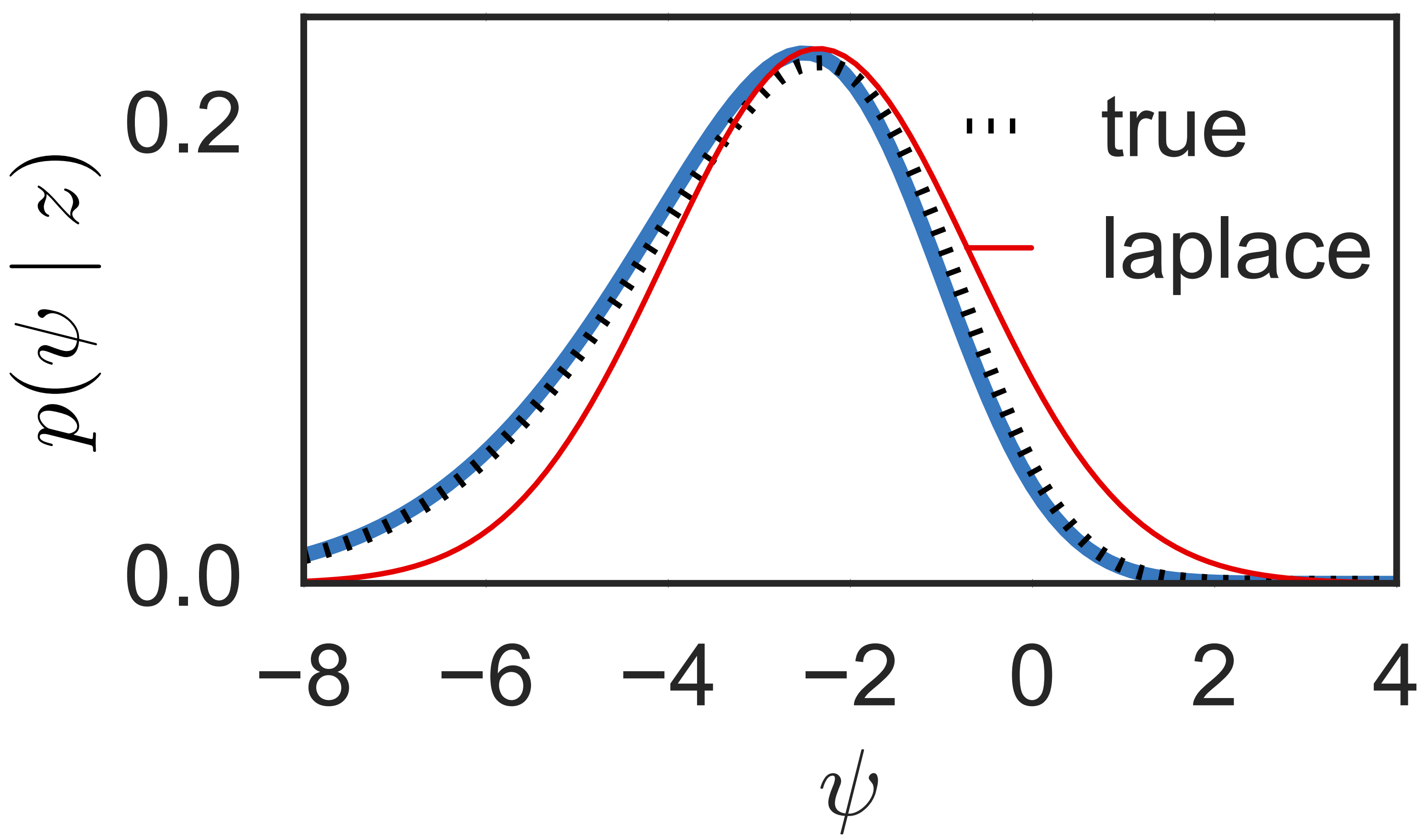
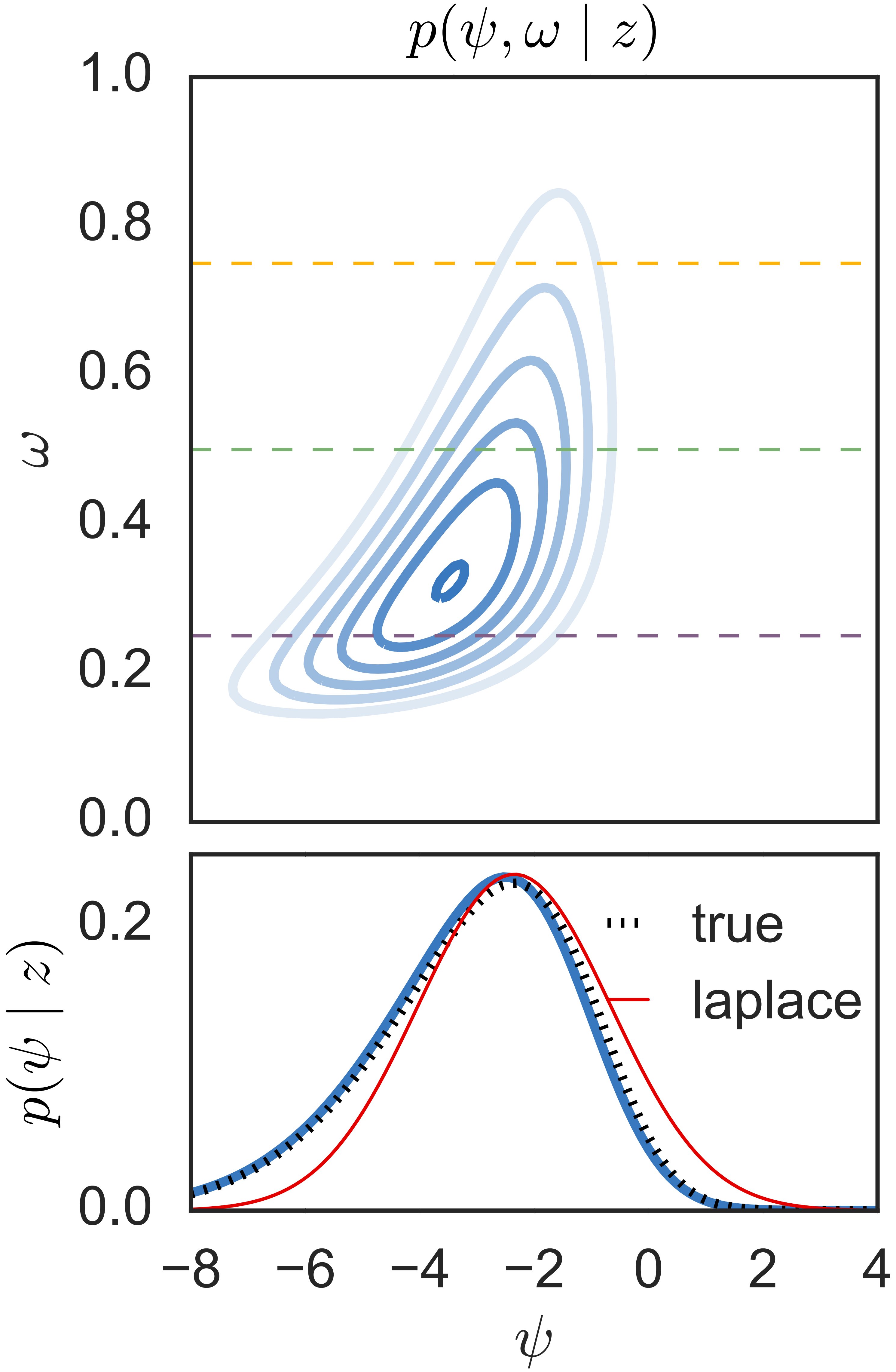
Augmentation schemes for GP Classification with a logit link function



