Bayesian Inference for Pólya Inverse Gamma Models

Jingyu He
Booth School of Business
University of Chicago

Nicholas Polson
Booth School of Business
University of Chicago

Jianeng Xu
Booth School of Business
University of Chicago

First Draft: August 2018
This Draft: September 2018

Abstract

Pólya-inverse Gamma (P-IG) distributions arise in high dimensional statistics and machine learning problems. They are constructed as infinite convolutions of reciprocal gamma distributions. Exponential Reciprocal Gamma (E-RG) distributions are a special case that apply to Dirichlet allocation, non-parametric Bayes and Gamma inference problems. Scalable fast sampling algorithms are available using simulation methods based on infinitely divisible distributions. To illustrate our methodology, we analyze a text analysis data from Zillow and a Gamma scale inference problem. Finally, we conclude with directions for future research.

Key Words: Pólya inverse Gamma, Logistic, GGC, Pólya Gamma, Exponential reciprocal Gamma, Latent Dirichlet allocation, Topic models.

1 Introduction

Pólya-inverse Gamma (P-IG) distributions arise as infinite convolutions of reciprocal gammas and are central to Bayesian inference in Dirichlet allocation and Gamma inference problems. P-IG distributions mimic the construction of Polya-Gamma distributions (Polson et al., 2013) which are central to logistic regression and classification problems. They arise as conditional distributions in MCMC sampling from a variety of models; such as latent Dirichlet allocation (Blei et al., 2003), topic models (Blei and Lafferty, 2006), non-parametric Bayes (Escobar and West, 1995), scale Gamma inference (Rossell et al., 2009). The goal of our paper is to provide the necessary theoretical and simulation algorithms for the implementation of P-IG distributions. Our approach will make use of simulation methods for infinitely divisible distributions.

Exponential Reciprocal Gamma (E-RG) distributions are a sub-class with direct application to scale Gamma inference problems, see Damsleth (1975); Bondesson (1992); Damien et al. (1995); Rossell et al. (2009); Miller (2018). Fast and scalable algorithms are developed that are directly tailored for P-IG distributions. Our approach builds on results of Hartman (1976) and Roynette and Yor (2005) who show that the exponential reciprocal gamma distribution can be represented as a scale mixture of normals, a class that has a long history as a latent variable representations in Bayesian inference, see Andrews and Mallows (1974), West (1987), Carlin and Polson (1991), Polson and Scott (2011) and Polson and Scott (2013).

To simulate from a P-IG conditional distribution we use results from infinitely divisible processes and Lévy densities, see Ferguson and Klass (1972); Bondesson (1982); Damien et al. (1995); Walker and Damien (2000). There are equivalent constructed scalable PG sampling, see Windle et al. (2014) and Glynn et al. (2019). To illustrate our methodology, we provide an application to Dirichlet-multinomal models (Glynn et al., 2019) and to Gamma scale inference (West, 1992; Miller, 2018). Our O-IG approach provides full posterior distribution and thus allows us to compare with approximate model algorithms proposed by Minka (2000) and the simulation approach of Miller (2018).

The rest of our paper is outlined as follows. Section 2 defines the class of Pólya-inverse Gamma distribution. Section 3 defines the Exponential reciprocal Gamma distribution with particular emphasis on applications to Dirichlet-Multinomial allocation and to scale inference in Gamma models. Section 4 provides an application to topic modeling and text analysis with Zillow data. Finally, Section 5 concludes with directions for future research.

2 Pólya-Inverse Gamma (P-IG) Models

The Pólya-inverse Gamma distribution P-IG$(d, 0)$ with density $p(\omega \mid \mathbf{d}, 0)$ is best defined by its Laplace transform (LT),

$$E\left[e^{-\omega t}\right] = \int_0^\infty e^{-\omega t} p(\omega \mid \mathbf{d}, 0) d\omega = \prod_{k=1}^{\infty} \left(1 + \frac{t}{d_k}\right) e^{-t/d_k}, \tag{1}$$

where $\mathbf{d} = (d_1, d_2, \cdots) > 0$ is a sequence of given positive constants, write $\omega \overset{D}{=}$ P-IG$(\mathbf{d}, 0)$.

