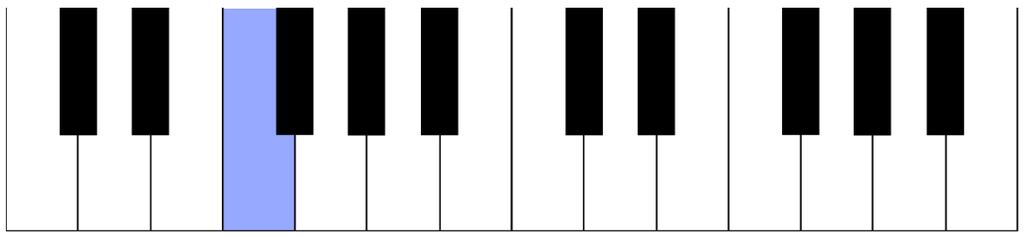
Piano Sounds

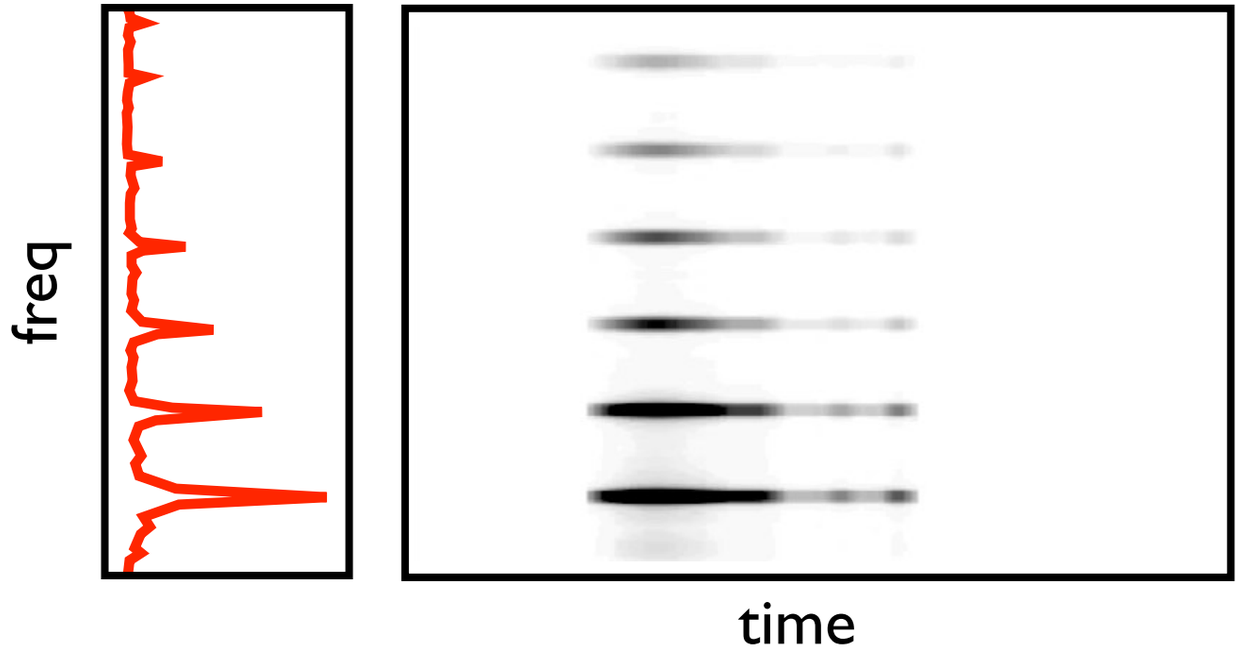
Piano Sounds



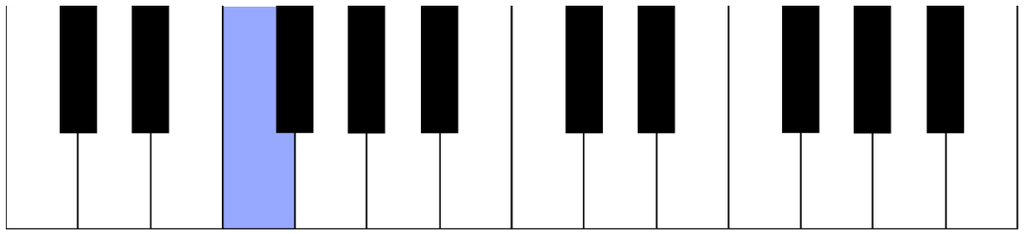


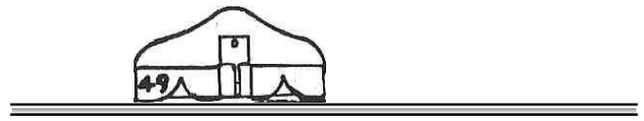
Spectral Shape

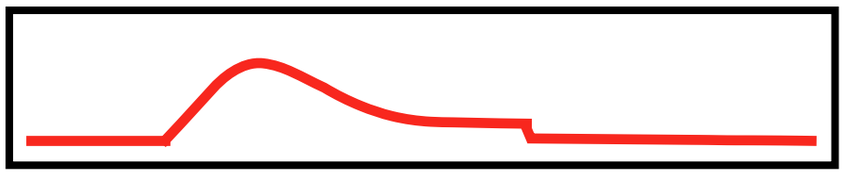


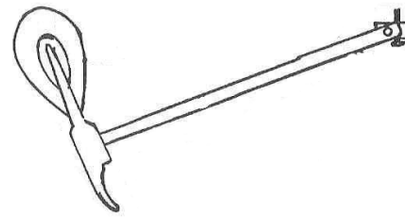


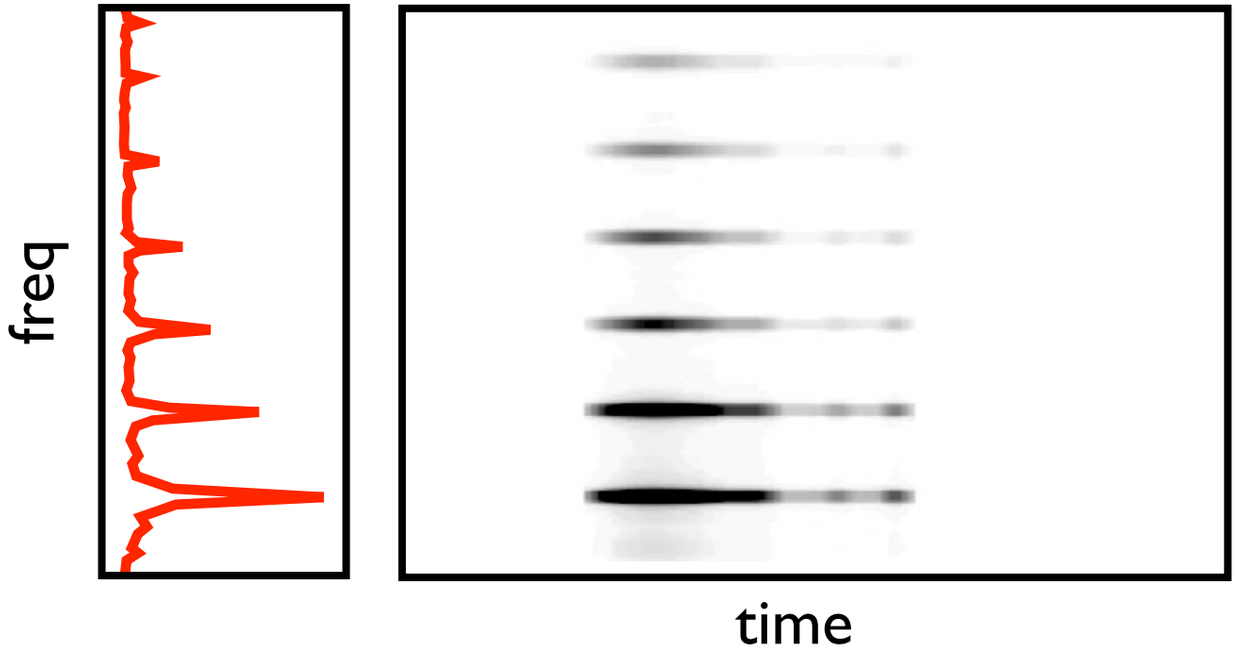
Temporal Shape



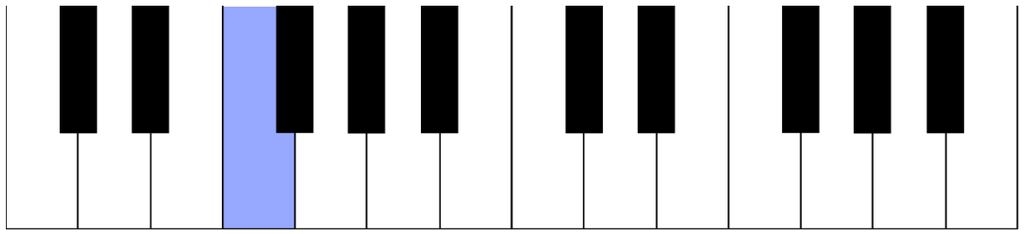


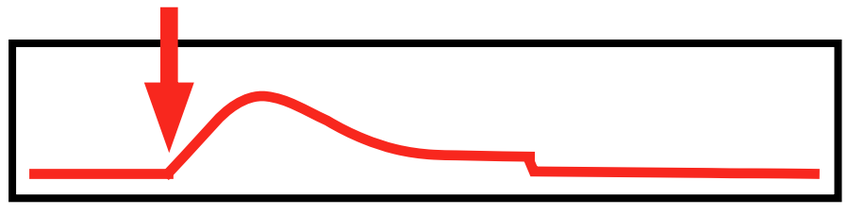


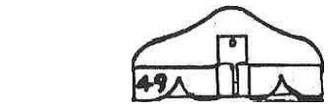


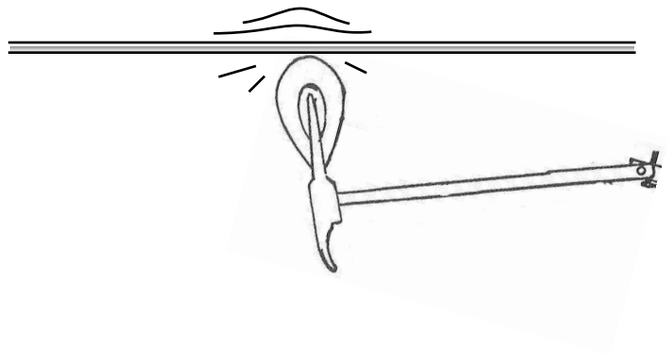


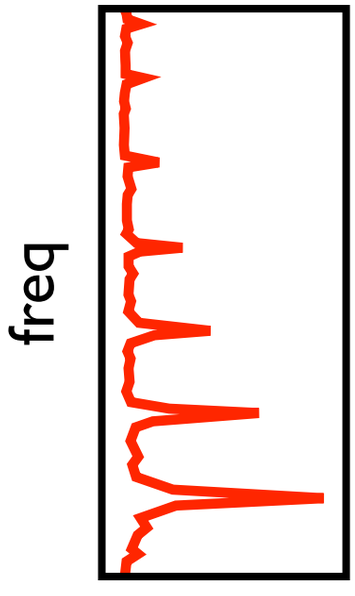
Temporal Shape