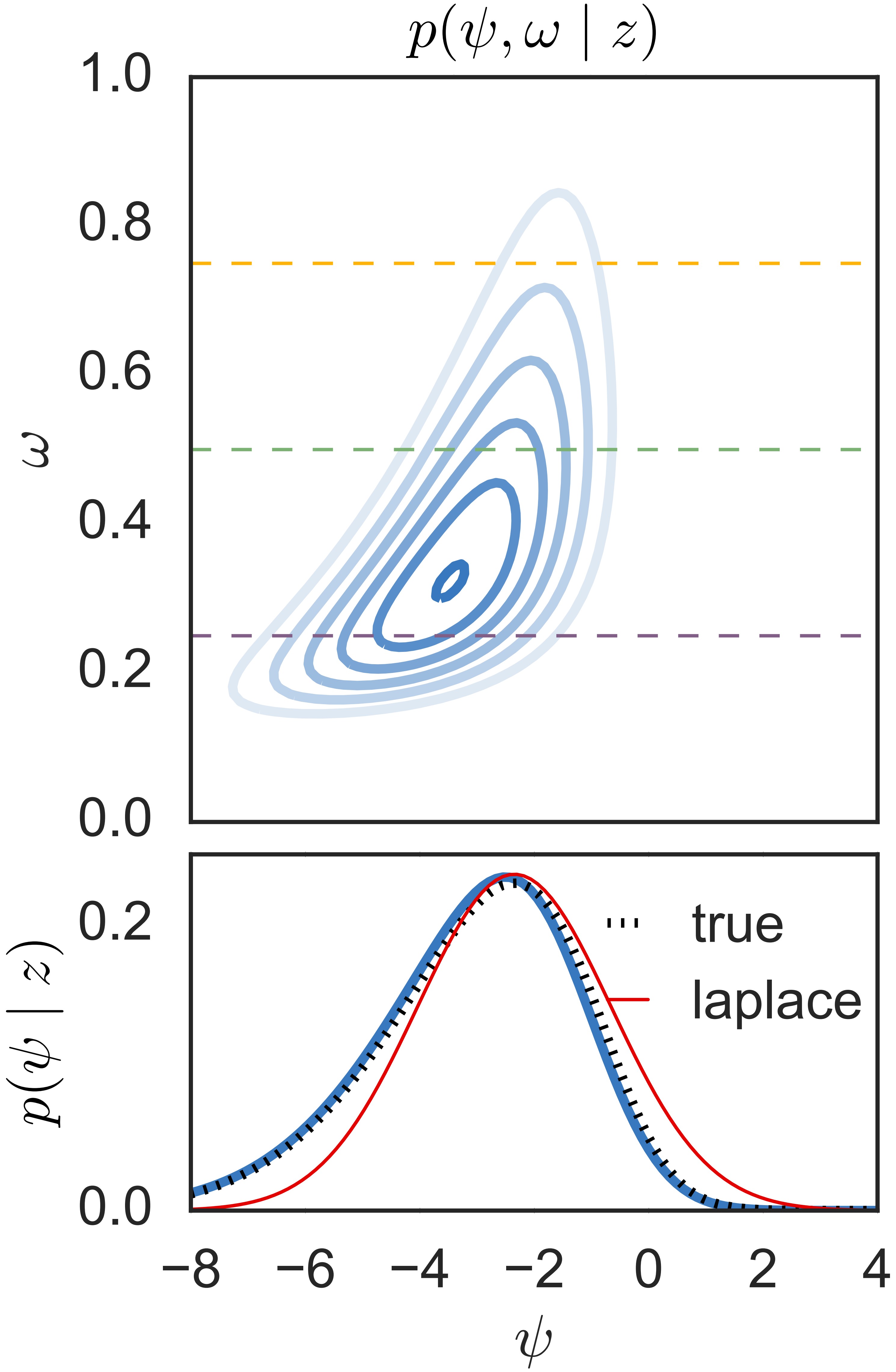
Polson, Scott, and Windle, 2013
Linderman**, Johnson*, and Adams, 2015*
https://github.com/slinderman/pypolyagamma