This distribution can be extended to the P-IG$(\mathbf{d}, c)$ distribution of the form which is defined as an exponential tilt of the P-IG$(\mathbf{d}, 0)$ density, of the form

$$p(\omega \mid \mathbf{d}, c) = \frac{\exp\left(-\frac{c^2}{2}\omega\right) p(\omega \mid \mathbf{d}, 0)}{E_\omega\left[\exp\left(-\frac{c^2}{2}\omega\right)\right]}. \tag{2}$$

The normalizing constant $E_\omega\left[\exp\left(-\frac{c^2}{2}\omega\right)\right]$ is calculated using the Laplace transform (1) as,

$$E_\omega\left[e^{-\frac{c^2}{2}\omega}\right] = \prod_{k=1}^{\infty} \left(1 + \frac{c^2}{2d_k}\right) e^{-\frac{c^2}{2d_k}}. \tag{3}$$

The following theorem characterizes the LT of the $p(\omega \mid \mathbf{d}, c)$ which proves useful in identifying the conditional posterior and providing simulation strategies based on GGC distribution (Bondesson, 1992).

Theorem 1. *The Laplace transform of P-IG$(\mathbf{d}, c)$ can be calculated from that of P-IG$(\mathbf{d}, 0)$ with the form*

$$E_{\omega|\mathbf{d},c}\left[e^{-t\omega}\right] = \prod_{k=1}^{\infty} \left(1 + \frac{t}{2d_k + c^2}\right) e^{-\frac{t}{2d_k + c^2}} e^{-t\left(\frac{1}{2d_k} - \frac{1}{2d_k + c^2}\right)}, \tag{4}$$

where δ denotes a Dirac measure. Hence, the P-IG$(\mathbf{d}, c)$ distribution is

$$\begin{aligned} \omega \mid \mathbf{d}, c &\overset{D}{=} \text{P-IG}(\mathbf{d} + c^2, 0) + \sum_{k=1}^{\infty} \delta\left(\frac{1}{d_k} - \frac{1}{d_k + c^2}\right) \\ \text{P-IG}(\mathbf{d} + c^2, 0) &\overset{D}{=} \sum_{k=1}^{\infty} \frac{Ga^{-1}(3/2)}{d_k + c^2} \end{aligned} \tag{5}$$

$$
\begin{aligned}
\omega \mid \mathbf{d}, c &\overset{D}{=} P\text{-}IG(\mathbf{d}+c^2, 0) + \sum_{k=1}^{\infty} \delta\left(\frac{1}{d_k} - \frac{1}{d_k+c^2}\right) \\
P\text{-}IG(\mathbf{d}+c^2, 0) &\overset{D}{=} \sum_{k=1}^{\infty} \frac{Ga^{-1}(3/2)}{2d_k+c^2} \neq \sum_{k=1}^{\infty} \frac{Ga^{-1}(3/2)}{2(d_k+c^2)}
\end{aligned} \tag{6}
$$

No additive of $\mathbf{d}+c^2$ here if we put it on denominator

Proof. The Laplace transform $E_{\omega|\mathbf{d},c}\left[e^{-t\omega}\right]$ can be directly calculated as follows

$$
\begin{aligned}
E_{\omega|\mathbf{d},c}\left[e^{-t\omega}\right] &= \frac{E_{\omega|\mathbf{d},0}\left[\exp\left(-\frac{c^2+t}{2}\omega\right)\right]}{E_{\omega}\left[\exp\left(-\frac{c^2}{2}\omega\right)\right]} = \frac{\prod_{k=1}^{\infty}\left(1+\frac{t+c^2}{d_k}\right)e^{-\frac{t+c^2}{d_k}}}{\prod_{k=1}^{\infty}\left(1+\frac{c^2}{d_k}\right)e^{-\frac{c^2}{d_k}}} \\
&= \prod_{k=1}^{\infty}\left(1+\frac{t}{d_k+c^2}\right)e^{-\frac{t}{d_k}} \\
&= \prod_{k=1}^{\infty}\left(1+\frac{t}{d_k+c^2}\right)e^{-\frac{t}{d_k+c^2}}e^{-t\left(\frac{1}{d_k}-\frac{1}{d_k+c^2}\right)}
\end{aligned} \tag{7}
$$