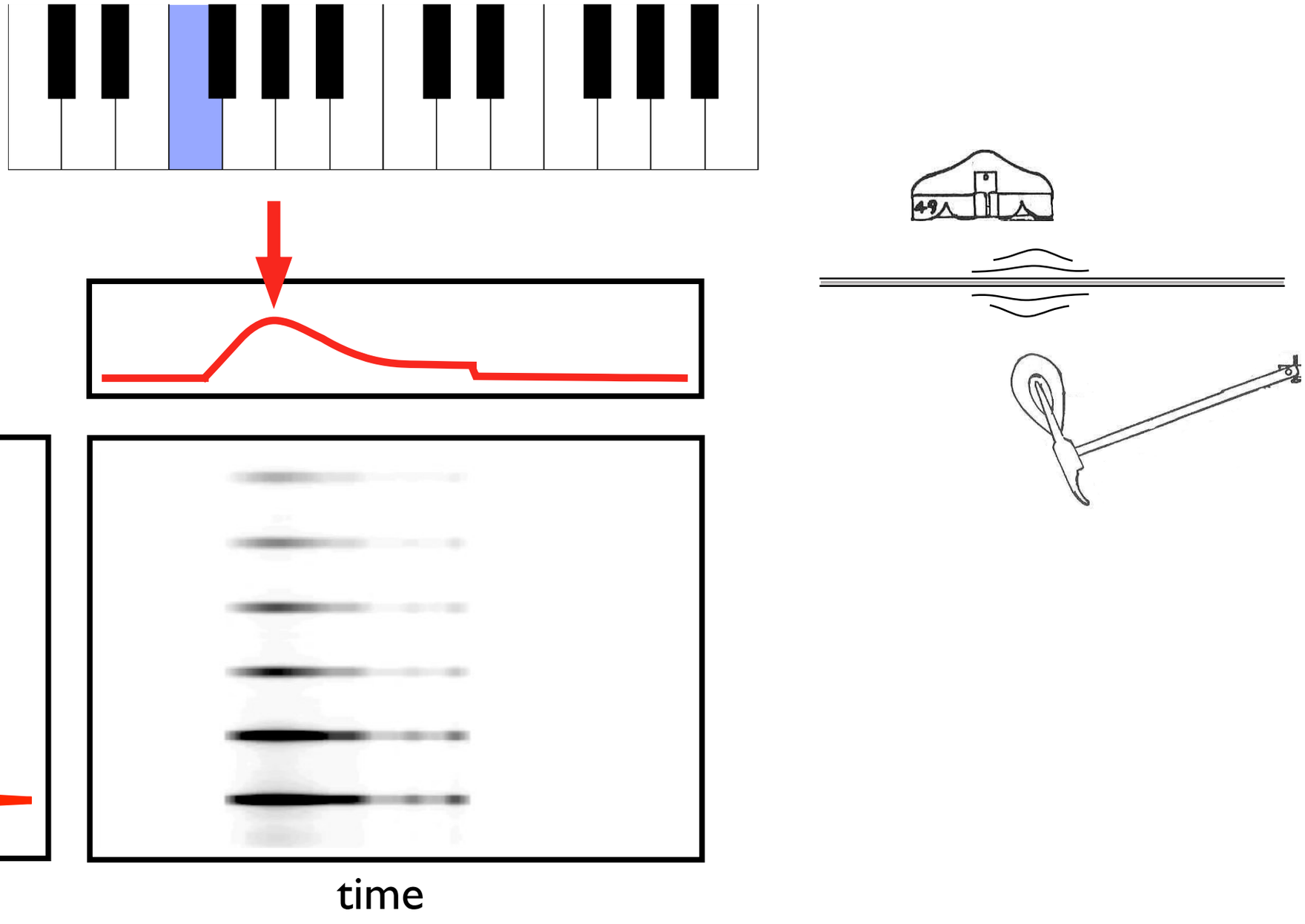




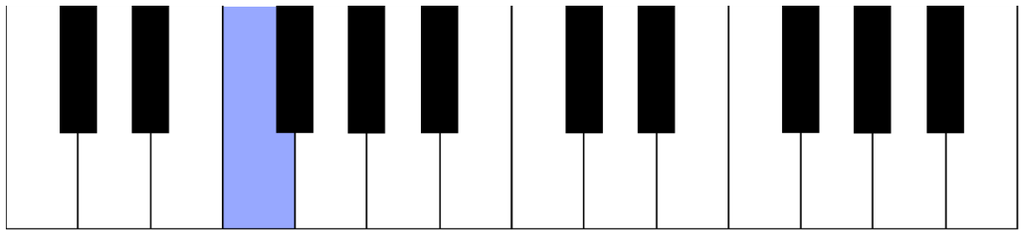
freq

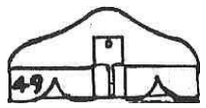
time

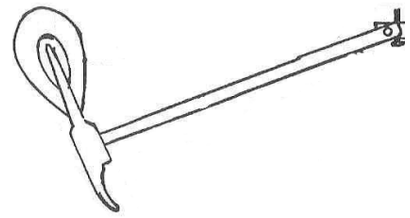
Temporal Shape

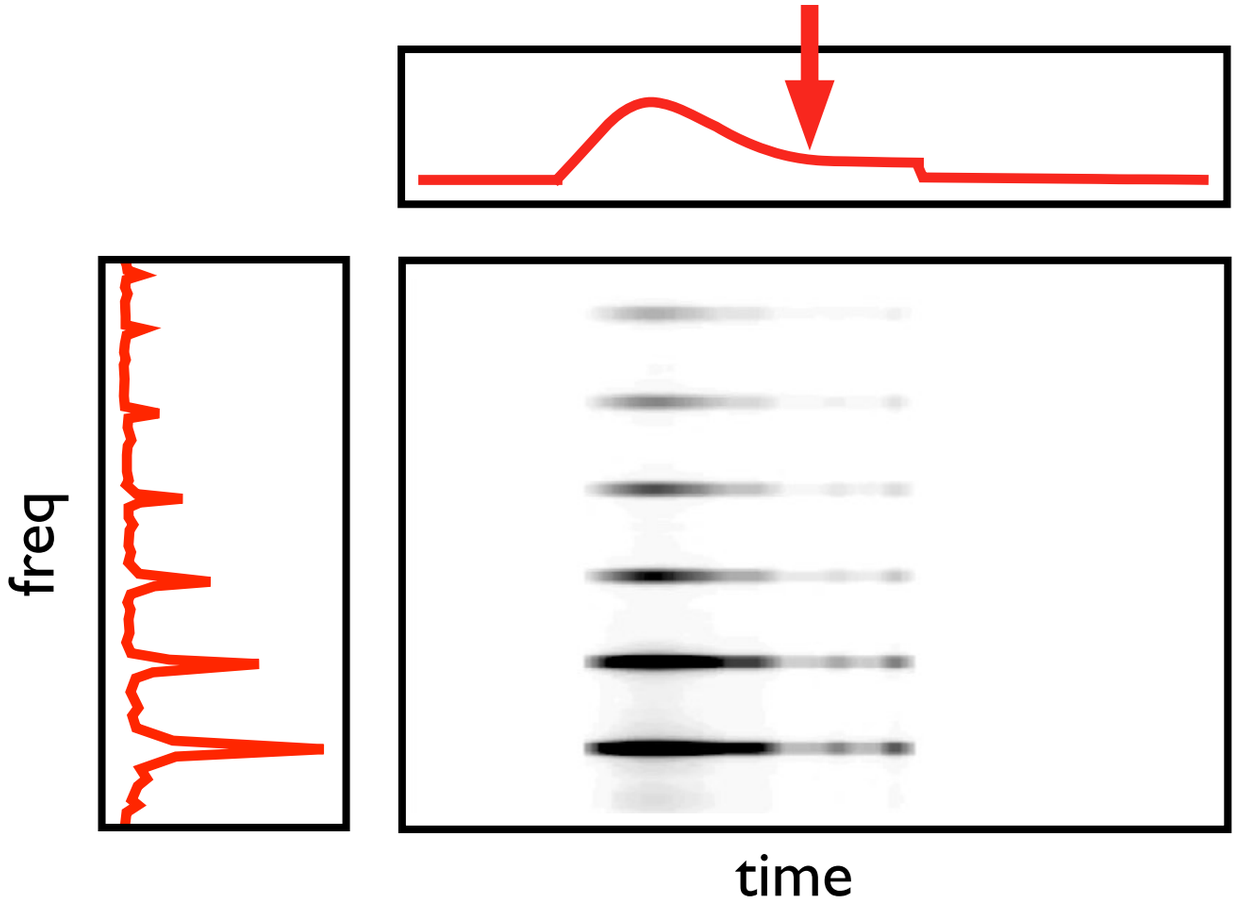


Temporal Shape

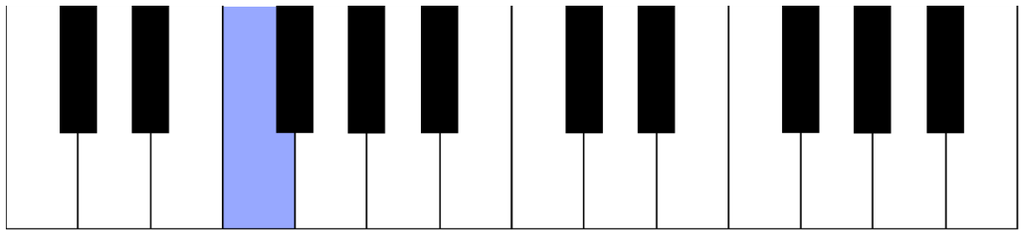


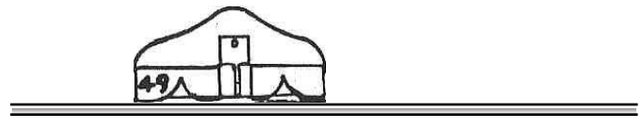


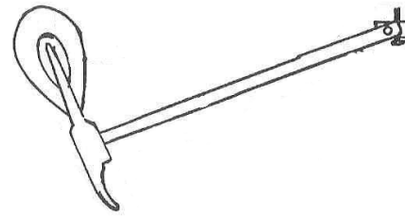


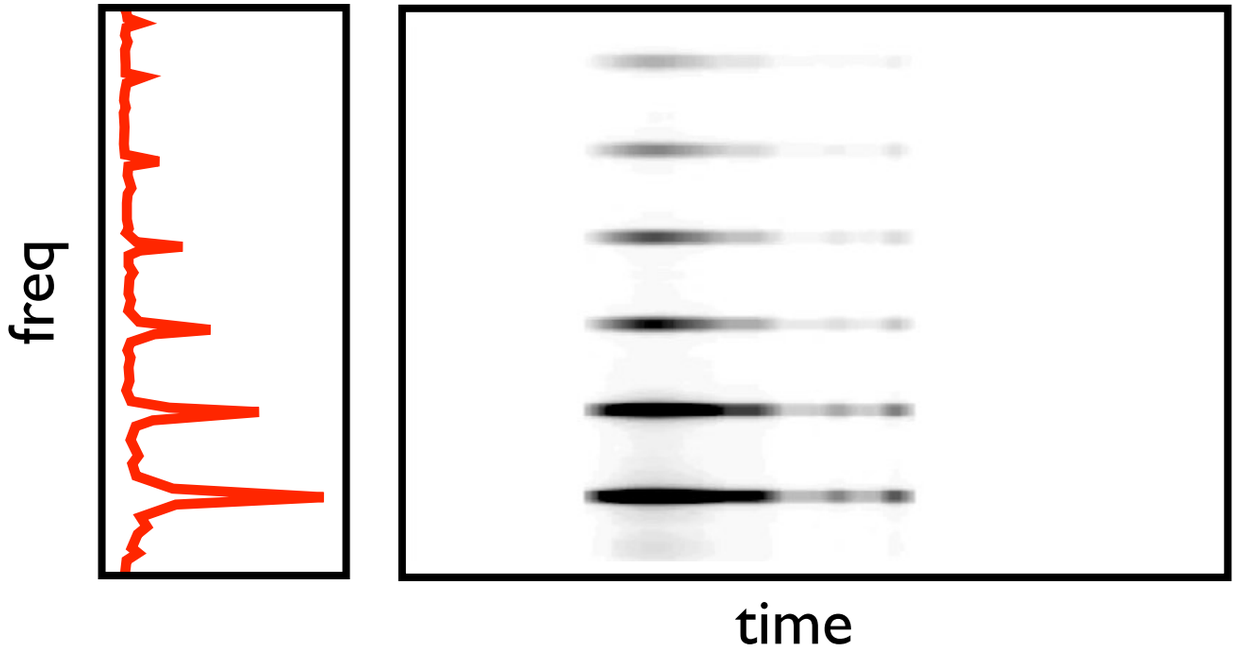


Temporal Shape

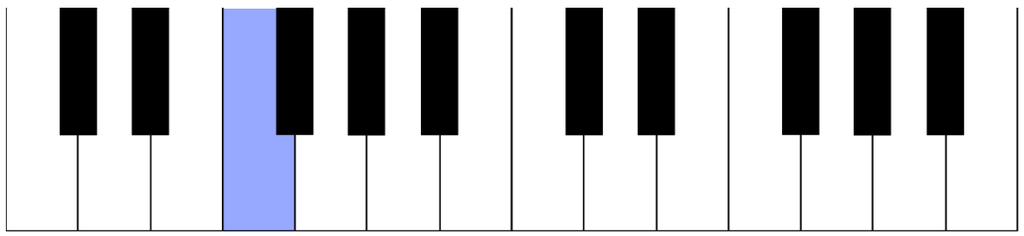


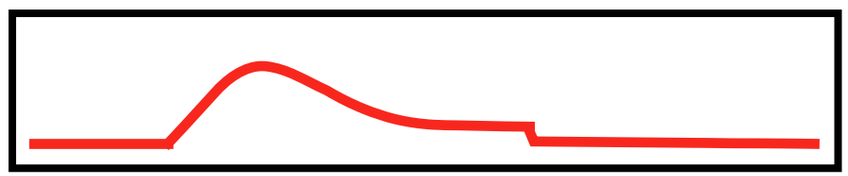


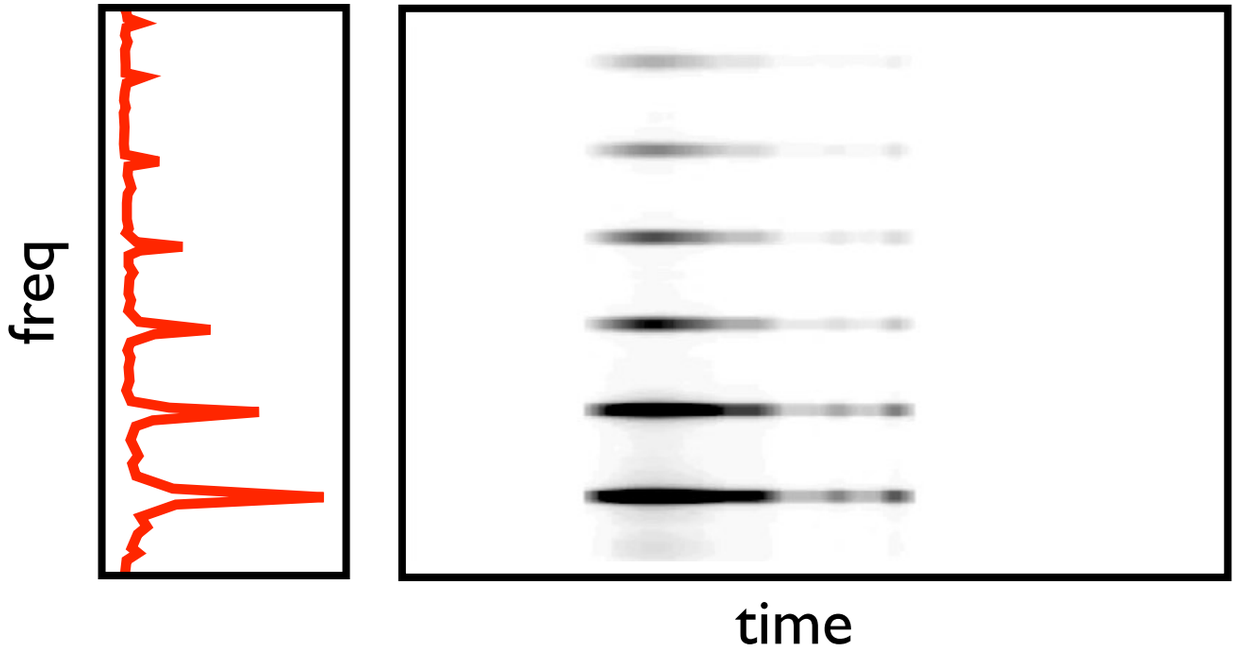