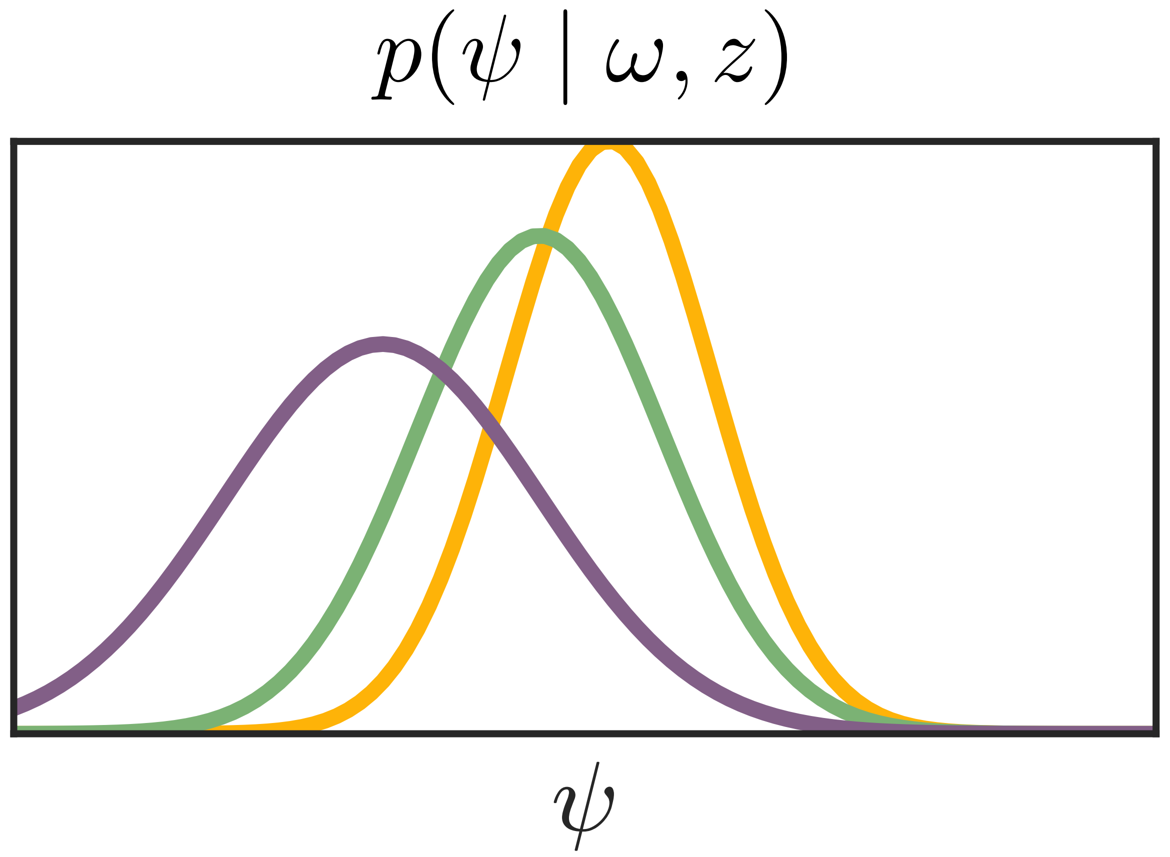
Augmentation schemes for GP Classification with a logit link function



Polson, Scott, and Windle, 2013
Linderman**, Johnson*, and Adams, 2015*
https://github.com/slinderman/pypolyagamma

Augmentation schemes for GP Classification with a logit link function





Polson, Scott, and Windle, 2013
Linderman**, Johnson*, and Adams, 2015*
https://github.com/slinderman/pypolyagamma

Pólya-gamma augmentation

Integral identity:

$$\frac{(e^{\psi})^a}{(1+e^{\psi})^b} = 2^{-b} e^{\kappa\psi} \int_0^{\infty} e^{-\omega\psi^2/2} p_{\mathrm{PG}}(\omega \,|\, b, 0) \, \mathrm{d}\omega$$

where

$$\kappa \triangleq a - b/2$$

Polson, Scott, and Windle, 2013

Pólya-gamma augmentation

Likelihood:

$$p(z \mid \psi) = c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}}$$

Polson, Scott, and Windle, 2013

Pólya-gamma augmentation

Likelihood:

$$p(z \mid \psi) = c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}}$$

Use integral identity to define auxiliary variables:

$$p(z, \psi) = p(\psi) c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}}$$

Polson, Scott, and Windle, 2013

Pólya-gamma augmentation

Likelihood:

$$p(z \,|\, \psi) = c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}}$$

Use integral identity to define auxiliary variables:

$$\begin{aligned} p(z, \psi) &= p(\psi) c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}} \\ &= \int_0^\infty p(\psi) c(z) 2^{-b(z)} e^{\kappa(z) - \omega \psi^2 / 2} p_{\mathrm{PG}}(\omega \,|\, b(z), 0) \, \mathrm{d}\omega \end{aligned}$$

Polson, Scott, and Windle, 2013

Pólya-gamma augmentation

Likelihood:

$$p(z \mid \psi) = c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}}$$

Use integral identity to define auxiliary variables:

$$\begin{aligned} p(z, \psi) &= p(\psi) c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}} \\ &= \int_0^\infty p(\psi) c(z) 2^{-b(z)} e^{\kappa(z) - \omega \psi^2 / 2} p_{\mathrm{PG}}(\omega \mid b(z), 0) \, \mathrm{d}\omega \\ &\triangleq \int_0^\infty p(z, \psi, \omega) \, \mathrm{d}\omega \end{aligned}$$

Polson, Scott, and Windle, 2013

Pólya-gamma augmentation

Likelihood:

$$p(z \mid \psi) = c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}}$$

Use integral identity to define auxiliary variables:

$$\begin{aligned}
p(z, \psi) &= p(\psi) c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}} \\
&= \int_0^\infty \underline{p(\psi)} c(z) 2^{-b(z)} \underline{e^{\kappa(z) - \omega \psi^2 / 2}} p_{\mathrm{PG}}(\omega \mid b(z), 0) \, \mathrm{d}\omega \\
&\triangleq \int_0^\infty p(z, \psi, \omega) \, \mathrm{d}\omega
\end{aligned}$$

Polson, Scott, and Windle, 2013

Pólya-gamma augmentation

Likelihood:

$$p(z \,|\, \psi) = c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}}$$

Use integral identity to define auxiliary variables:

$$\begin{aligned} p(z, \psi) &= p(\psi) c(z) \frac{(e^{\psi})^{a(z)}}{(1 + e^{\psi})^{b(z)}} \\ &= \int_0^\infty \underline{p(\psi)} c(z) 2^{-b(z)} \underline{e^{\kappa(z) - \omega \psi^2 / 2}} \, \underline{p_{\mathrm{PG}}(\omega \,|\, b(z), 0)} \, \mathrm{d}\omega \\ &\triangleq \int_0^\infty p(z, \psi, \omega) \, \mathrm{d}\omega \end{aligned}$$

Polson, Scott, and Windle, 2013

Discrete count likelihoods of this form

$$\mathrm{Bern}(z \mid \sigma(\psi)) = \sigma(\psi)^z (1 - \sigma(\psi))^{1-z} = \frac{(e^\psi)^z}{1 + e^\psi}$$

where $\sigma(\psi) = \dfrac{e^\psi}{1 + e^\psi}$

$$\mathrm{Bin}(z \mid N, \sigma(\psi)) \propto \sigma(\psi)^z (1 - \sigma(\psi))^{N-z} = \frac{(e^\psi)^z}{(1 + e^\psi)^N}$$

$$\mathrm{NB}(z \mid N, \sigma(\psi)) \propto \sigma(\psi)^z (1 - \sigma(\psi))^N = \frac{(e^\psi)^z}{(1 + e^\psi)^{N+z}}$$