Should it be

$$
\begin{aligned}
E_{\omega|\mathbf{d},c}\left[e^{-t\omega}\right] &= \frac{E_{\omega|\mathbf{d},0}\left[\exp\left(-\frac{c^2+t}{2}\omega\right)\right]}{E_{\omega}\left[\exp\left(-\frac{c^2}{2}\omega\right)\right]} = \frac{\prod_{k=1}^{\infty}\left(1+\frac{t+c^2}{2d_k}\right)e^{-\frac{t+c^2}{2d_k}}}{\prod_{k=1}^{\infty}\left(1+\frac{c^2}{2d_k}\right)e^{-\frac{c^2}{2d_k}}} \\
&= \prod_{k=1}^{\infty}\left(1+\frac{t}{2d_k+c^2}\right)e^{-\frac{t}{2d_k}} \\
&= \prod_{k=1}^{\infty}\left(1+\frac{t}{2d_k+c^2}\right)e^{-\frac{t}{2d_k+c^2}}e^{-t\left(\frac{1}{2d_k}-\frac{1}{2d_k+c^2}\right)}
\end{aligned} \tag{8}
$$

□

2.1 Remarks

Remark 1. *Inspired by Laplace transform of Pólya-Gamma (PG) distribution, which has Hadamard factorization $\prod_{k=1}^{\infty}\left(1+d_k^{-1}t\right)^{-b}$ and corresponds to an infinite convolution of Gammas, the P-IG factorization is of Weierstrass for*

$$
E\left[e^{-\omega t}\right] = \prod_{k=1}^{\infty}\left(1+\frac{t}{d_k}\right)e^{-t/d_k} \tag{9}
$$

Remark 2. *Let* $\omega \sim 1/(\gamma^{3/2})$ *(inverse Gamma* $(\frac{3}{2}, 1)$*) with density*

$$p(\omega) = \frac{1}{a^3\sqrt{2\pi\omega^5}} e^{-\frac{1}{2a^2\omega}}, \tag{10}$$

note that

$$E\left[\exp\left(-\frac{t^2}{a}\omega\right)\right] = \left(1+\frac{t}{a}\right)e^{-\frac{t}{a}}, \quad \frac{d}{d\lambda}E[e^{-\lambda^2/4\gamma^{3/2}}] = -\lambda e^{-\lambda} \tag{11}$$

See Appendix A for the derivation.

Hence random variables with $E(-e^{\frac{t^2}{2}\Omega})$ being equal to Equation 9 can be simulated via

$$\omega \stackrel{D}{=} \sum_{k=1}^{\infty} \frac{Ga^{-1}(3/2)}{d_k}, \tag{12}$$

which is an infinity sum of inverse gamma distribution $Ga^{-1}\left(\frac{3}{2}, 1\right)$. By Bondesson (1982), we can truncate the series at T and use a normal approximation for the tail.

3 Exponential Reciprocal Gamma (E-RG) Models

The Exponential-Reciprocal Gamma (E-RG) distribution is as a special case of a Pólya-inverse Gamma distribution. We write $\omega \stackrel{D}{=} \text{E-RG}(a)$ for $a \geq 0$ if its Laplace transform is given by a vector of Gamma functions

$$E_{\omega|a}\left[e^{-\omega t}\right] = \frac{\Gamma(a)}{\Gamma(a+t)}, \quad t > 0. \tag{13}$$

The Hadamard factorization is

$$\frac{e^{-\gamma t}}{\Gamma(t+1)} = \prod_{k=1}^{\infty}\left(1+\frac{t}{k}\right)e^{-\frac{t}{k}}. \tag{14}$$

The Hadamard-Weierstrass factorization of the reciprocal Gamma distribution is given by

$$\frac{\Gamma(a)}{\Gamma(a+t)} = e^{-\psi(a)t}\prod_{k=0}^{\infty}\left(1+\frac{t}{a+k}\right)e^{-\frac{t}{a+k}} \tag{15}$$

where $\psi(a)$ is digamma function. Hence an equivalent definition is

$$E_{\omega|a}\left[e^{-\omega t}\right] = e^{-\psi(a)t}\prod_{k=0}^{\infty}\left(1+\frac{t}{a+k}\right)e^{-\frac{t}{a+k}}. \tag{16}$$

This coincides with equation (14) with $\psi(1) = -\gamma$ and $a = 1$. (J: constant $\Gamma(a)$ missing?)