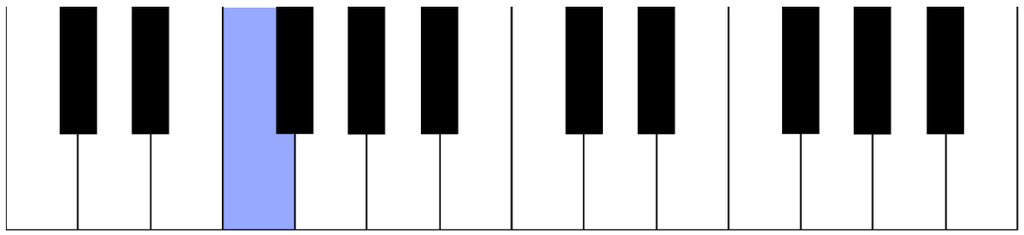
Parameters

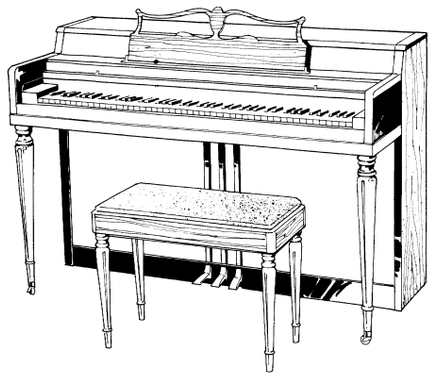


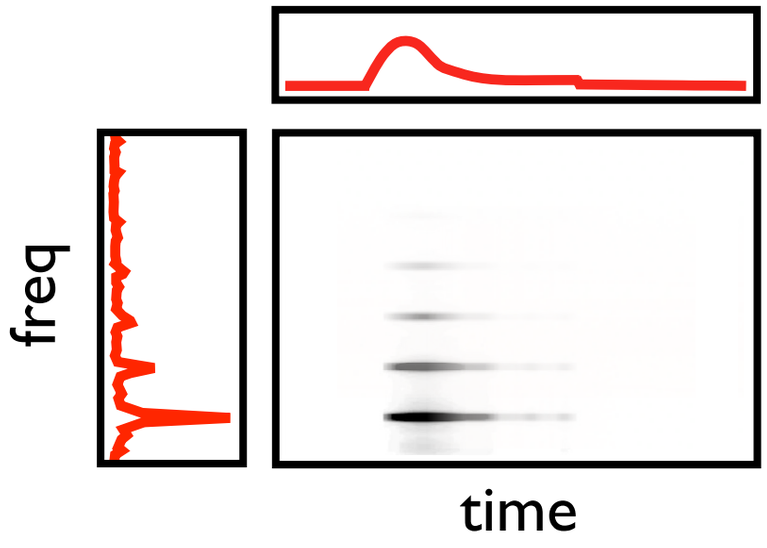


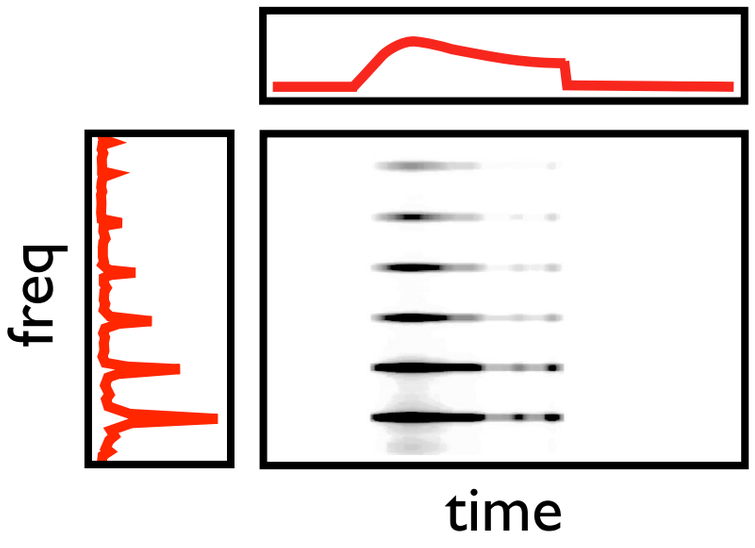


Parameters

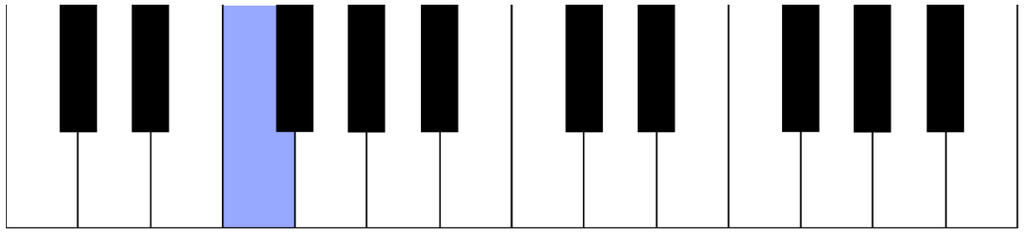


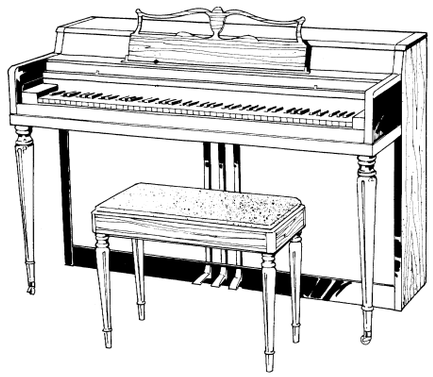


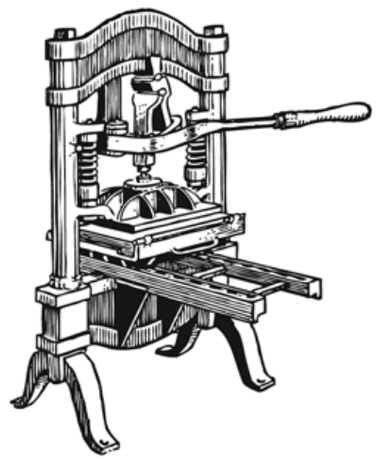


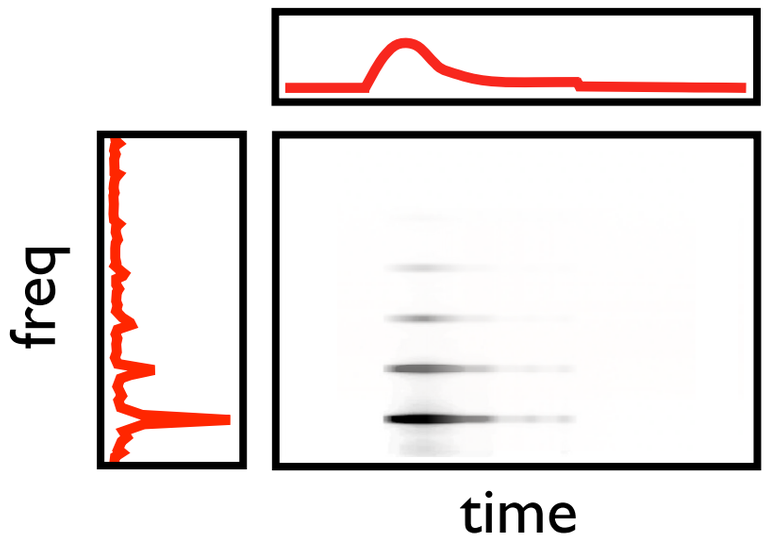


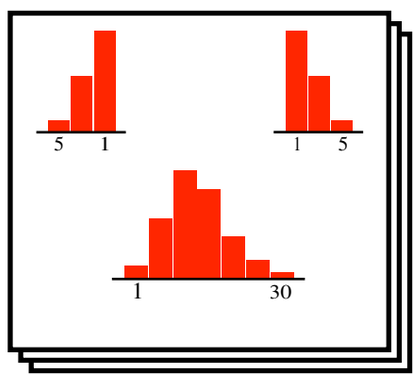
Parameters

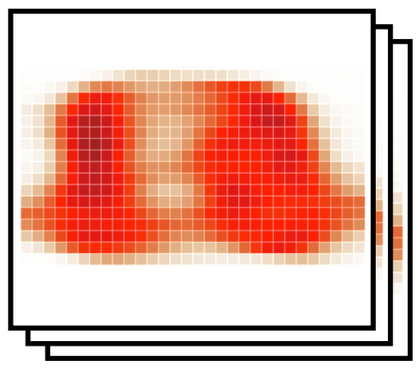




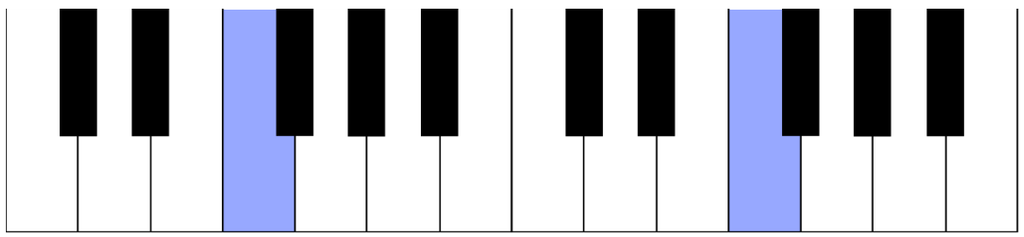


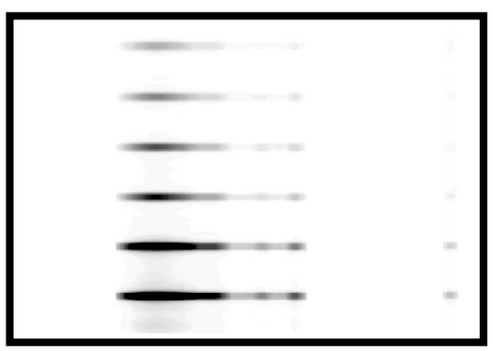


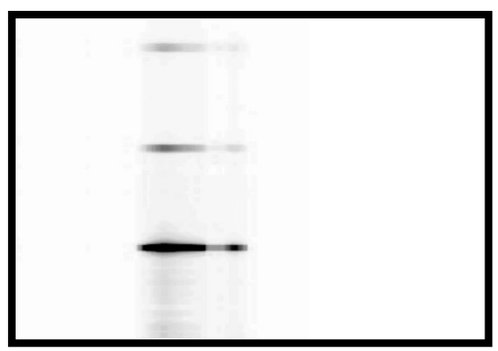