Discrete count likelihoods of this form

$$\mathrm{Bern}(z \,|\, \sigma(\psi)) = \sigma(\psi)^z (1 - \sigma(\psi))^{1-z} = \frac{(e^\psi)^z}{1 + e^\psi}$$

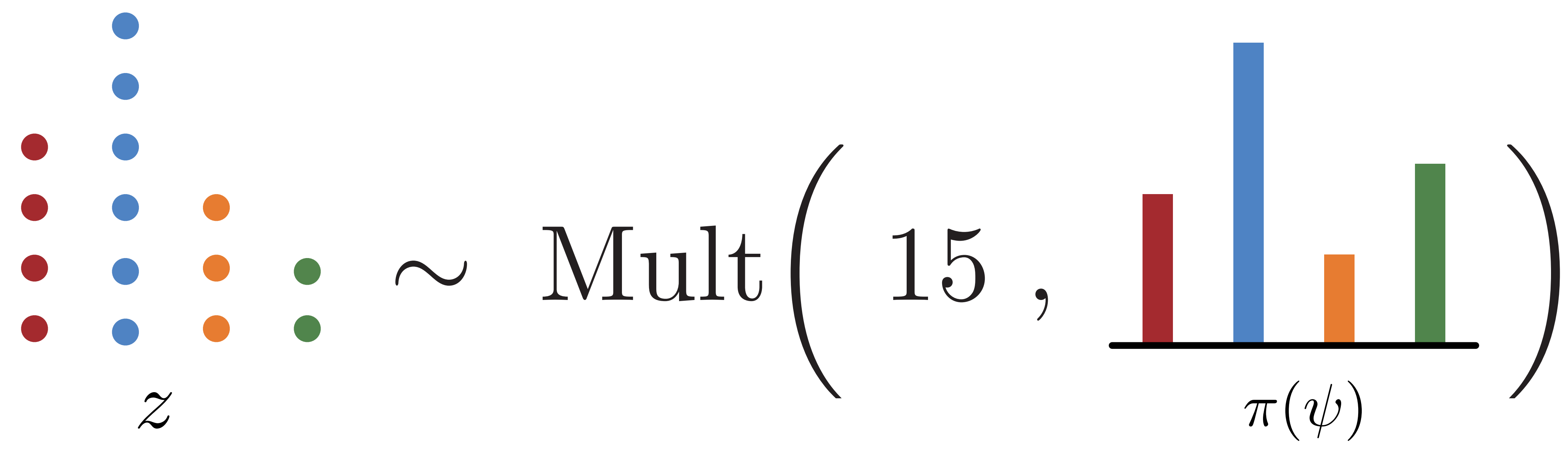
where $\sigma(\psi) = \dfrac{e^\psi}{1 + e^\psi}$

$$\mathrm{Bin}(z \,|\, N, \sigma(\psi)) \propto \sigma(\psi)^z (1 - \sigma(\psi))^{N-z} = \frac{(e^\psi)^z}{(1 + e^\psi)^N}$$

$$\mathrm{NB}(z \,|\, N, \sigma(\psi)) \propto \sigma(\psi)^z (1 - \sigma(\psi))^{N} = \frac{(e^\psi)^z}{(1 + e^\psi)^{N+z}}$$

What about **categorical** or **multinomial** data?

Pólya-gamma augmentation for multinomials



Probability mass function:

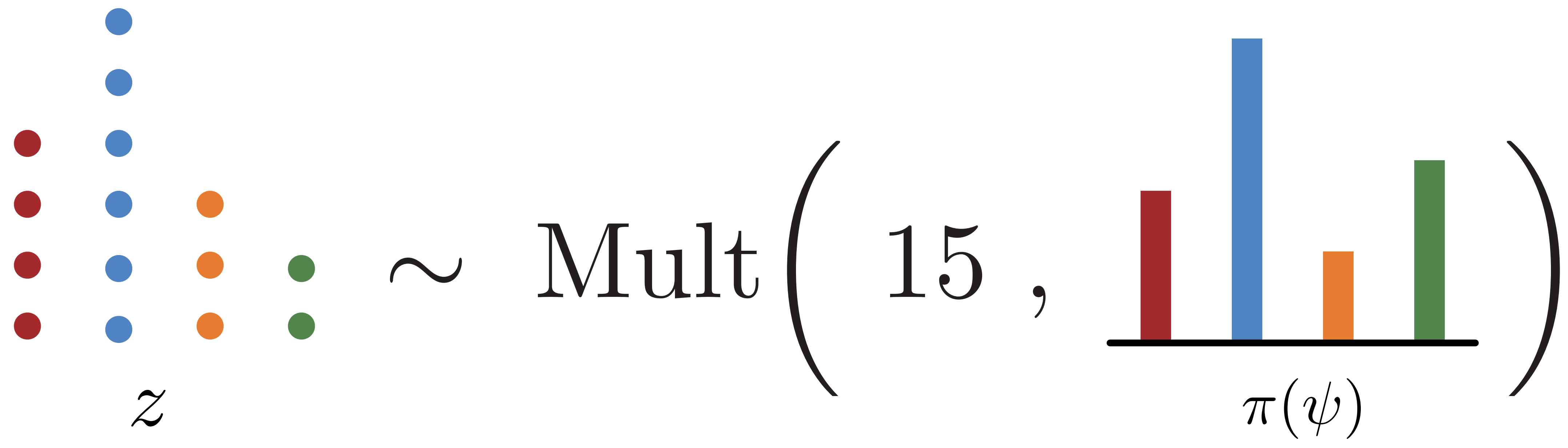
$$p(z \,|\, \psi) = c(z) \prod_k \pi_k(\psi)^{z_k}$$

Rewrite it as:

$$p(z \,|\, \psi) = c(z) \prod_k \sigma(\psi_k)^{a_k(z)} (1 - \sigma(\psi_k))^{b_k(z)}$$

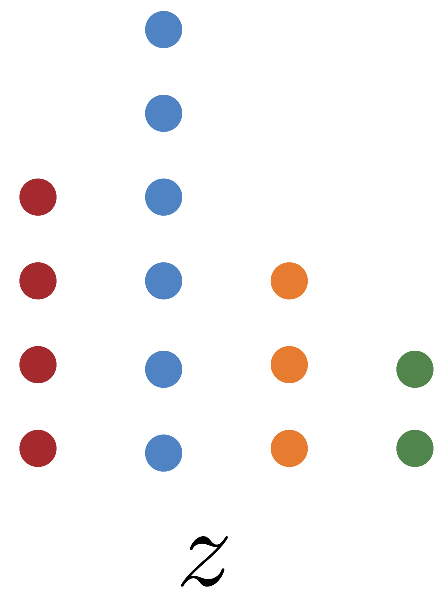
Logistic-normal doesn't suffice.