Hartman (1976) and Roynette and Yor (2005) discuss the scale mixture of normals representation of $\Gamma(a)/\Gamma(a+s) = E\left[e^{-\omega s^2}\right]$. This is clearly related to the Laplace transform identity in (14). See Stefanski (1991) for further discussion of relationships between scale mixture of normals representations and Laplace transforms.

write

$$\frac{\Gamma(a)}{\Gamma(a+s)} = \int_0^\infty e^{-\frac{1}{2}s^2\omega} p(\omega \mid a) d\omega \tag{17}$$

as a scale mixture of normals

3.1 Bayes Inference for a Gamma Scale Model

need work

The Gibbs sampler is

1. $s \mid a, \omega \sim$Normal
2. $\omega \mid s, a \sim \frac{e^{-s^2\omega} p_a(\omega)}{\Gamma(a)}$

$$\begin{aligned}
\int_0^\infty e^{-\omega t} \frac{e^{-s^2\omega} p_a(\omega)}{p(a \mid e^{s\psi(a)})/\Gamma(a+s)} &= \int_0^\infty e^{-(s^2+t)\omega} \frac{p_a(\omega)}{p(a \mid e^{s\psi(a)})} \Gamma(a+s) d\omega \\
&= \frac{\Gamma(a) e^{-\sqrt{s^2+t^2}\psi(a)}}{\Gamma(a+\sqrt{s^2+t^2})} \frac{\Gamma(a+s)}{\Gamma(a) e^{-\psi(a)}} \\
&= \frac{\Gamma(a+s)}{\Gamma(a+\sqrt{s^2+t^2})} e^{-\sqrt{s^2+t^2}} \\
&= \prod_{k=1}^\infty \left(1 + \frac{s}{d_k}\right) e^{-s/d_k}
\end{aligned} \tag{18}$$

Notice that

$$\int_0^\infty \frac{1}{\Gamma(a+t)} dt < \infty \quad \forall a \geq 0. \tag{19}$$

3.2 Latent Dirichlet Allocation

Suppose $\mathbf{n} = (n_1, \cdots, n_K)$. Let $N = \sum_k n_k$

$$p(\mathbf{n} \mid \alpha) = \frac{\Gamma\left(\sum_k \alpha_k\right)}{\Gamma\left(\sum_k n_k + \alpha_k\right)} \prod_k \frac{\Gamma\left(n_k + \alpha_k\right)}{\Gamma\left(\alpha_k\right)} \tag{20}$$

Need to learn α,

$$\frac{e^{-\gamma\alpha}}{\Gamma\left(\alpha+1\right)} = e^{-s\psi(\alpha)} \prod_{k=0}^{\infty}\left(1+\frac{s}{a+k}\right)e^{-\frac{s}{a+k}} \tag{21}$$

where $\psi(\alpha)$ is digamma function and $\psi(1) = -\gamma$.

$$\frac{\Gamma\left(\alpha\right)}{\Gamma\left(a+s\right)} = e^{-s\psi(\alpha)} \prod_{k=0}^{\infty}\left(1+\frac{s}{a+k}\right)e^{-\frac{s}{a+k}} \tag{22}$$

then

$$\frac{\Gamma\left(\alpha\right)}{\Gamma\left(\alpha+n\right)} = \frac{\left(\alpha+n\right)Beta\left(\alpha+1,n\right)}{\alpha\Gamma\left(n\right)} \tag{23}$$

$$p\left(k \mid \alpha, n\right) = \binom{n}{k} n!\alpha^{k}\frac{\Gamma\left(\alpha\right)}{\Gamma\left(\alpha+n\right)} \tag{24}$$