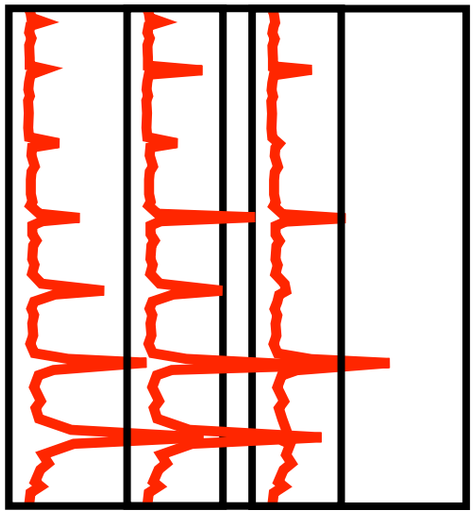


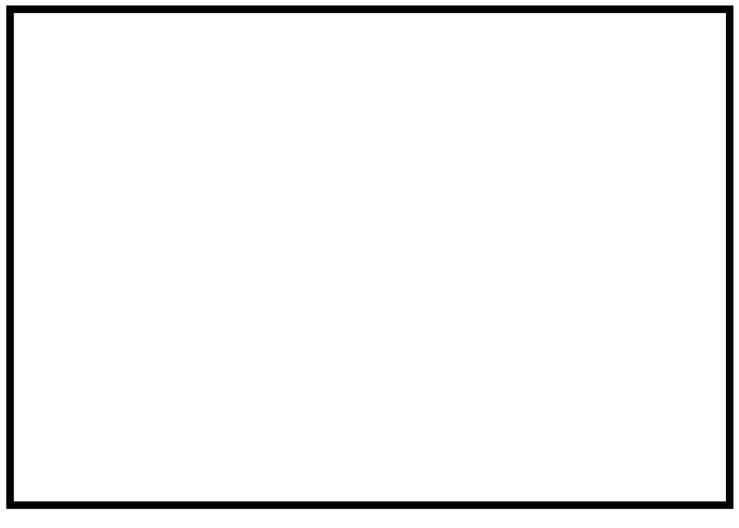
Polyphony



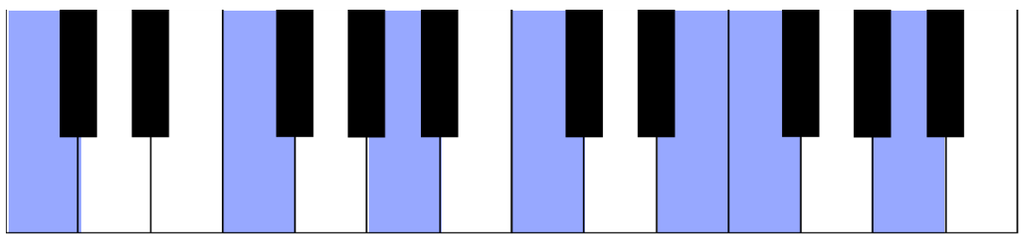


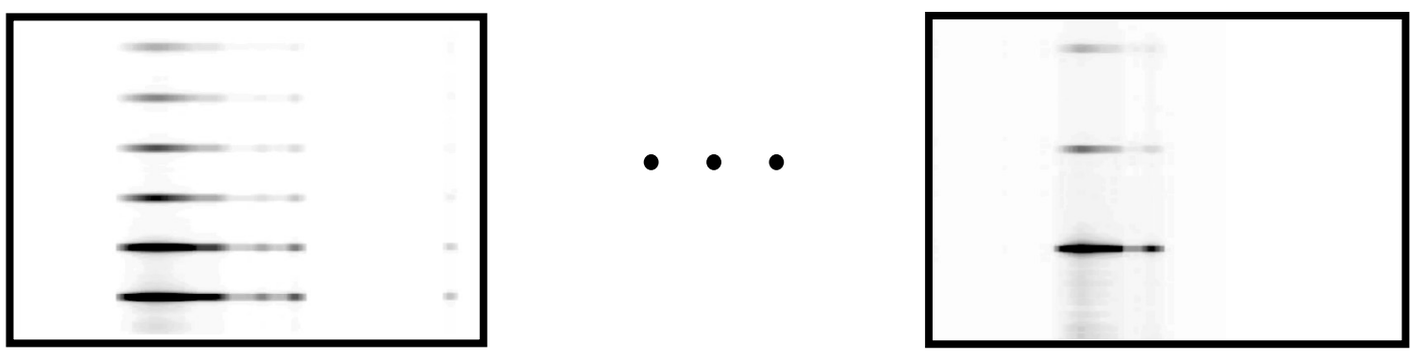




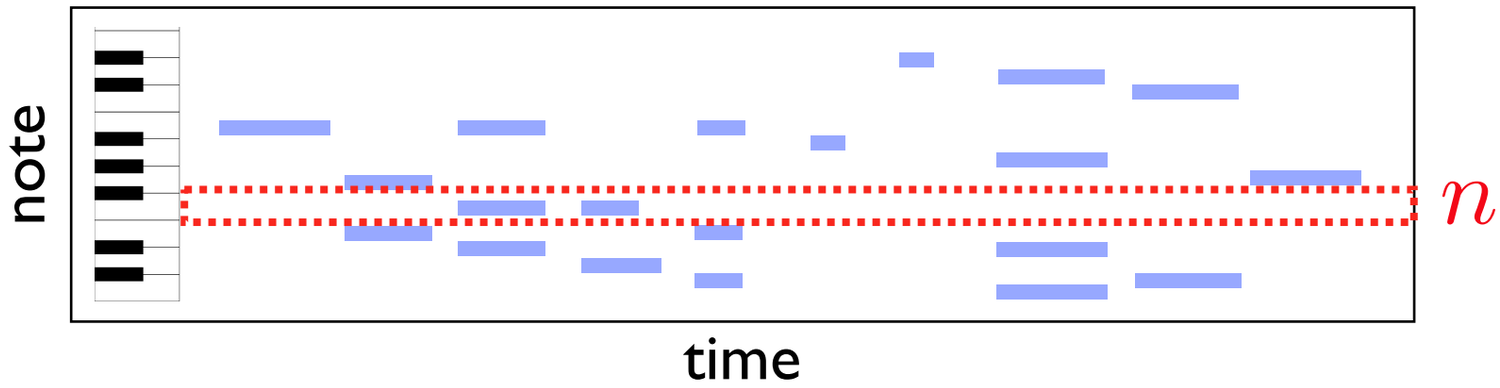


Polyphony

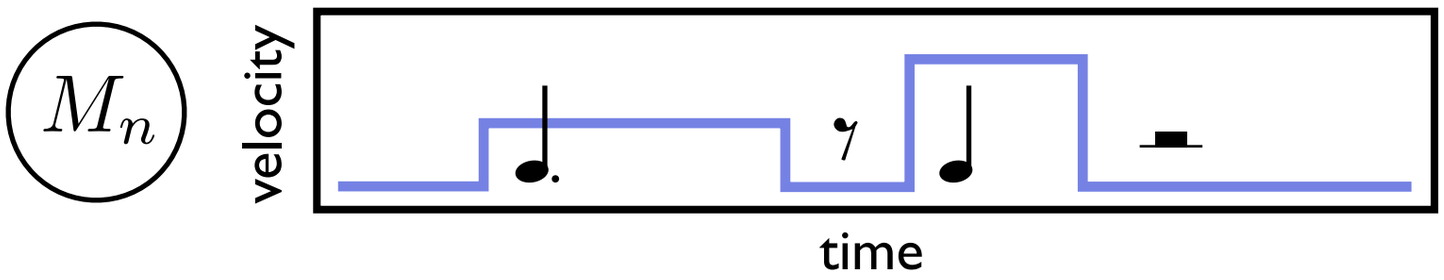




Generative Model

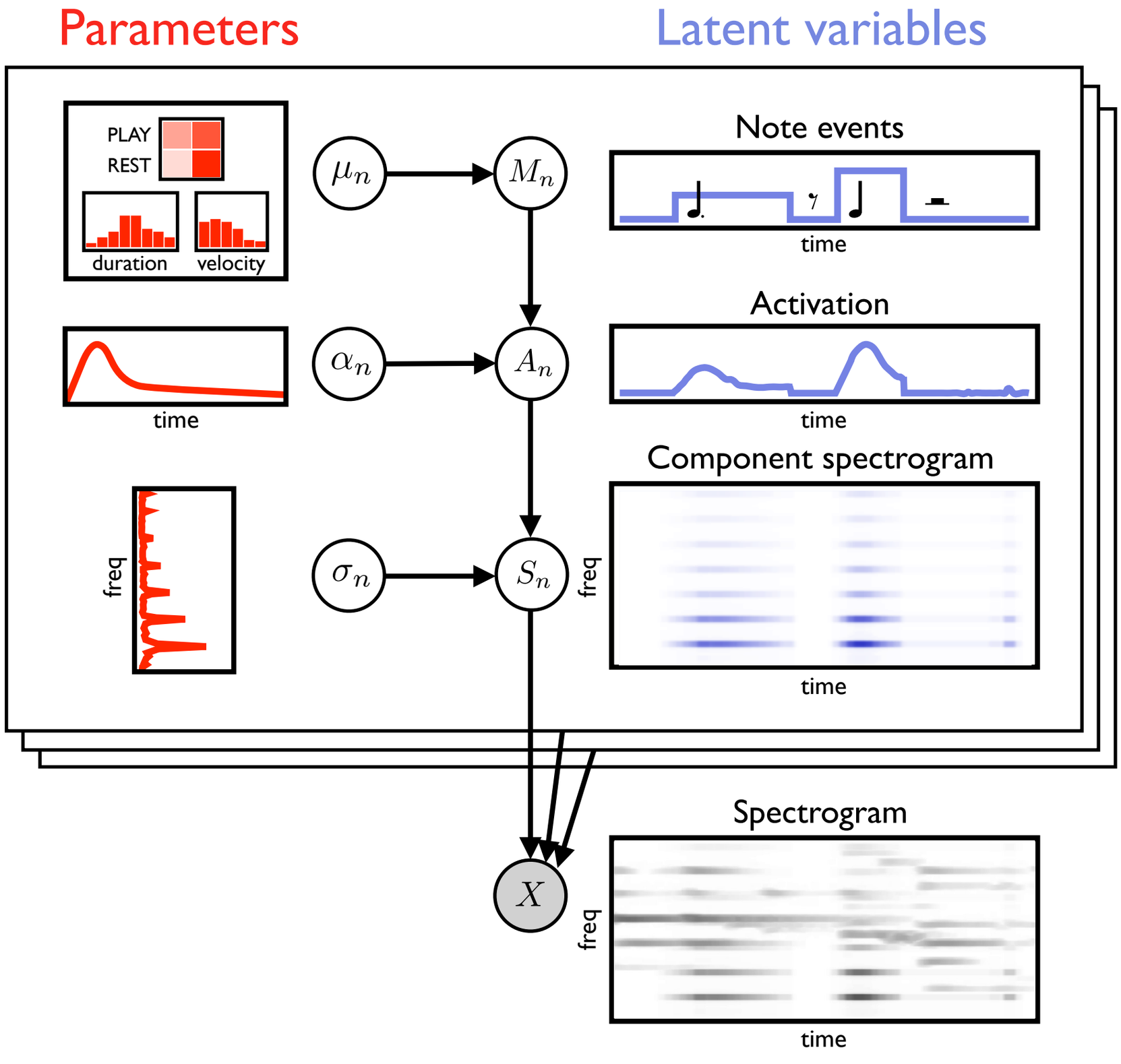


Note events



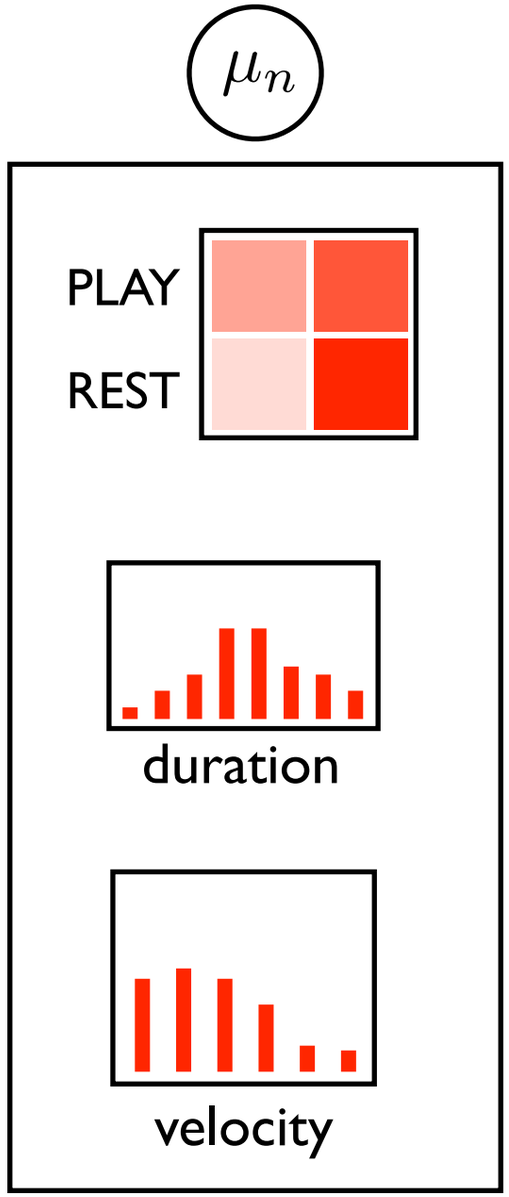
Berkeley N L P

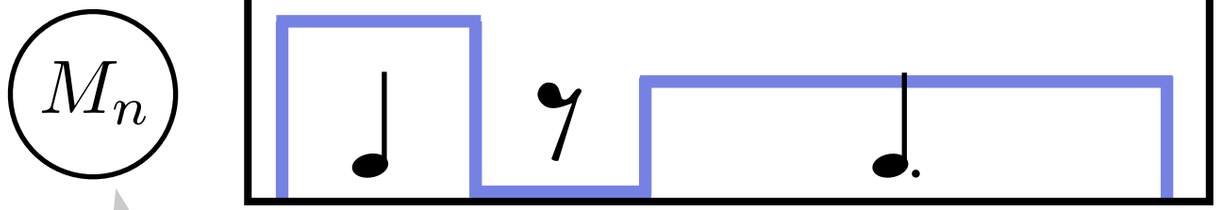
Generative Model

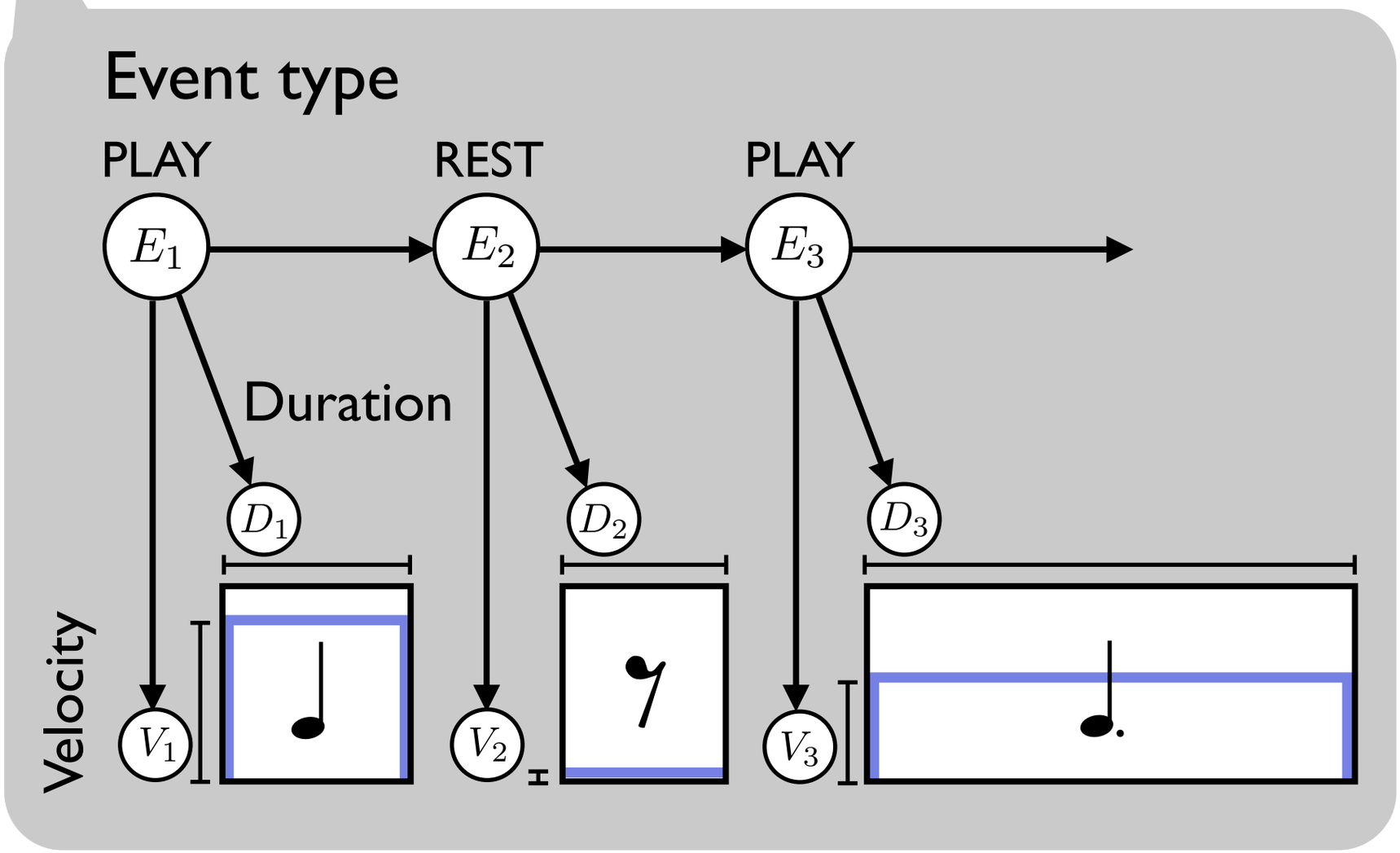


Berkeley N L P

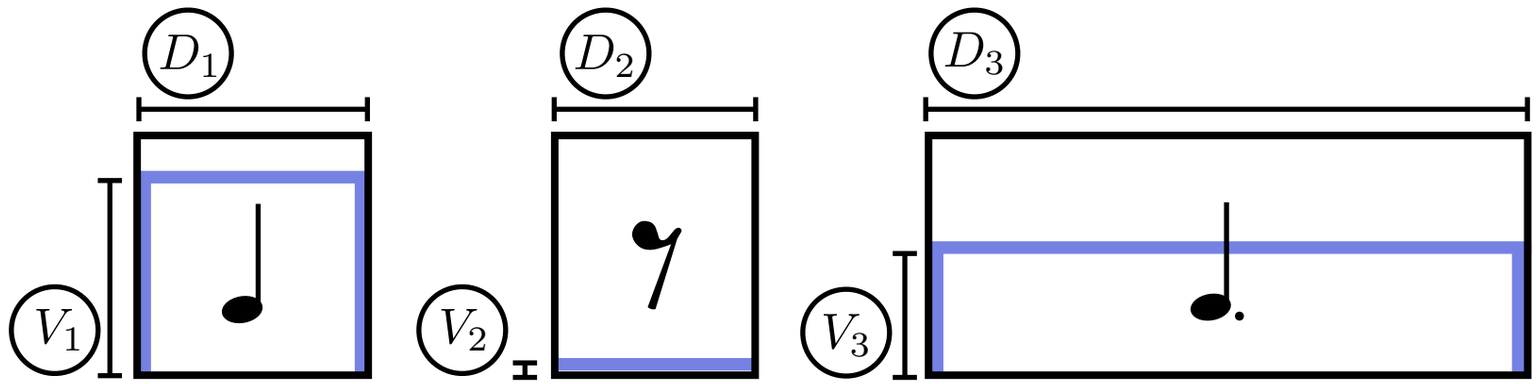
Note Event Model