Reducing multinomial to sequence of binomials

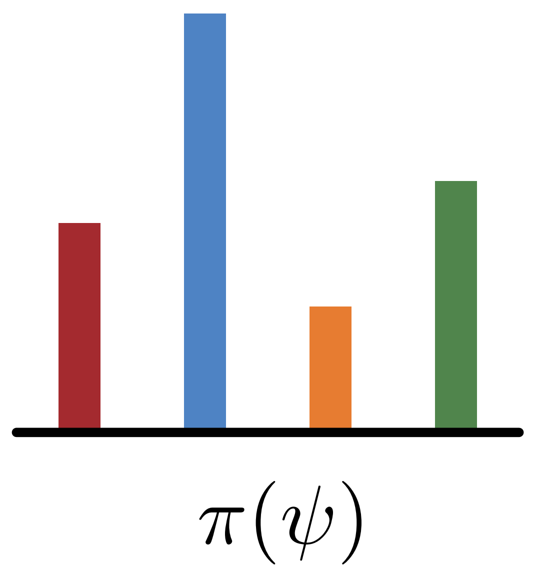


Linderman, Johnson*, and Adams. NIPS (2*

Reducing multinomial to sequence of binomials



$\sim \mathrm{Mult}\Big(15 ,$

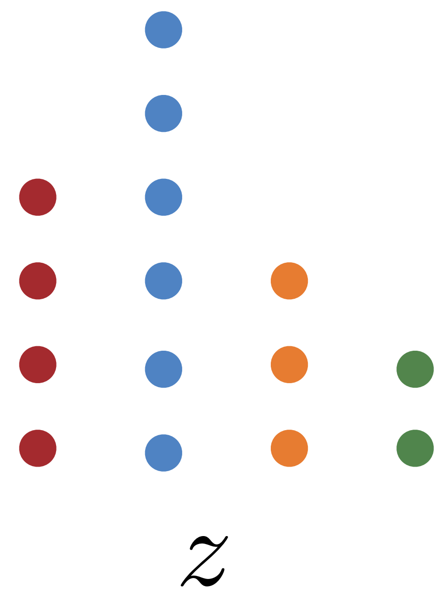


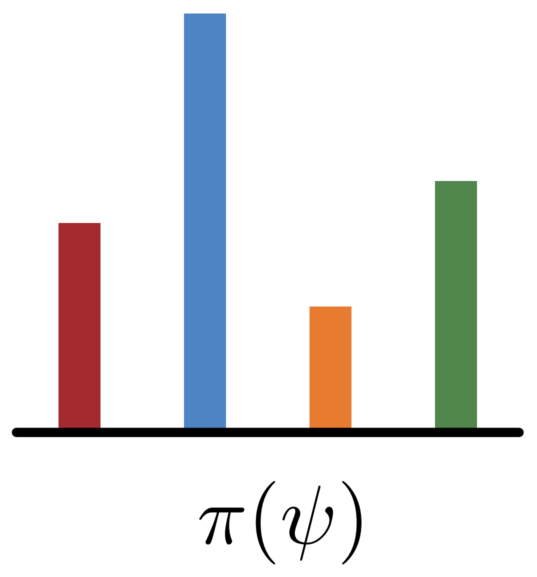
$\Big)$

$$z_1 \sim \mathrm{Bin}\,(15, \pi_1)$$

Linderman, Johnson*, and Adams. NIPS (2*

Reducing multinomial to sequence of binomials



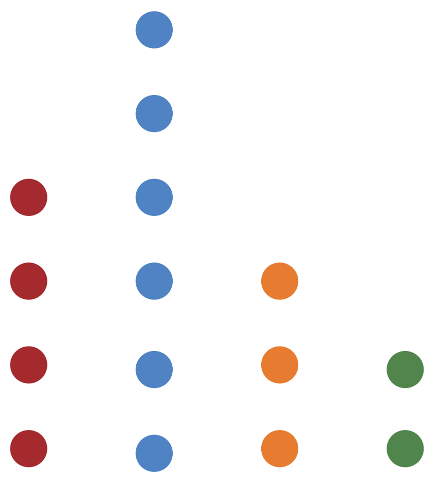
$\sim \mathrm{Mult}\big(15,$  $\big)$

$$z_1 \sim \mathrm{Bin}\,(15, \pi_1)$$

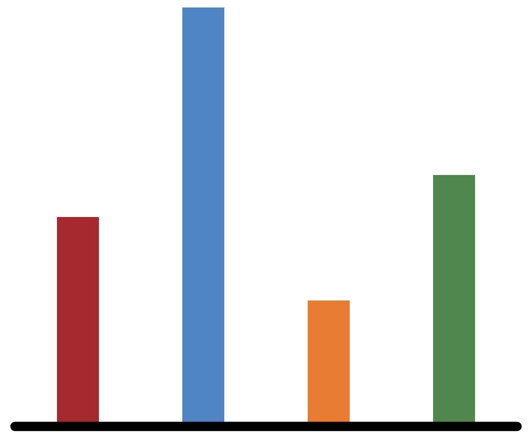
$$z_2 \mid z_1 \sim \mathrm{Bin}\left(15 - z_1, \frac{\pi_2}{1 - \pi_1}\right)$$

