Gibbs Sampling version of SSVI.

① Algorithm : Gibbs Sampling

- initialize $W, H, S^W, \Pi^W, S^H, \pi^H$
- for $j = 1, 2, \ldots, \text{burnin}, \ldots$ do:
  - Sample S^W_{fk} } using collapsed gibbs sampling derived before
  - Sample S^H_{kt}
  - $\rightarrow$ compute $\phi_{ftk} = \dfrac{W_{fk} S^W_{fk} H_{kt} S^H_{kt}}{\sum_l W_{fl} S^W_{fl} H_{lt} S^H_{lt}}$
  - sample $W_{fk} \sim \text{Gamma}\left(a + \sum_t X_{ft}\phi_{ftk},\ b + S^W_{fk} \sum_t H_{kt} S^H_{kt}\right)$
  - sample $H_{kt} \sim \text{Gamma}\left(c + \sum_f X_{ft}\phi_{ftk},\ d + S^H_{kt} \sum_f W_{fk} S^W_{fk}\right)$
  - sample $\pi^H_k \sim \text{Beta}\left(\frac{a_0^H}{K} + \sum_t S^H_{kt},\ \frac{b_0^H (K-1)}{K} + T - \sum_t S^H_{kt}\right)$
  - sample $\pi^W_k \sim \text{Beta}\left(\frac{a_0^W}{K} + \sum_f S^W_{fk},\ \frac{b_0^W (K-1)}{K} + F - \sum_f S^W_{fk}\right)$
- end for

— Derivations

- $p(W_{fk} \mid X_f, Z_f, \ldots) \propto p(W_{fk}, X, Z, S^W, S^H, H, \pi^W, \pi^H)$

$$\propto \underset{\substack{\uparrow \\ \text{gamma prior}}}{p(W_{fk})} \cdot \underset{\substack{\uparrow \\ \text{poisson}}}{p(Z_f \mid W_f, H, S^W_f, S^H)}$$

see SSVI derivation

$$\vdots$$

$$= \Gamma\left(a + \sum_t \cdots,\ b + S^W_{fk} \sum_t \cdots\right)$$

basically all derivations are like SSVI, just have prior in front.
(for $p(H_{kt} \mid \cdots)$, $p(\pi^H_k \mid \cdots)$, $p(\pi^W_k \mid \cdots)$)