3.3 Scale Inference of Gamma

$$p(y \mid \beta) = \frac{y^{\beta-1}e^{-y}}{\Gamma(\beta)}, \quad y > 0 \tag{25}$$

Prior

$$p(\beta) = c\beta^{a-1}e^{-\gamma\beta}, \quad \text{where } \gamma = \psi(1), \text{ Euler's const.} \tag{26}$$

By Weierstrass

$$\frac{1}{\Gamma(\beta)} = \beta e^{\gamma\beta}\prod_{j=1}^{\infty}\left(1+\frac{\beta}{j}\right)e^{-\frac{\beta}{j}} \tag{27}$$

Hence, the posterior is

$$p(\beta \mid y) \propto \beta^{a}e^{-\beta(-\log y)}\prod_{j=1}^{\infty}\left(1+\frac{\beta}{j}\right)e^{-\frac{\beta}{j}} \tag{28}$$

where $\prod_{j=1}^{\infty}\left(1+\frac{\beta}{j}\right)e^{-\frac{\beta}{j}}$ is convolution of inverse gammas, $\sum_{j=1}^{\infty}\frac{1/2\gamma_{3/2}}{j}$ need work

$$\frac{b^{a}y^{a-1}}{\Gamma(a)}e^{-by} \tag{29}$$

The posterior is

$$p(a \mid y) \propto \frac{y^{a-1}}{\Gamma(a)}p(a) \tag{30}$$

exponential-reciprocal gamma

$$\frac{1}{\Gamma(a)} = ae^{-\gamma a} \prod_{k=1}^{\infty} \left(1 + \frac{a}{k}\right) e^{-\frac{a}{k}} \tag{31}$$

Hence

$$\begin{aligned} p(a \mid y) &\propto b^a a y^{a-1} e^{-\gamma a} \prod_{k=1}^{\infty} \left(1 + \frac{a}{k}\right) e^{-\frac{a}{k}} \\ &\propto ae^{-a(\gamma - \log y - \log b)} \sum_{k=1}^{\infty} \frac{Ga^{-1}(3/2)}{k} \end{aligned} \tag{32}$$

So the Gibbs sampler is

$$p(a \mid y) \propto ae^{-a(\gamma - \log y - \log b)} E_\omega \left[e^{-a\omega}\right] \tag{33}$$

sample

$$\begin{aligned} &a \mid \omega \sim Gamma(2, \gamma - \log y - \log b) \\ &\omega \mid a \sim \text{P-IG Truncated sum of gammas} \end{aligned} \tag{34}$$

$$\omega = \sum_{k=1}^{\infty} \frac{Ga^{-1}(3/2)}{k} \tag{35}$$

therefore

$$E[e^{-s\omega}] = \prod_{k=1}^{\infty} \left(1 + \frac{s}{k}\right) e^{-\frac{s}{k}} \tag{36}$$

Posterior

$$p(\omega \mid a) = \frac{e^{-a\omega} p(\omega)}{E[e^{-a\omega}]} = \frac{e^{-a\omega} p(\omega)}{\prod_{k=1}^{\infty} \left(1 + \frac{a}{k}\right) e^{-\frac{a}{k}}} \tag{37}$$

The Laplace transform is

$$\begin{aligned} E_{\omega|a}[e^{-s\omega}] &= \int_0^\infty e^{-(a+s)\omega} p(\omega) d\omega / E[e^{-a\omega}] \\ &= \prod_{k=1}^{\infty} \left(\frac{1 + \frac{s+a}{k}}{1 + \frac{a}{k}}\right) e^{-\frac{s}{k}} \\ &= \prod_{k=1}^{\infty} \left(1 + \frac{s}{k+a}\right) e^{-\frac{s}{k+a}} e^{-s\left(\frac{1}{k} - \frac{1}{k+a}\right)} \end{aligned} \tag{38}$$

Hence, the conditional distribution of the latent variable is

$$\omega \mid a = \sum_{k=1}^{\infty} \frac{Ga^{-1}(3/2)}{k+a} + \delta_{\{\frac{1}{k} - \frac{1}{k+a}\}} \tag{39}$$

where δ denotes a Dirac measure.