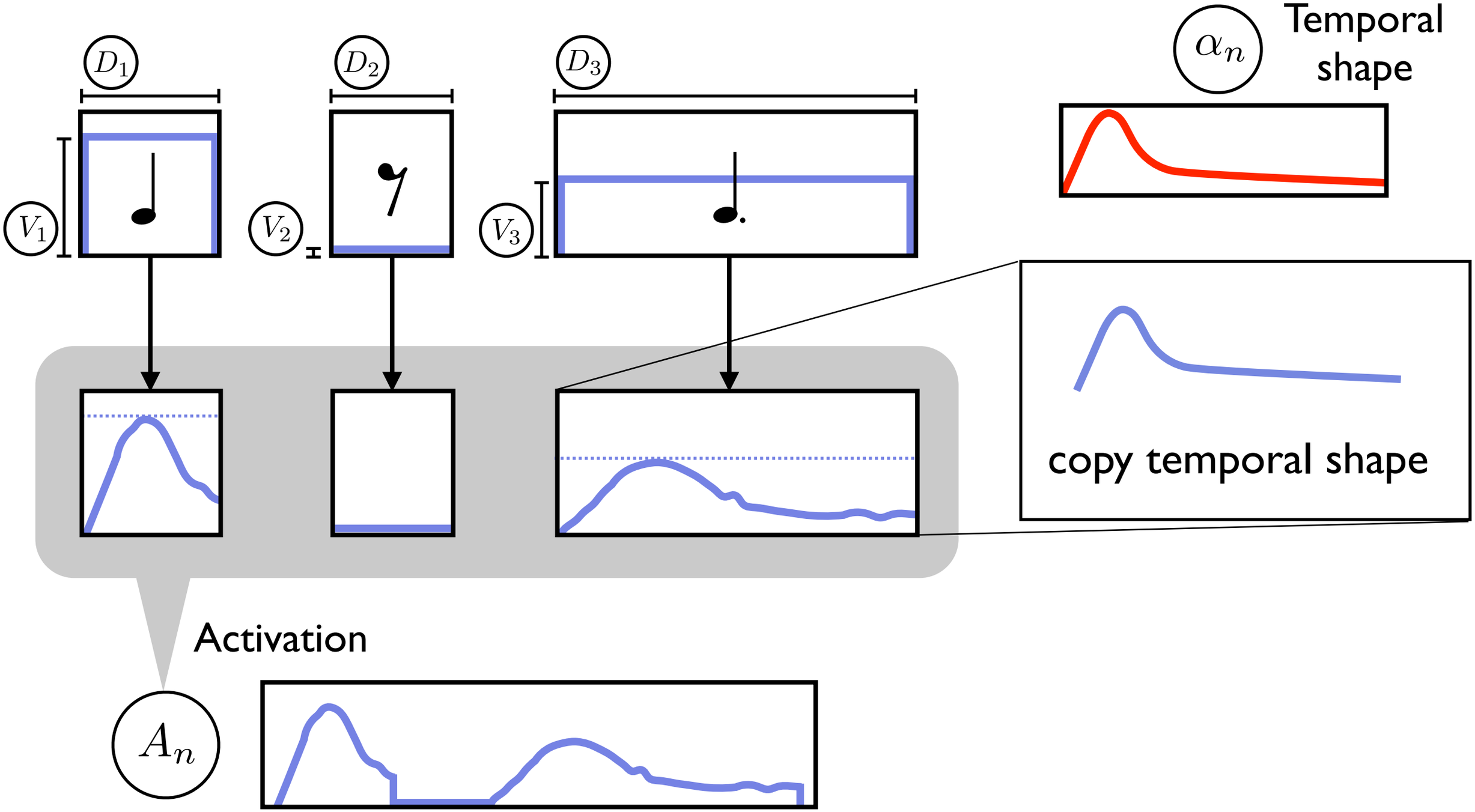




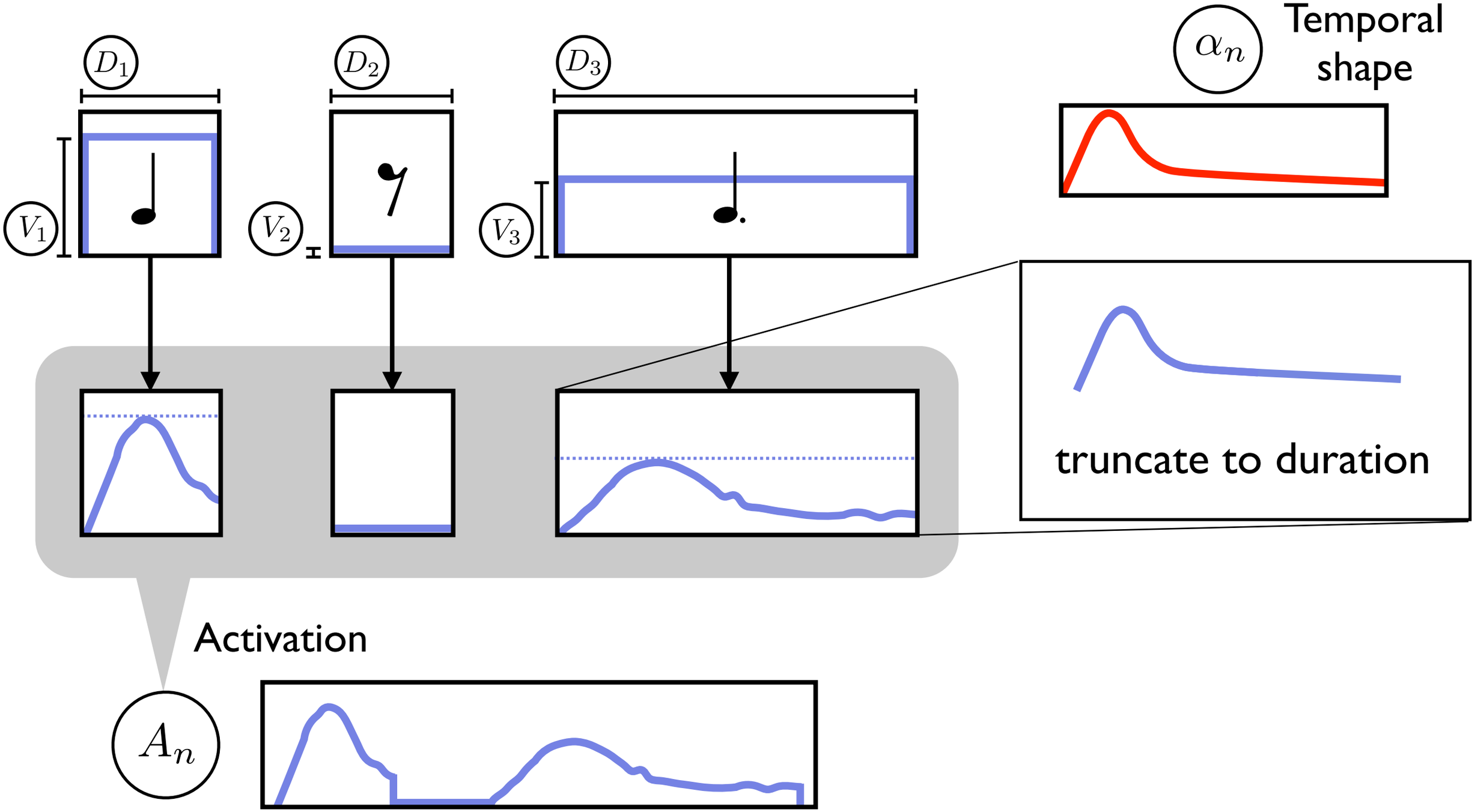
Activation Model



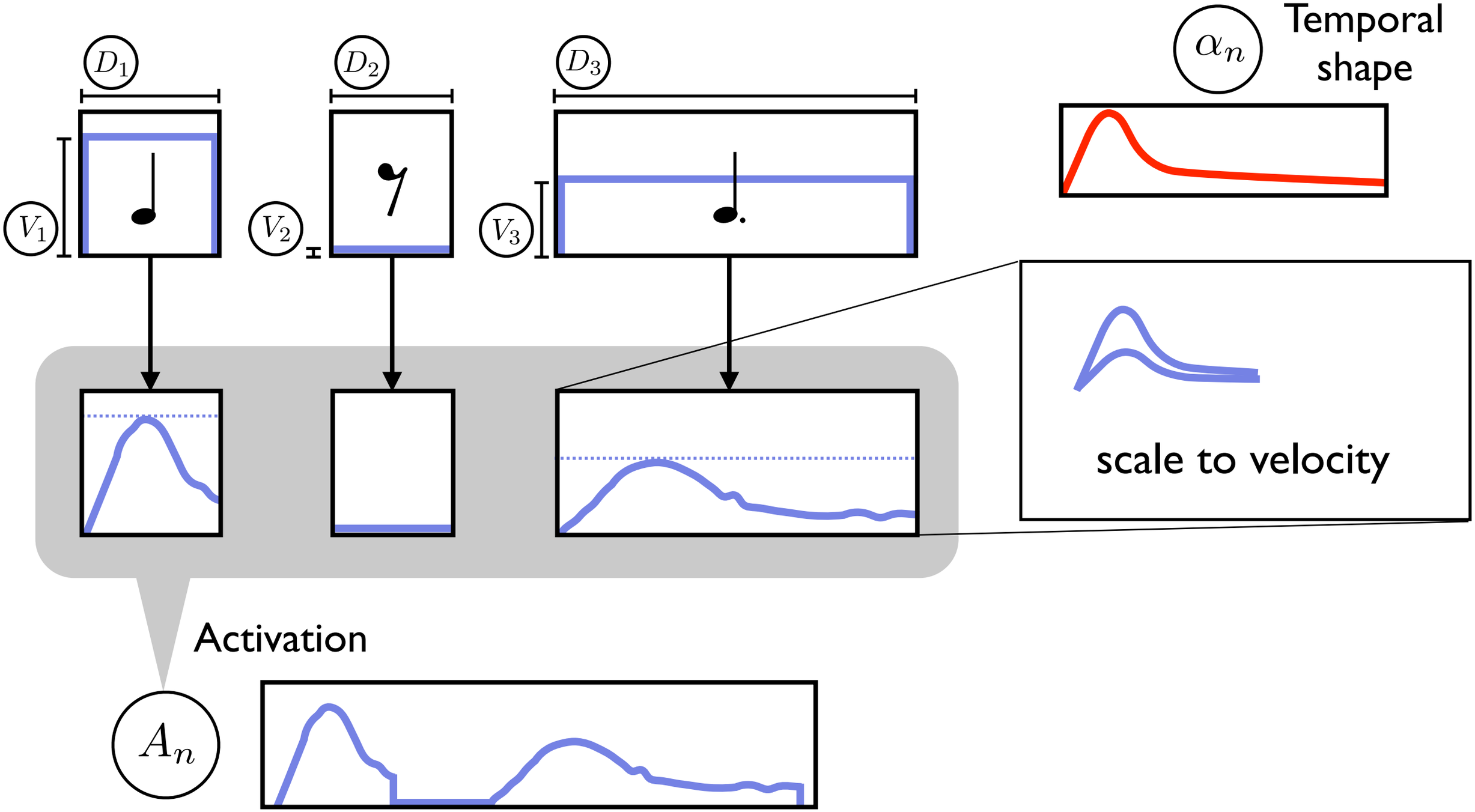
Activation Model



Activation Model

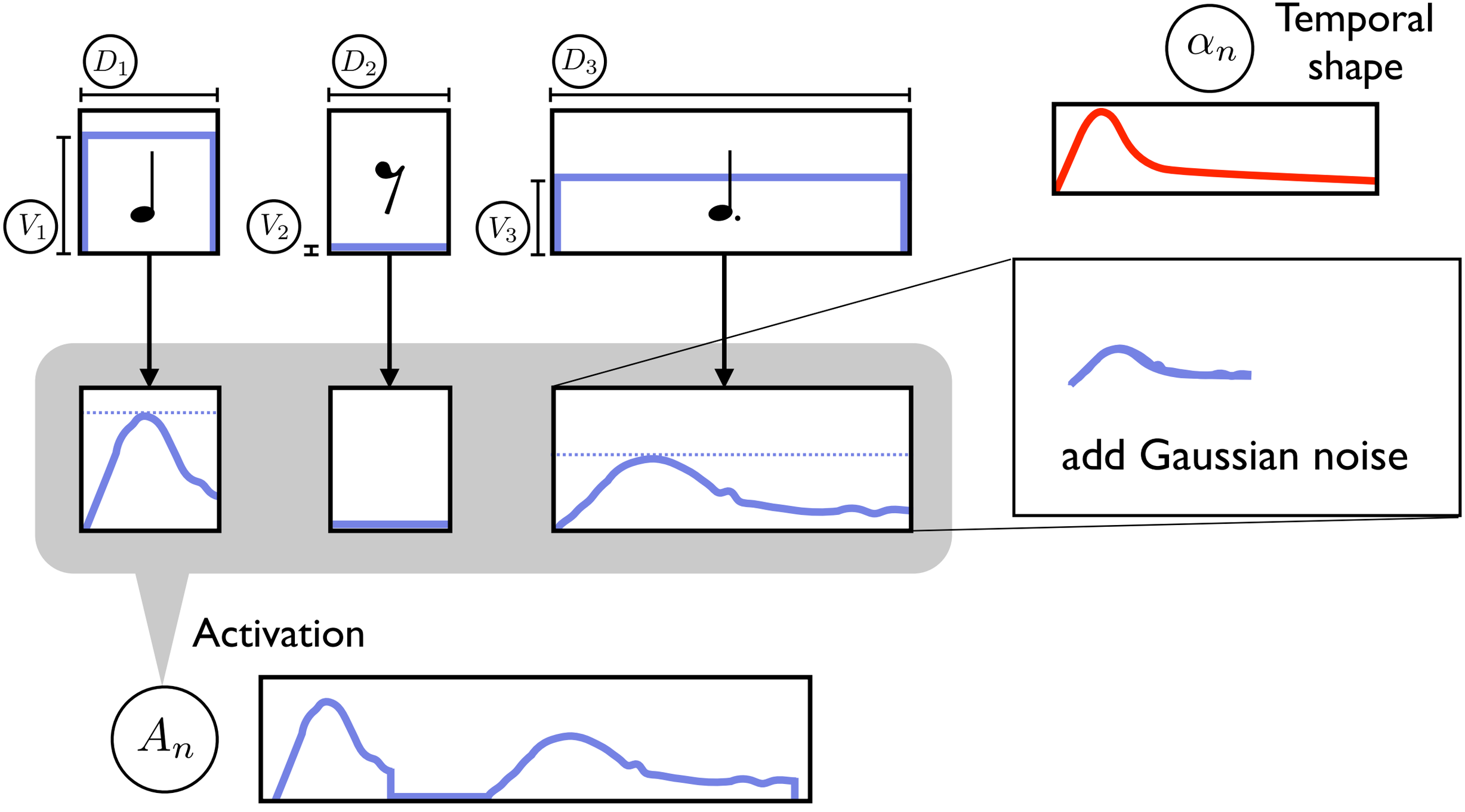


Activation Model

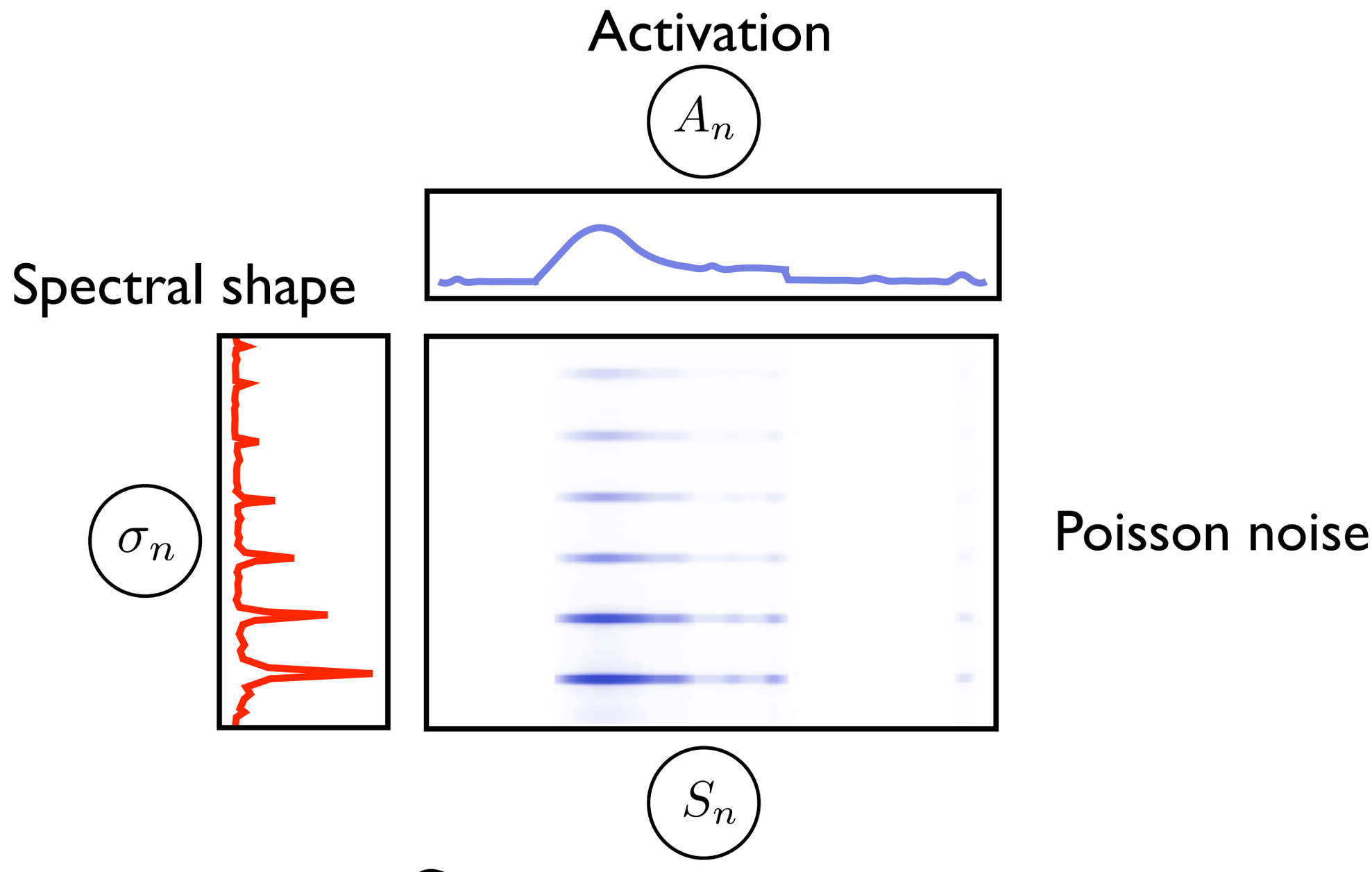


Berkeley N L P

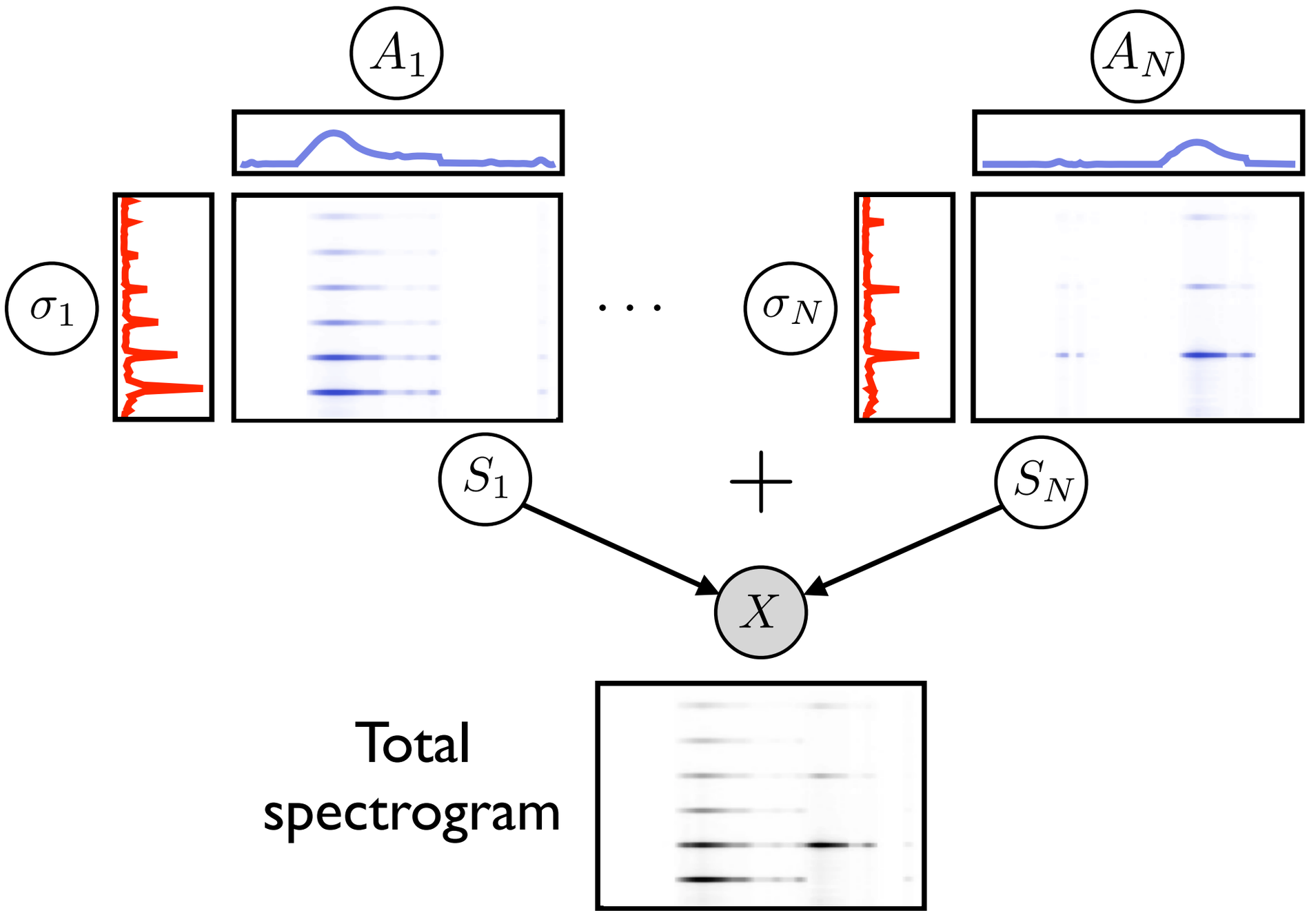
Activation Model



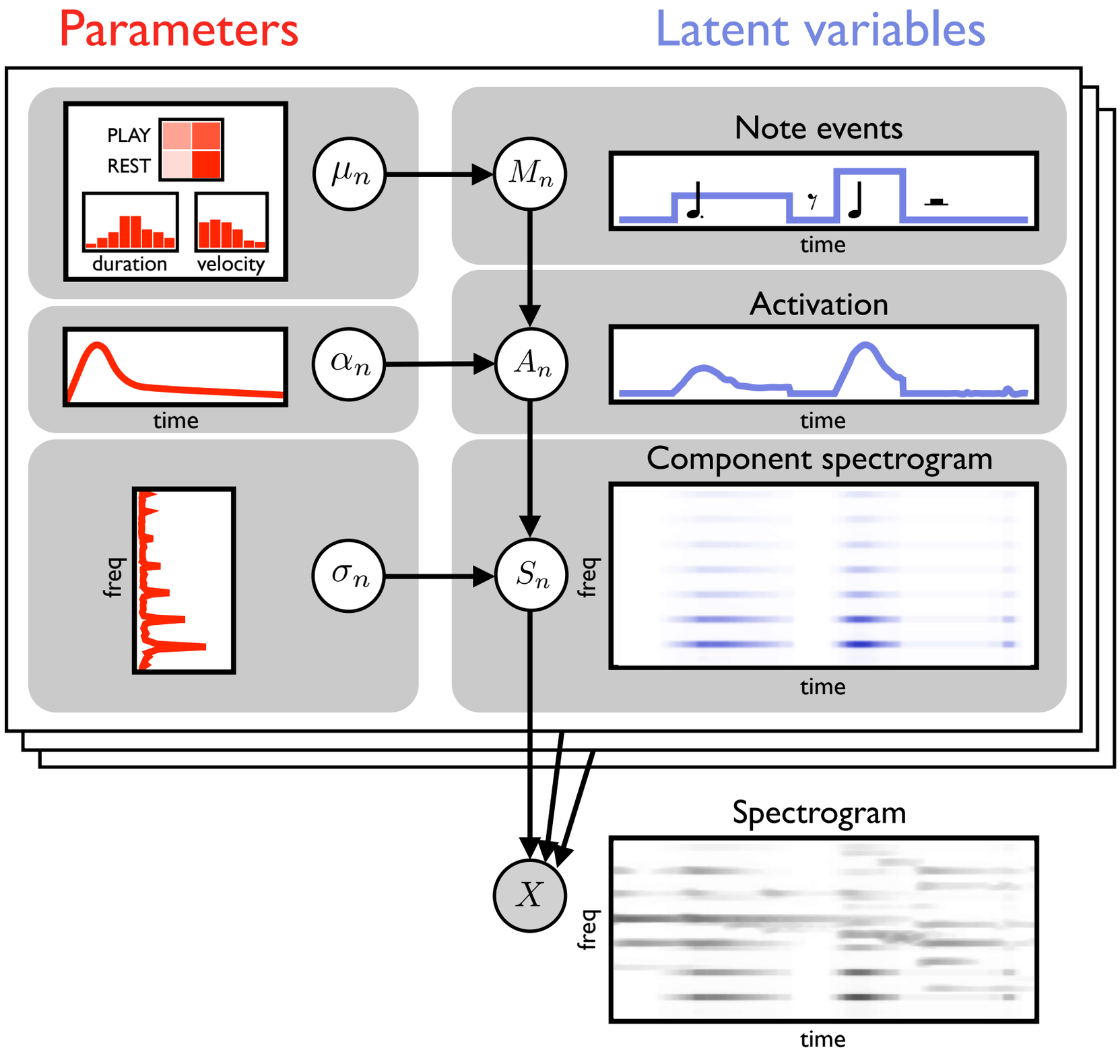
Component Spectrogram Model



Total Spectrogram Model

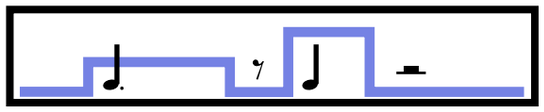


Learning and Inference



Learning and Inference

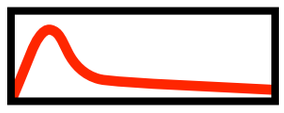