Linderman*, Johnson*, and Adams. NIPS (2

Reducing multinomial to sequence of binomials



$z \sim \mathrm{Mult}\left(15, \pi(\psi)\right)$



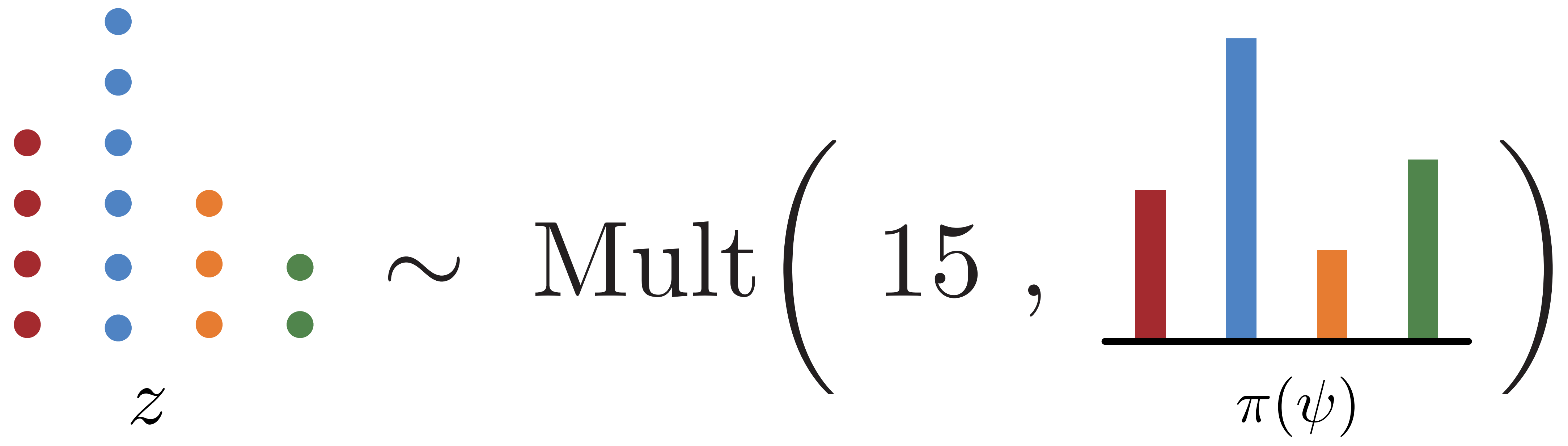
$$z_1 \sim \mathrm{Bin}\left(15, \pi_1\right)$$

$$z_2 \mid z_1 \sim \mathrm{Bin}\left(15 - z_1, \frac{\pi_2}{1 - \pi_1}\right)$$

$$z_3 \mid z_2, z_1 \sim \mathrm{Bin}\left(15 - z_1 - z_2, \frac{\pi_3}{1 - \pi_1 - \pi_2}\right)$$

Linderman, Johnson*, and Adams. NIPS (2*

Reducing multinomial to sequence of binomials



$$z_1 \sim \mathrm{Bin}\left(15, \pi_1\right)$$

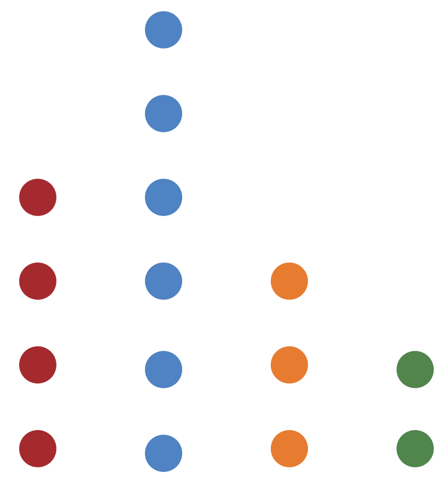
$$z_2 \mid z_1 \sim \mathrm{Bin}\left(15 - z_1, \frac{\pi_2}{1 - \pi_1}\right)$$

$$z_3 \mid z_2, z_1 \sim \mathrm{Bin}\left(15 - z_1 - z_2, \frac{\pi_3}{1 - \pi_1 - \pi_2}\right)$$

$$z_4 = 15 - z_1 - z_2 - z_3$$

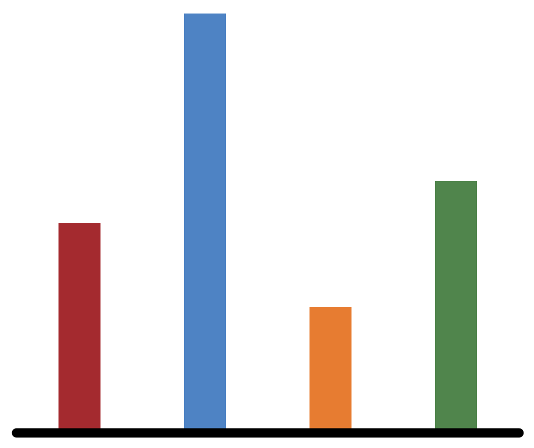
Linderman, Johnson*, and Adams. NIPS (2*

Reducing multinomial to sequence of binomials



z

$\sim \mathrm{Mult}\Big(\, 15 \,,$



$\pi(\psi)$

$\Big)$

$$z_1 \sim \mathrm{Bin}\left(15, \underbrace{\pi_1}_{\sigma(\psi_1)}\right)$$

$$z_2 \mid z_1 \sim \mathrm{Bin}\left(15 - z_1, \underbrace{\frac{\pi_2}{1 - \pi_1}}_{\sigma(\psi_2)}\right)$$

$$z_3 \mid z_2, z_1 \sim \mathrm{Bin}\left(15 - z_1 - z_2, \underbrace{\frac{\pi_3}{1 - \pi_1 - \pi_2}}_{\sigma(\psi_3)}\right)$$

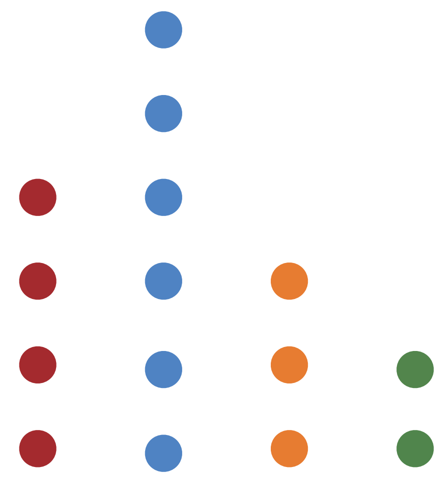
$$z_4 = 15 - z_1 - z_2 - z_3$$

logistic stick-breaking transformation

$\mathbb{R}^{K-1} \to \Delta_{K-1}$

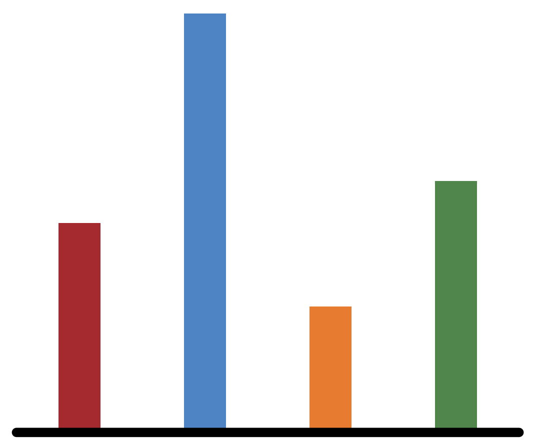
Linderman, Johnson*, and Adams. NIPS (2*