4 Simulations

4.1 Sampling P-IG distributions

There is a useful connection to the class of generalized gamma convolutions (GGC), which are infinitely divisible and hence straightforward to sample from. Bondesson (1992) defines the class of GGC distributions on $[0, \infty)$ as those with Laplace transform

$$E[e^{-sX}] = \exp\left[-as + \int_0^\infty \log\left(\frac{z}{z+s}\right) U(dz)\right] \tag{40}$$

with (left-extremely) $a \geq 0$ and $U(dz)$ a non-negative measure on $(0, \infty)$ (with finite mass on any compact set of $[0, \infty)$) such that $\int_0^1 |\log(z)| U(dz) < \infty$ and $\int_1^\infty z^{-1} U(dz) < \infty$. The $\sigma-$ finite measure U is chosen so that

$$\psi(s) = \int_0^\infty \log\left(1 + \frac{s}{z}\right) U(dz) = \int_0^\infty \int_0^\infty \left(1 - e^{-sz}\right) t^{-1} e^{-tz} U(dz) < \infty \tag{41}$$

U is often refered to as the Thorin measure (and can have infinite mass). The corresponding Lévy measure is $t^{-1} \int_0^\infty e^{tz} U(dz)$. This representation is basis of many approximate sampling techniques such as Bondesson (1992), Damien et al. (1995)

Generalized Gamma convolution (GGC) can be approximated by a compounded Poisson process where $N(dy)$ is completely monotone with $N(dy) = E\left[e^{-yz}\right] = \int e^{-yz} U(dz)$.

$$E[e^{-sX}] = \psi(s) = \exp\left\{-as + \int_0^\infty (e^{-sy} - 1) N(dy)\right\} \tag{42}$$

This defines a probability distribution, with density $p_{\text{E-RG}}(\omega \mid a)$ for all $a \geq 0$.

P-IG distribution is a special case of GGC. Suppose $U(dz) = \sum_{k=0}^\infty \delta_{1/d_k}(dz)$, when $\sum_{k=1}^\infty d_k^{-1} < \infty$, we have

$$\exp \int_0^\infty \log\left(\frac{z}{z+t}\right) U(dz) + \sum_{k=1}^\infty d_k^{-1} t = \prod_{k=1}^\infty \left(1 + \frac{t}{d_k}\right) e^{-t/d_k} \tag{43}$$

and the method of section 4.1 of Bondesson (1992) directly apply here. Bondesson (1992) also consider the approximation error when truncated to a finite sum.

5 Additional notes

$$(1+as)^{-\lambda} = \exp\left\{\lambda \int_0^\infty (e^{-sy}-1)\frac{1}{y}e^{-\frac{y}{a}}dy\right\} \tag{44}$$

Pólya-Gamma PG(b, c)

$$\begin{aligned}\psi(s) = E\left[e^{-sX}\right] &= \exp\left\{b\int_0^\infty (e^{-sy}-1)\frac{1}{y}\left[\sum_{k=1}^\infty e^{-d_k y}\right]dy\right\} \\ &= \prod_{k=1}^\infty \left(1+d_k^{-1}s\right)^{-b}, \quad \text{where } d_k = 2\left(k-\frac{1}{2}\right)^2\pi^2 + \frac{1}{2}c^2\end{aligned} \tag{45}$$

Hence

$$N(dy) = \frac{1}{y}\sum_{k=1}^\infty e^{-d_k y} \tag{46}$$

Section 4.3 of Bondesson,

$$g(u) = e^{-u}, \quad U(d\tau) = \sum_{k=1}^\infty \delta_{d_k} \tag{47}$$

Tail approximation by a normal, $\sum_{k=1}^\infty d_k^{-1} < \infty$, set $b = U[0,\infty)$ and $M(d\tau) = b^{-1}U(d\tau)$. Simulate points T_i from Póisson(b) process on $(0,\infty)$ and iid T_i from M, the random variable is

$$X = \sum_{i=1}^\infty e^{-T_i}V_i \tag{48}$$

where $V_i = \tau_i^{-1}W$ and $W \sim \Gamma(1,1)$, i.e. exponential. Then truncate the sum.