Note events update:



$$M|A, \alpha, \mu$$

Semi-Markov dynamic program

Temporal shapes update:



$$\alpha|A, M$$

Closed form update

Activations update:

$$A|M, X, \alpha, \sigma$$

Exponentiated gradient ascent

Spectral shapes update:

$$\sigma|A, X$$

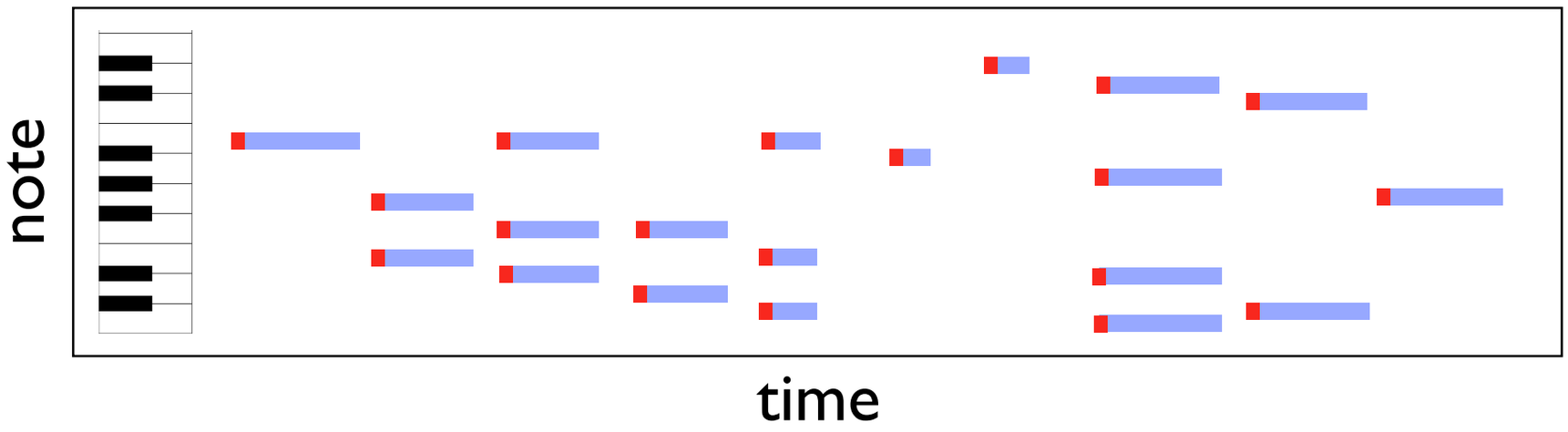
Exponentiated gradient ascent

Experiments

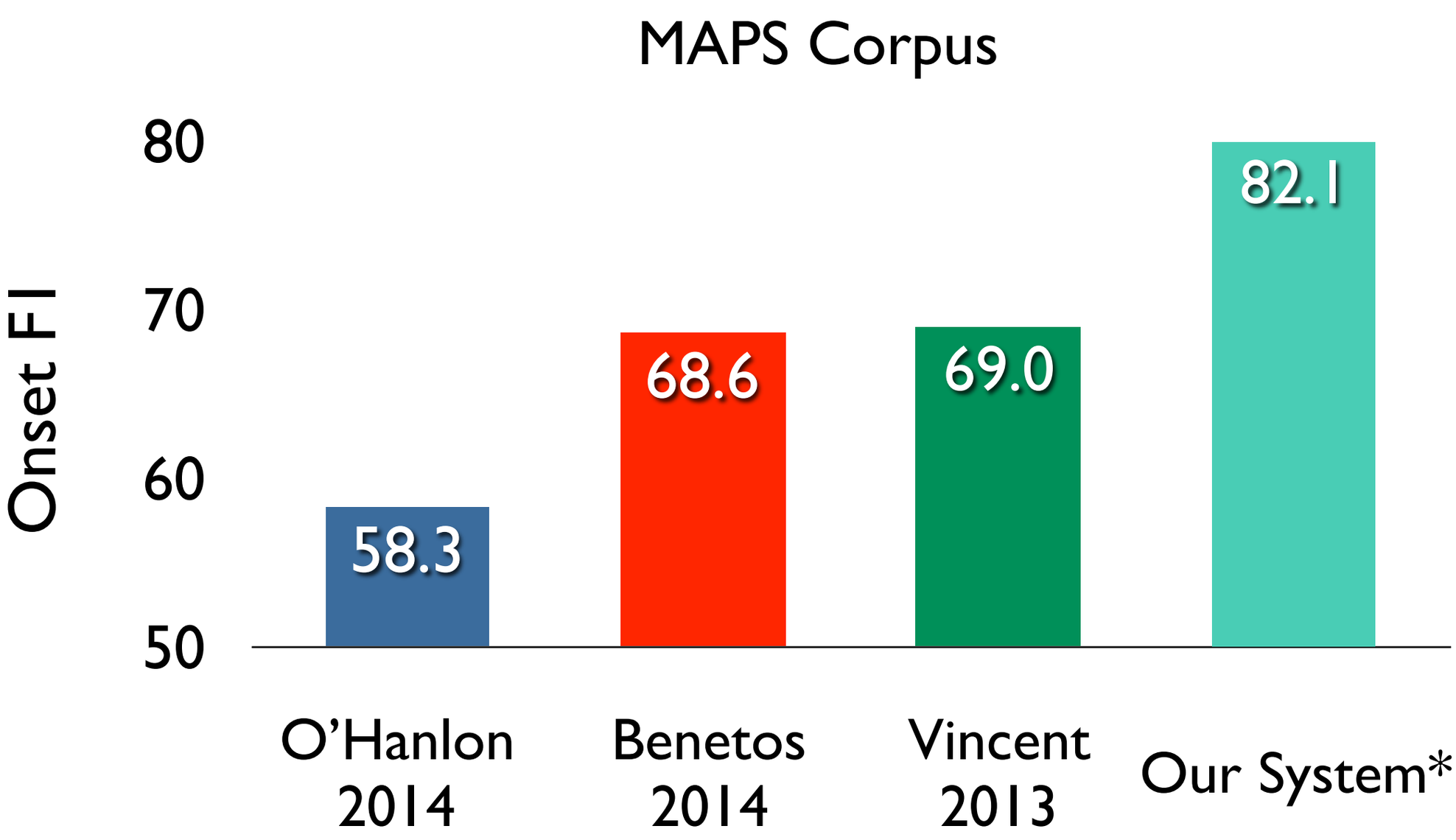
Berkeley N L P

Evaluation

Onset F1



Results



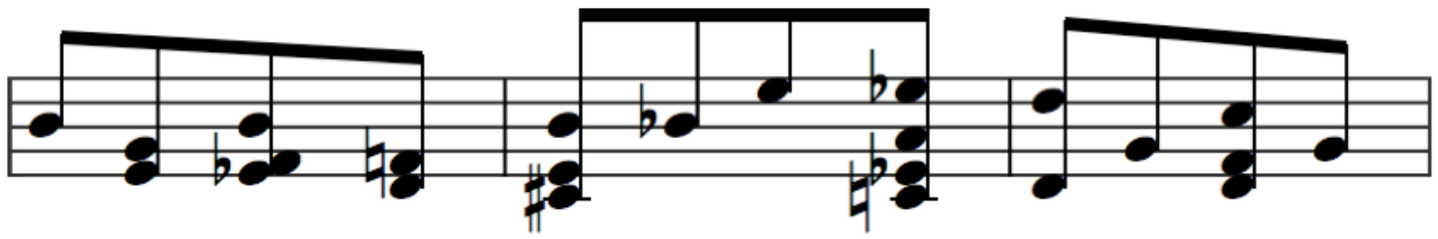
[Valentin et al. 2010]

[Berg-Kirkpatrick et al. 2014]

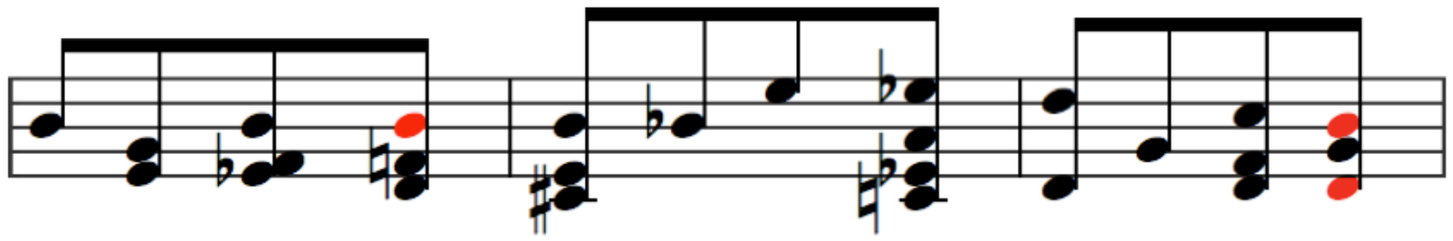
Demo!