Reducing multinomial to sequence of binomials



z

$\sim \mathrm{Mult}\Big(15 \,,$



$\pi(\psi)$

$\Big)$

logistic stick-breaking transformation

$\mathbb{R}^{K-1} \to \Delta_{K-1}$

$$z_1 \sim \mathrm{Bin}\,(15, \pi_1)$$

$\sigma(\psi_1)$

$$z_2 \,|\, z_1 \sim \mathrm{Bin}\left(15 - z_1, \frac{\pi_2}{1-\pi_1}\right)$$

$\sigma(\psi_2)$

$$z_3 \,|\, z_2, z_1 \sim \mathrm{Bin}\left(15 - z_1 - z_2, \frac{\pi_3}{1-\pi_1-\pi_2}\right)$$

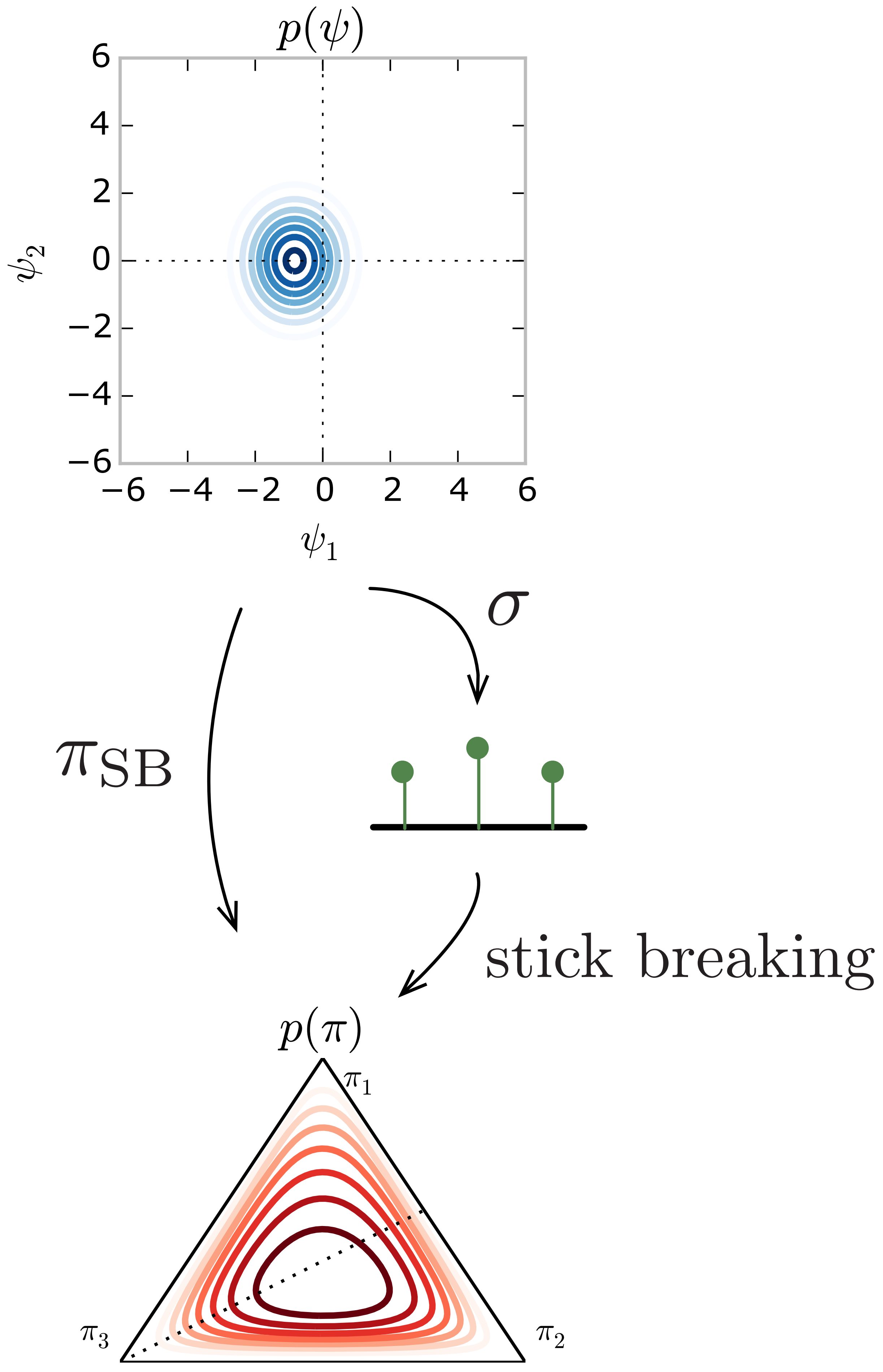
$\sigma(\psi_3)$

$$z_4 = 15 - z_1 - z_2 - z_3$$