6 Applications

7 Discussion

Pólya-inverse Gamma distributions are a flexible class with a wide range of Bayesian applications. NP-Bayes, latent Dirichlet allocations, Topic models and Gamma scale inference all involve posterior distributions that depend on ratios of gamma functions. P-IG distributions provide a natural latent data augmentation strategy that expresses theire models as scale mixture of normals. This allows MCMC and fast scalable algorithms for implementation. They arise as latent variable conditional distributions for MCMC sampling. We show that P-IG distributions can be represented as scale mixture of normal, see Bhattacharya et al. (2016) for introduction.

Bayesian classification and allocation applications involve exponential reciprocal gamma distributions (E-RG) which provide a natural class of conjugate priors for gamma inference problems.

There are a number of avenues for future research. In particular, regularized scale allocation models can be implemented using P-IG and E-RG distributions using data augmentation methods of Polson and Scott (2013).

8 Notes on 09/25

$$\left(1 + a\sqrt{s}\right) e^{-\sqrt{x}} = \int_0^\infty e^{-tx} \frac{a + 2t(1-a)}{4\sqrt{\pi} t^{5/2}} e^{-\frac{1}{4}t} dt \tag{49}$$

If $a = 1$, PIG

$$(1 + \sqrt{x}) e^{-\sqrt{x}} = \int_0^\infty e^{-tx} \frac{1}{4\sqrt{\pi} t^{5/2}} e^{-\frac{1}{4}t} dt \tag{50}$$

If $0 < a < 1$, Posterior of PIG

$$= \int_0^\infty e^{-tx} \left\{ a \frac{e^{-\frac{1}{4}t}}{4\sqrt{\pi} t^{5/2}} + (1-a) \frac{e^{-\frac{1}{4}t}}{4\sqrt{\pi} t^{3/2}} \right\} \tag{51}$$

Inverse Gamma $3/2$ and $1/2$, therefore

$$\prod_{k=1}^{\infty} \left(1 + \frac{t}{d_k + c^2}\right) e^{-\frac{t}{d_k}} = \sum_{k=1}^{\infty} \frac{a/\gamma_{3/2} + (1-a)/\gamma_{1/2}}{d_k} \tag{52}$$

Remark 2: Bondersonn

$$E_{\omega|b} \left\{ e^{-t^2 \frac{\omega}{a}} \right\} = \left(1 + \frac{t}{a}\right) e^{-\frac{t}{b}} = \exp\left(\log\left(1 + \frac{t}{a}\right) - \frac{t}{b} \right) \tag{53}$$

Remark 3: P-IG = GGC, Thorin measure $U(\infty) = \infty$

References

Andrews, D. F. and C. L. Mallows (1974). Scale mixtures of Normal distributions. *Journal of the Royal Statistical Society, B*, 99–102.

Bhattacharya, A., A. Chakraborty, and B. K. Mallick (2016). Fast sampling with Gaussian scale mixture priors in high-dimensional regression. *Biometrika*.

Blei, D. M. and J. D. Lafferty (2006). Dynamic topic models. In *Proceedings of the 23rd international conference on Machine Learning*. ACM.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research 3*, 993–1022.

Bondesson, L. (1982). On simulation from infinitely divisible distributions. *Advances in Applied Probability 14*(4), 855–869.

Bondesson, L. (1992). Generalized Gamma Convolutions and related classes of distributions and densities. *Lecture Notes in Statistics 76*.

Carlin, B. P. and N. G. Polson (1991). Inference for Non-Conjugate Bayesian models using the Gibbs sampler. *Canadian Journal of Statistics 19*(4), 399–405.

Damien, P., P. W. Laud, and A. F. Smith (1995). Approximate random variate generation from infinitely divisible distributions with applications to Bayesian inference. *Journal of the Royal Statistical Society B*, 547–563.

Damsleth, E. (1975). Conjugate classes for Gamma distributions. *Scandinavian Journal of Statistics*, 80–84.

Escobar, M. D. and M. West (1995). Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association 90*(430), 577–588.