Transcription

Reference



Predicted



Resynthesized Examples

Grieg input

Grieg resynth piano

Grieg resynth guitar

Resynthesized Examples

Chopin input

Chopin resynth piano

Chopin resynth guitar

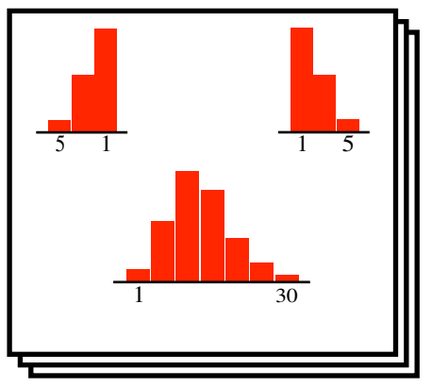
Resynthesized Examples

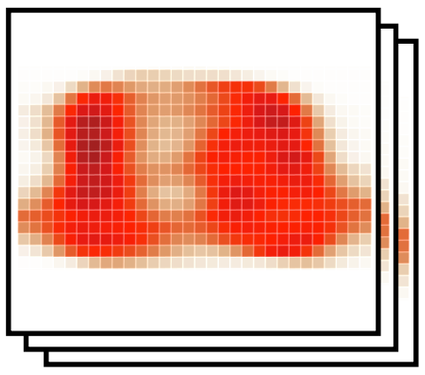
Beethoven input

Beethoven resynth piano

Beethoven resynth guitar

System Output Example





taken ill and taken away -- I remember

Transcription

Google Tesseract

and Ch’: priftmer anhc bar. Jacob Lazarus and his IHP1 uh: prifoner. were both together when! rcccivcd lhczn. I fold eievén pair of than for xiirce guincas, and dclivcrcd the rcll'l:.in-d:r hack lo :11: prifuner. 1 fold ftvcn pairof filk to Mark Simpcr : nncpuir of mixcd. and. mo pair of Ifircad to lhz: foolnun, and on: pair of zhrzad to lh: barber. '

Q: What is the foolmarfs name?

Fraum Mgfzr. I dun’: know.

Hairy Hzrvir. l was flandingar the Camp Icr waizin far the thcrrilfs ufliceruo employ in: : Mo 3‘: daughter came for me to 0 am! take the prifoncr. 1 Wm! to |hc Old aailcy

Ocular

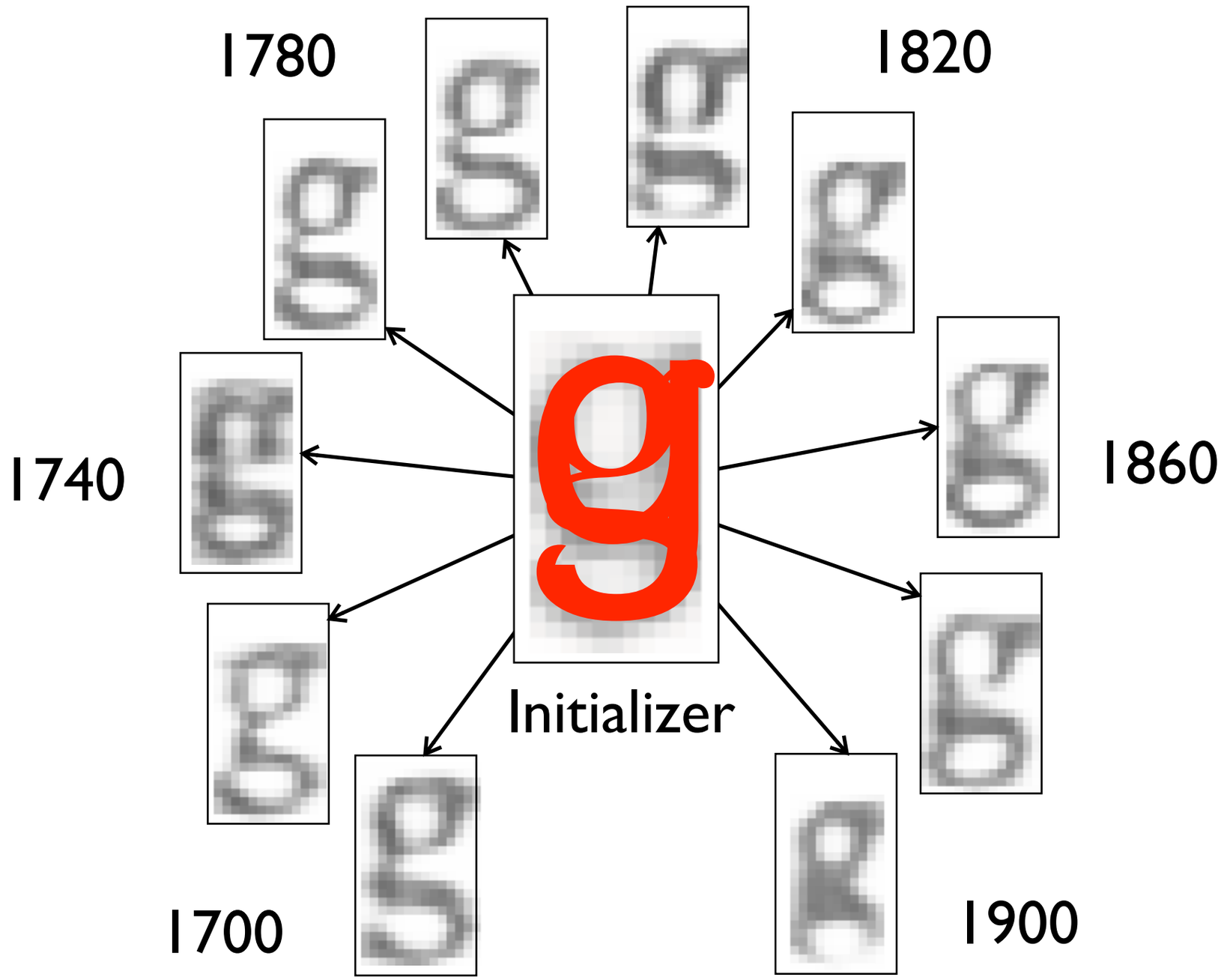
the prisoner at the bar. Jacob Lazarus and his wife, the prisoners were both together when I received them. I sold eleven pair of them for three guineas, and delivered the remainder back to the prisoner. I sold, seven pair of silk to Mark Simpert one pair of mixed, and two pair of thread to the footman, and one pair of thread to the barber,

Ms. What in the footman's name?

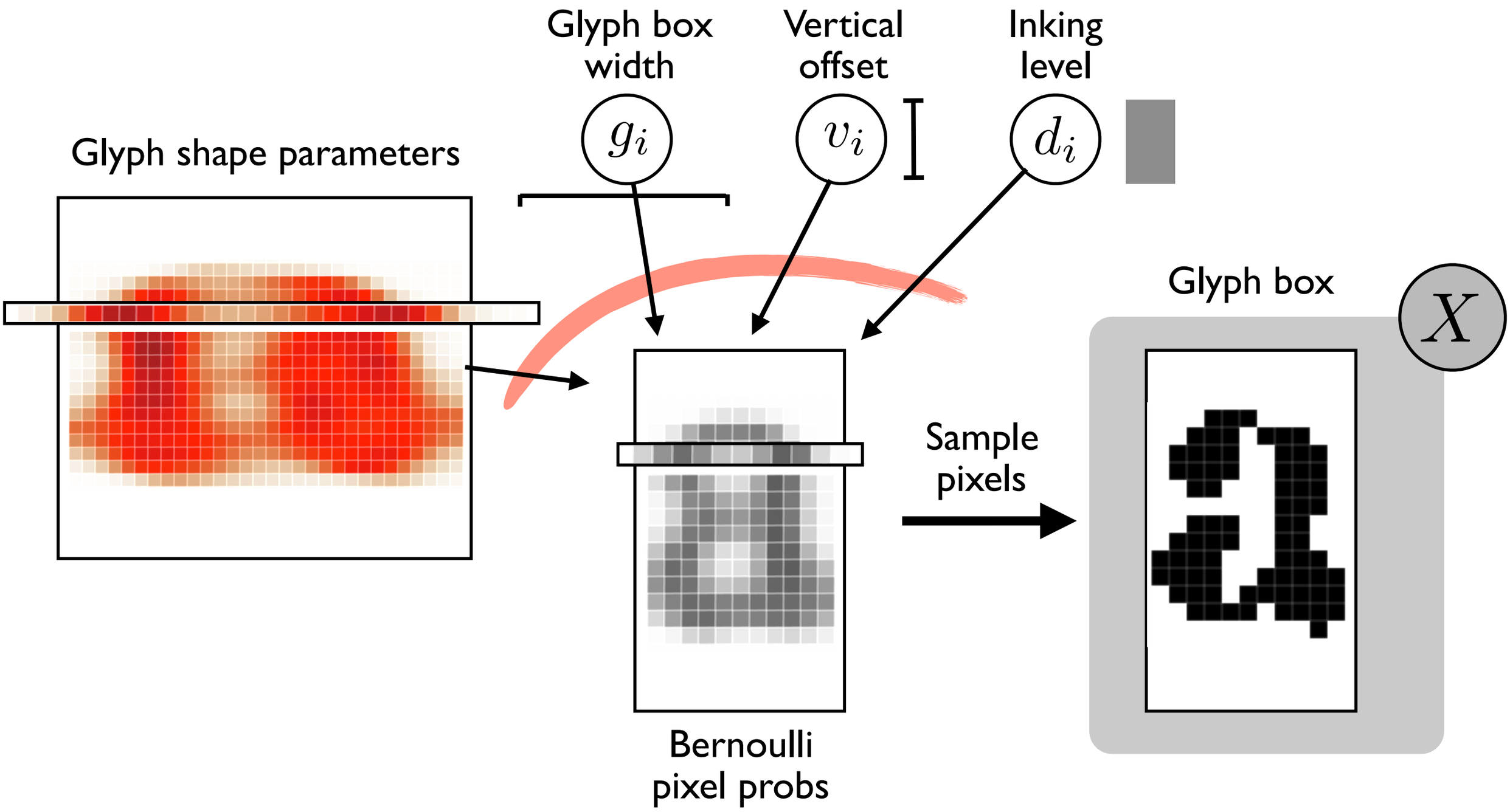
Franco Asyut, I don't know-

Nearly Norris. I was standing at the Compter waiting for the sherrill's officers to employ me a Moses's daughter came for me to go and take the prisoner. I went to the Old Bailey