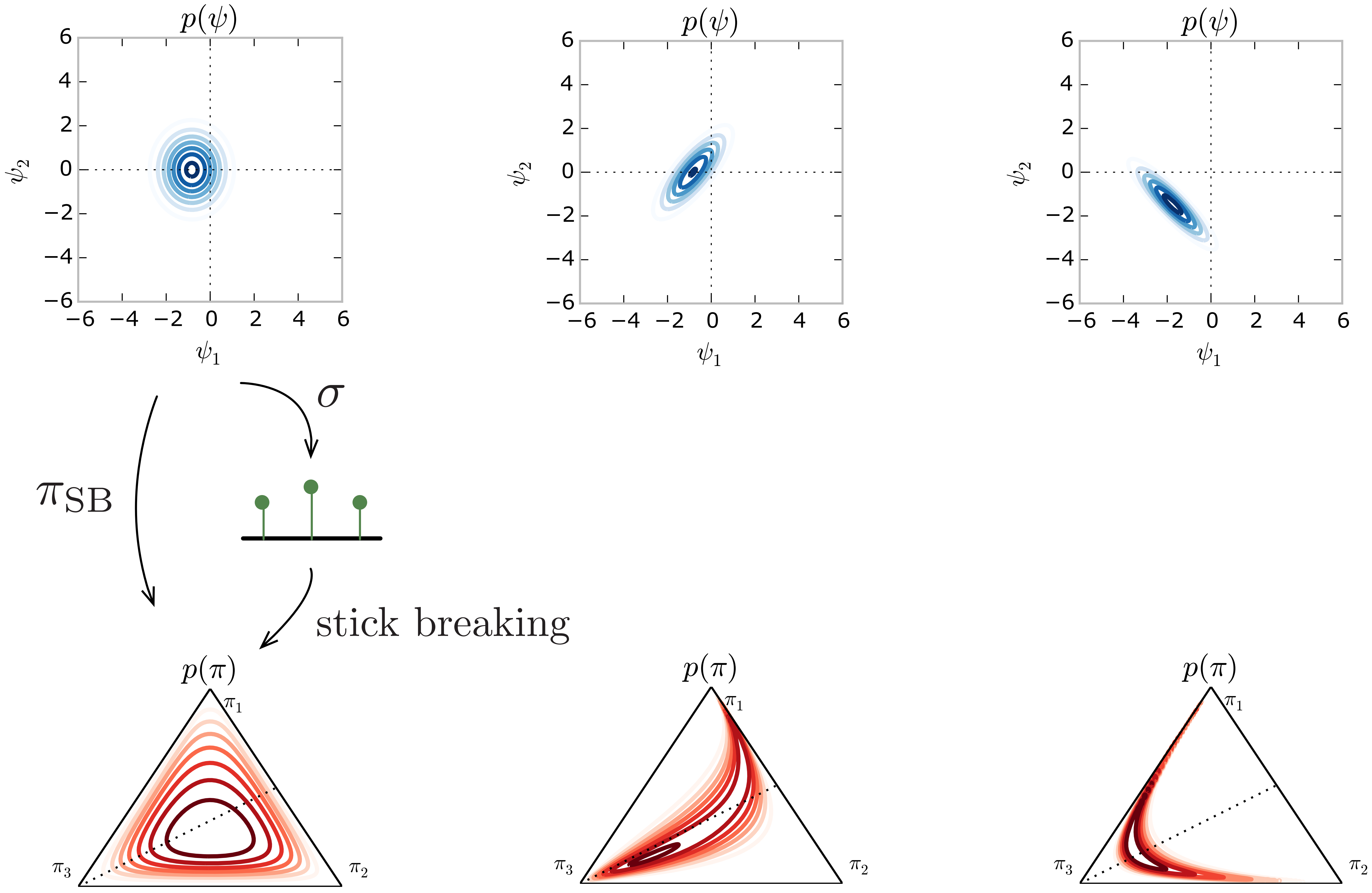
$$p(z \,|\, \psi) = \prod_{k=1}^{K-1} \mathrm{Bin}\,(z_k \,|\, N_k(z_{1:k-1}), \sigma(\psi_k))$$

Linderman, Johnson*, and Adams. NIPS (2*

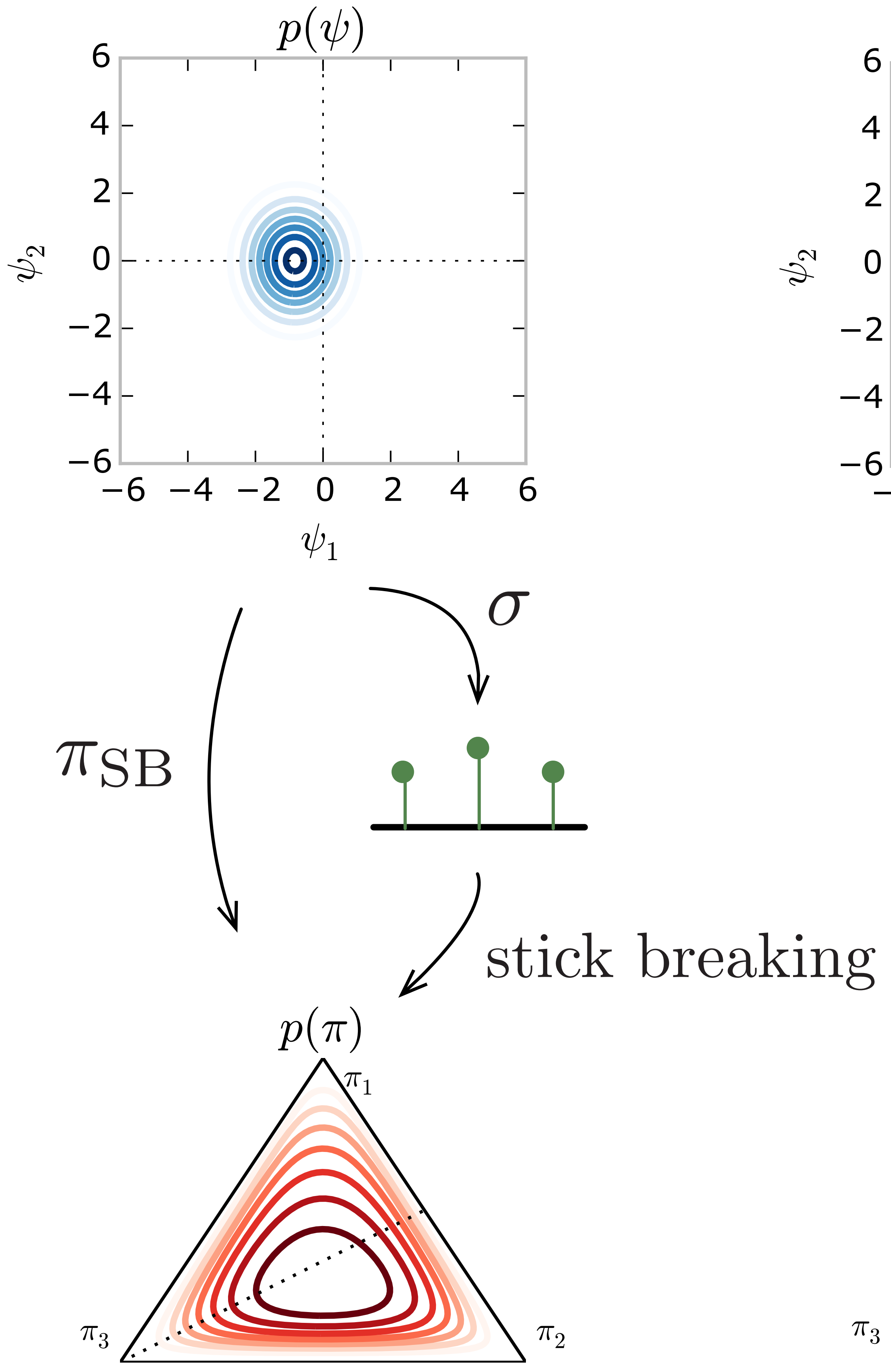
Correlated distributions on the simplex



Correlated distributions on the simplex



Correlated distributions on the simplex



Other Applications:
- correlated topic models
- dynamic topic models
- multinomial GP's
- multinomial time series

Linderman, Johnson*, and Adams. NIPS (2015)*
Elibol, …, Linderman, …, Doshi-Velez. JMLR (2016)