Ferguson, T. S. and M. J. Klass (1972). A representation of independent increment processes without Gaussian components. *Annals of Mathematical Statistics 43*(5), 1634–1643.

Glynn, C., S. T. Tokdar, B. Howard, and D. L. Banks (2019, 03). Bayesian analysis of dynamic linear topic models. *Bayesian Anal. 14*(1), 53–80.

Hartman, P. (1976). Completely monotone families of solutions of n-th order linear differential equations and infinitely divisible distributions. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze 3*(2), 267–287.

Miller, J. W. (2018). Fast and accurate approximation of the full conditional for Gamma shape parameters. *arXiv:1802.01610*.

Minka, T. (2000). Estimating a Dirichlet distribution. *Technical report, MIT*.

Polson, N. G. and J. G. Scott (2013). Data augmentation for Non-Gaussian regression models using variance-mean mixtures. *Biometrika 100*(2), 459–471.

Polson, N. G., J. G. Scott, and J. Windle (2013). Bayesian inference for logistic models using Pólya–Gamma latent variables. *Journal of the American Statistical Association 108*(504), 1339–1349.

Polson, N. G. and S. L. Scott (2011). Data augmentation for Support Vector Machines. *Bayesian Analysis 6*(1), 1–23.

Rossell, D. et al. (2009). GaGa: a parsimonious and flexible model for differential expression analysis. *The Annals of Applied Statistics 3*(3), 1035–1051.

Roynette, B. and M. Yor (2005). Couples de Wald indéfiniment divisibles. Exemples liés à la fonction gamma d'Euler et à la fonction zeta de Riemann. *Annales de l'institut Fourier 55*(4), 1219–1284.

Stefanski, L. A. (1991). A normal scale mixture representation of the logistic distribution. *Statistics & Probability Letters 11*(1), 69–70.

Walker, S. and P. Damien (2000). Representations of Lévy processes without Gaussian components. *Biometrika 87*(2), 477–483.

West, M. (1987). On scale mixtures of Normal distributions. *Biometrika 74*(3), 646–648.

West, M. (1992). Hyperparameter estimation in Dirichlet process mixture models. *Working paper, Duke University*.

Windle, J., N. G. Polson, and J. G. Scott (2014). Sampling Polya-Gamma random variates: alternate and approximate techniques. *arXiv:1405.0506*.

A Roynette and Yor (2005) Proof

This section we prove the Laplace transform of $Ga^{-1}\left(\frac{3}{2}, \frac{1}{2}\right)$ distribution. Note that if $x \sim Ga^{-1}\left(\frac{3}{2}, \frac{1}{2}\right)$, density $f(x) = \frac{1}{\sqrt{2\pi x^5}} e^{-\frac{1}{2x}}$, the Laplace transform is

$$E\left[\exp\left(-\frac{t^2}{2}x\right)\right] = (1+t)e^{-t} \tag{54}$$

See page 1252 of Roynette and Yor (2005)

$$E\left[\exp\left(-\frac{\lambda^2}{2}\frac{1}{2\gamma_{\frac{3}{2}}}\right)\right] = \frac{1}{\Gamma\left(\frac{3}{2}\right)}\int_0^\infty e^{-\lambda^2/4x} x^{\frac{1}{2}} e^{-x} dx \tag{55}$$

$$\frac{\partial}{\partial\lambda} E\left[e^{-\lambda^2/4\gamma_{\frac{3}{2}}}\right] = -\frac{\lambda}{2\Gamma\left(\frac{3}{2}\right)}\int_0^\infty e^{-\lambda^2/4x} x^{-\frac{1}{2}} e^{-x} dx = -\lambda e^{-\lambda} \tag{56}$$

$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}$,

$$\frac{\partial}{\partial\lambda}\left[E\left(e^{-\frac{1}{2}\lambda^2 H_1^{(2)}}\right)\right] = \frac{\partial}{\partial\lambda}\{(1+\lambda)e^{-\lambda}\} = -\lambda e^{-\lambda} \tag{57}$$

Therefore we can apply Bondesson (1982) to simulate by short-noise process.