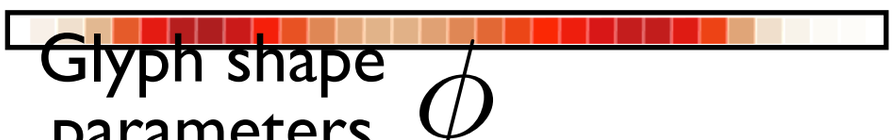
Learned Fonts



Rendering Model



Log-linear Interpolation



Glyph shape parameters ϕ

Bernoulli pixel probs θ

Log-linear Interpolation

Glyph shape parameters ϕ

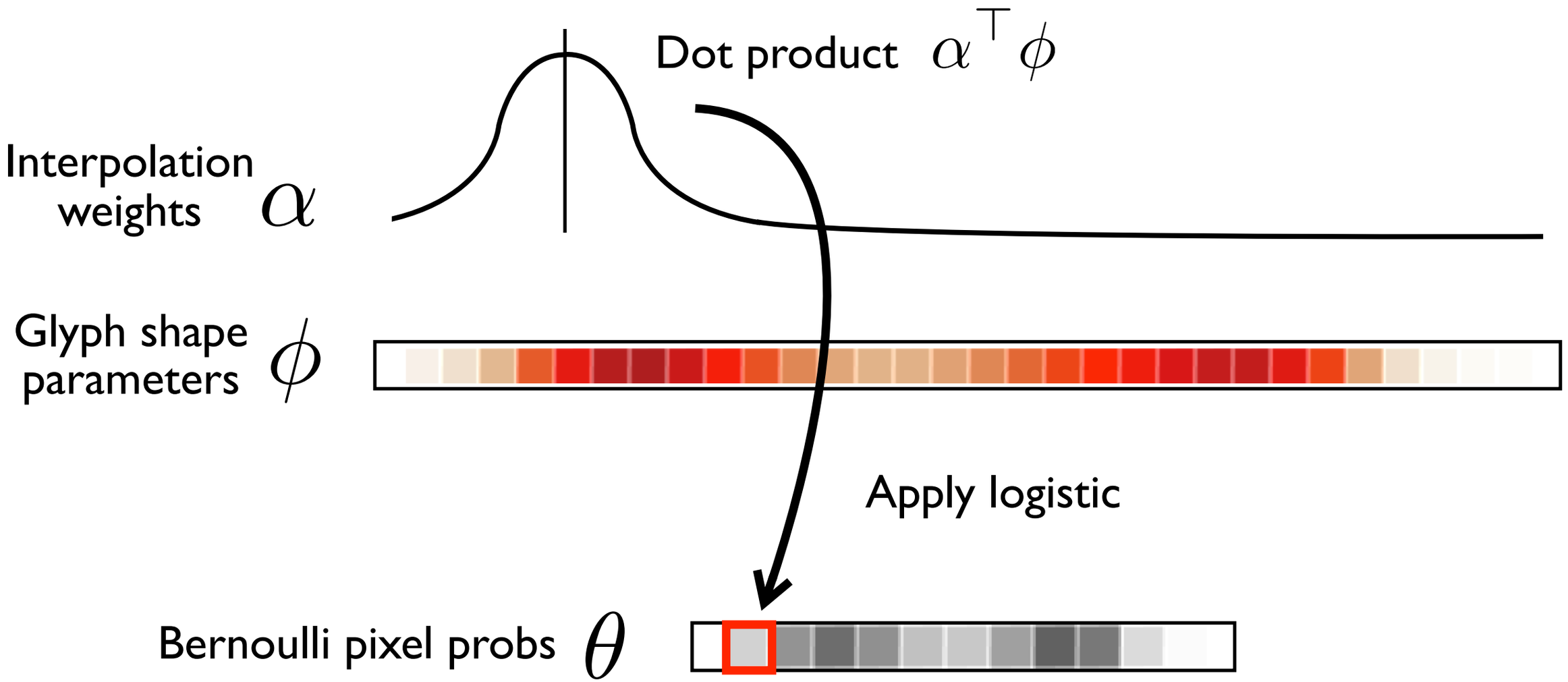
Bernoulli pixel probs θ

Log-linear Interpolation

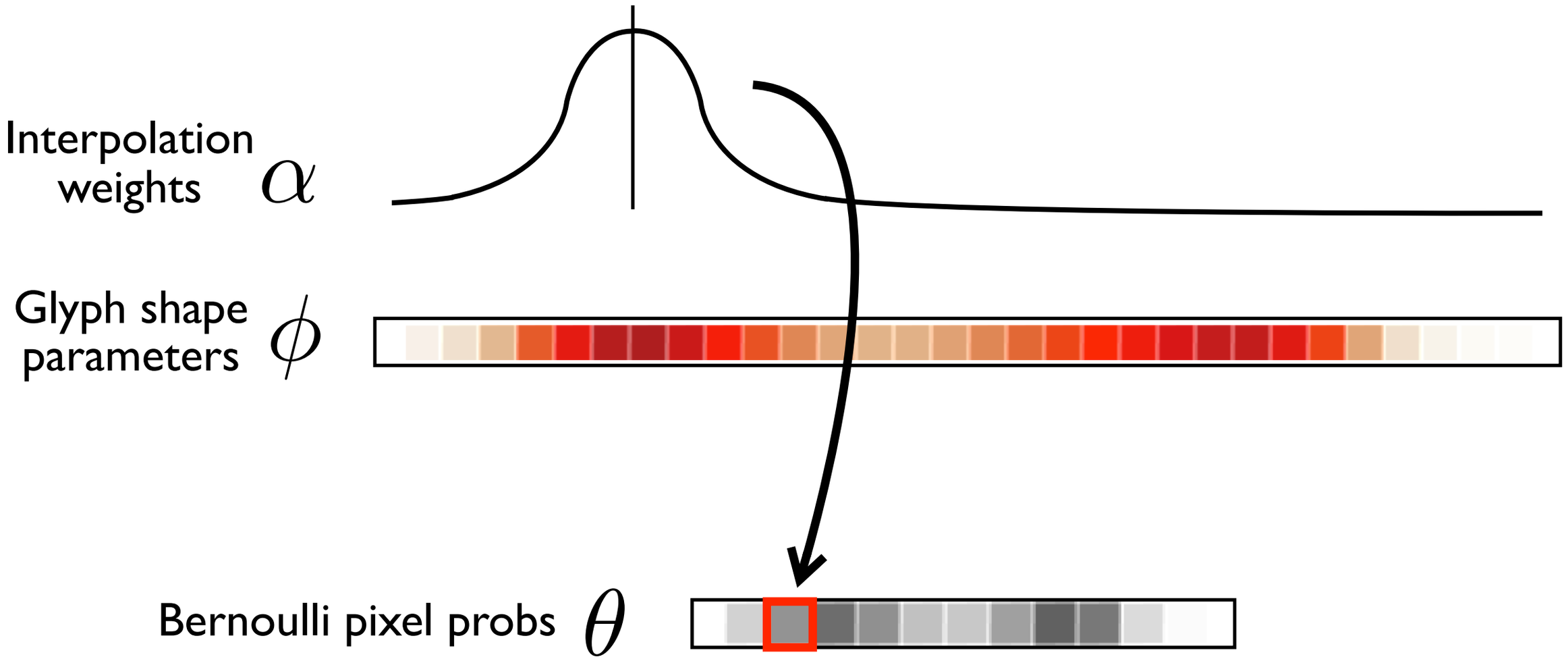
Glyph shape parameters ϕ

Bernoulli pixel probs θ

Log-linear Interpolation



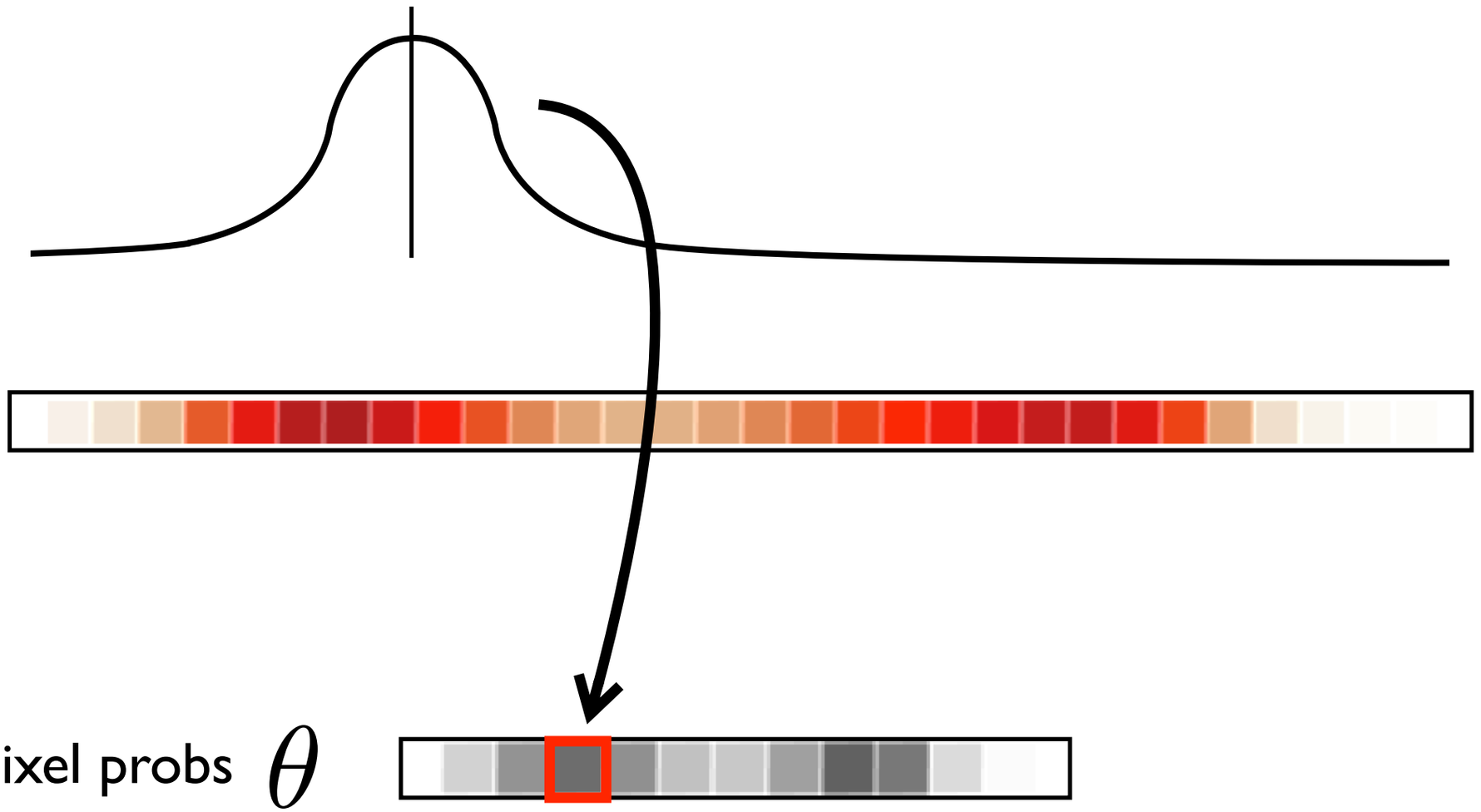
Log-linear Interpolation



Log-linear Interpolation

Interpolation weights α

Glyph shape parameters ϕ

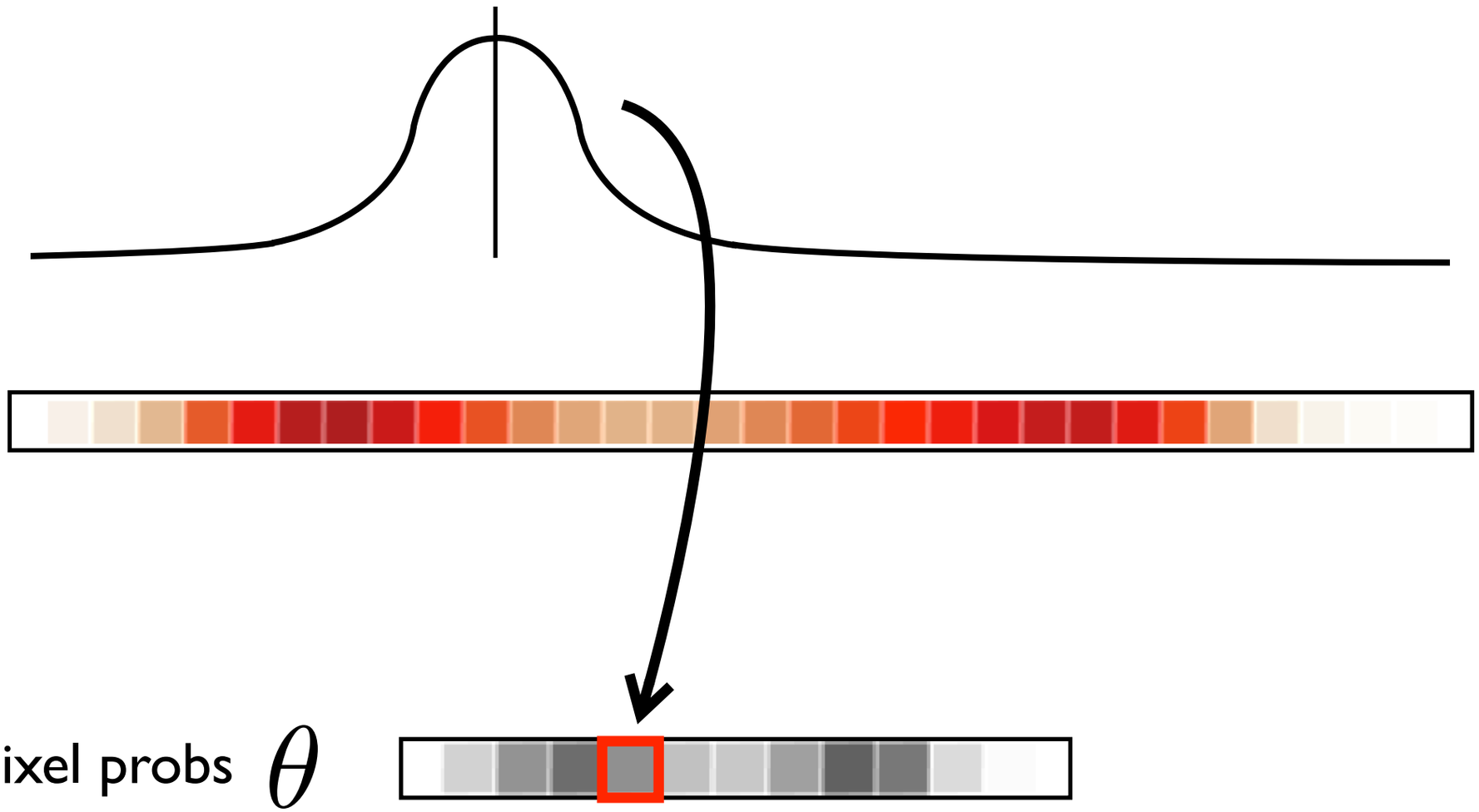


Bernoulli pixel probs θ

Log-linear Interpolation

Interpolation weights α

Glyph shape parameters ϕ



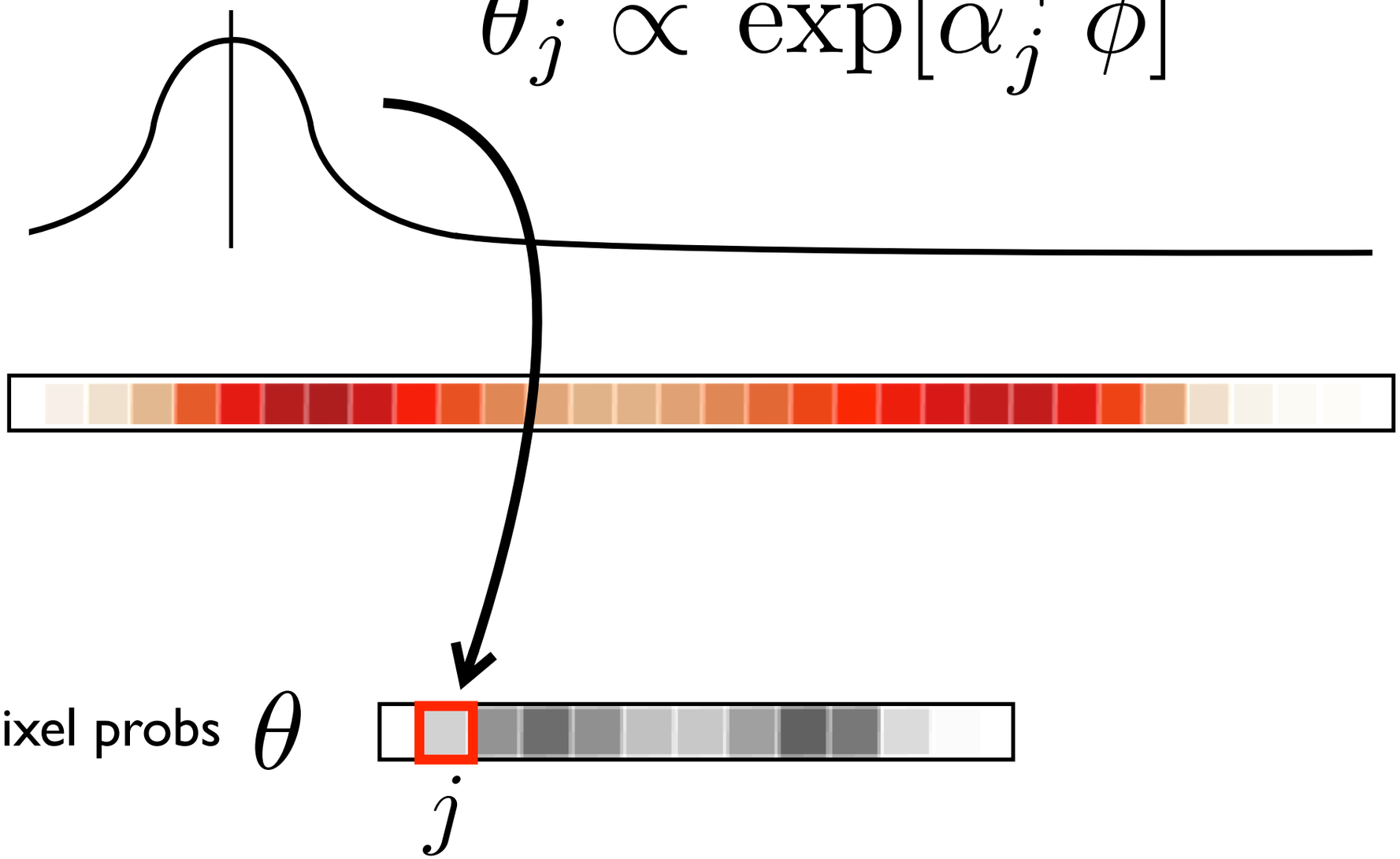
Bernoulli pixel probs θ

Log-linear Interpolation

$$\theta_j \propto \exp[\alpha_j^\top \phi]$$

Interpolation weights α_j

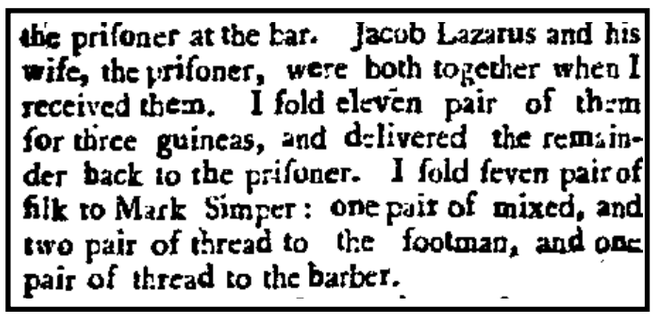
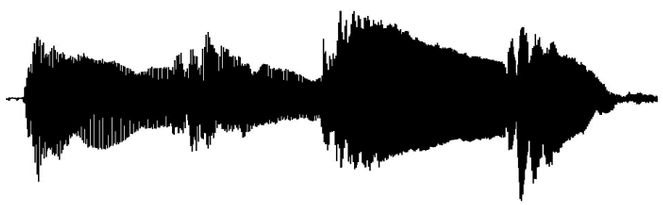
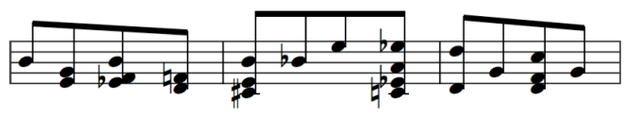
Glyph shape parameters ϕ



Bernoulli pixel probs θ

j

Unsupervised Transcription


the prisoner at the bar. Jacob Lazarus and his wife, the prisoner, were both together when I received them. I sold eleven pair of them for three guineas, and delivered the remainder back to the prisoner. I sold, seven pair of silk to Mark Simper one pair of mixed, and two pair of thread to the footman, and one pair of thread to the barber,