

# Constraining new physics from Higgs measurements with Lilith: update to LHC Run 2 results

Thi Nhung Dao<sup>1</sup>, Sabine Kraml<sup>2\*</sup>, Duc Ninh Le<sup>1</sup>, Loc Tran Quang<sup>1</sup>

**1** Institute For Interdisciplinary Research in Science and Education, ICISE,  
590000, Quy Nhon, Vietnam

**2** Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes,  
CNRS/IN2P3, 53 Avenue des Martyrs, F-38026 Grenoble, France

\* sabine.kraml@lpsc.in2p3.fr

June 10, 2019

## Abstract

Lilith is public python library for constraining new physics from Higgs signal strength measurements. We here present version 2.0 of Lilith together with an updated database which includes the full set of ATLAS and CMS Run 2 Higgs results for  $36 \text{ fb}^{-1}$ . Both the code and the XML database where extended from the ordinary Gaussian approximation employed in Lilith-1.1 to using variable Gaussian and Poisson distributions. Moreover, Lilith can now make use of correlation matrices of arbitrary dimension. We provide detailed validations of the implemented experimental results as well as a status of global fits for *i)* reduced Higgs couplings and *ii)* Two-Higgs-doublet models of Type-I and Type-II. Lilith-2.0 is available on GitHub and ready to be used to constrain a wide class of new physics scenarios.

## 1 Introduction

Introduce Higgs couplings fits and Lilith [?] .....

.....

.....

.....

.....

.....

.....

## 2 Extended XML format for experimental input

In the *Lilith* database, every single experimental result is stored in a separate XML file. This allows to easily select the results to use in a fit, and it also makes maintaining and updating the database rather easy.

The root tag of each XML file is `<expmu>`, which has two mandatory attributes, `dim` and `type` to specify the type of signal strength result. Production and decay modes are specified via `prod` and `decay` attributes either directly in the `<expmu>` tag or as efficiencies in `<eff>` tags. Additional (optional) information can be provided in `<experiment>`, `<source>`, `<sqrts>`, `<CL>` and `<mass>` tags. Taking the  $H \rightarrow \gamma\gamma$  result from the combined ATLAS and CMS Run 1 analysis [?] as a concrete example, the structure of the XML file is

```
<expmu decay="gammagamma" dim="2" type="n">
  <experiment>ATLAS-CMS</experiment>
  <source type="publication">CMS-HIG-15-002; ATLAS-HIGG-2015-07</source>
  <sqrts>7+8</sqrts>
  <mass>125.09</mass>
  <CL>68%</CL>

  <eff axis="x" prod="ggH">1.</eff>
  <eff axis="y" prod="VVH">1.</eff>

  <!-- (...) -->
</expmu>
```

where `<!-- (...) -->` is a placeholder for the actual likelihood information. For a detailed description, we refer to the original *Lilith* manual [?]. In the following, we assume that the reader is familiar with the basic syntax.

So far, the likelihood information could be specified in one or two dimensions in the form of [?]: 1D intervals given as best fit with  $1\sigma$  error; 2D likelihood contours described as best fit point and parameters  $a, b, c$  which parametrize the inverse of the covariance matrix; or full likelihood information as 1D or 2D grids of  $-2\log L$ . The first two options, 1D intervals and 2D likelihood contours, declared as `type="n"` in the `<expmu>` tag, employ an ordinary Gaussian approximation; in the 1D case, asymmetric errors are accounted for by putting together two one-sided Gaussians with the same mean but different variances, while the 2D case assumes symmetric errors. This does not always allow to describe the experimental data (i.e. the true likelihood) very well. Full 2D likelihood grids would be much better but are rarely available.

In order to treat asymmetric uncertainties in a better way, we have extended the XML format and fitting procedure in *Lilith* to Gaussian distributions of variable width (“variable Gaussian”) as well as generalized Poisson distributions. The declaration is `type="vn"` for variable Gaussian or `type="p"` for Poisson distribution in the `<expmu>` tag. Both work for 1D and 2D data with the same syntax. Moreover, in order to make use of the  $N$ -dimensional ( $N > 2$ ) correlation matrices which both ATLAS and CMS have started to provide, we have added a new XML format for correlated signal strengths in more than two dimensions. This can be used with the ordinary or variable Gaussian approximation for the likelihood. In the following we give explicit examples for the different possibilities.

## 1D likelihood parameterization

For 1D data, the format remains the same as in [?]. For example, a signal strength  $\mu(ZH, b\bar{b}) = 1.12^{+0.50}_{-0.45}$  is implemented as

```
<bestfit>1.12</bestfit>
<param>
  <uncertainty side="left">-0.45</uncertainty>
  <uncertainty side="right">0.50</uncertainty>
</param>
```

The `<bestfit>` tag contains the best-fit value, while the `<uncertainty>` tag contains the left (negative) and right (positive)  $1\sigma$  errors.<sup>1</sup> The choice of likelihood function is done by setting `type="n"` for ordinary, 2-sided Gaussian (as in `Lilith-1.1`); `type="vn"` for a variable Gaussian; or `type="p"` for a Poisson distribution in the `<expmu>` tag.

## 2D likelihood parameterization

For `type="vn"` and `type="p"`, signal strengths in 2D with a correlation are now described in an analogous way as 1D data. For example,  $\mu(\text{ggH}, WW) = 1.10^{+0.21}_{-0.20}$  and  $\mu(\text{VBF}, WW) = 0.62^{+0.36}_{-0.35}$  with a correlation of  $\rho = -0.08$  can be implemented as

```
<expmu decay="WW" dim="2" type="vn">
  <eff axis="x" prod="ggH">1.0</eff>
  <eff axis="y" prod="VBF">1.0</eff>
  <bestfit>
    <x>1.10</x>
    <y>0.62</y>
  </bestfit>
  <param>
    <uncertainty axis="x" side="left">-0.20</uncertainty>
    <uncertainty axis="x" side="right">+0.21</uncertainty>
    <uncertainty axis="y" side="left">-0.35</uncertainty>
    <uncertainty axis="y" side="right">+0.36</uncertainty>
    <correlation>-0.08</correlation>
  </param>
</expmu>
```

Here, the `<eff>` tag is used to declare the `x` and `y` axes. The `<bestfit>` tag specifies the location of the best-fit point in the `(x,y)` plane. The `<uncertainty>` tags contain the left (negative) and right (positive)  $1\sigma$  errors for the `x` and `y` axes, and finally the `<correlation>` tag specifies the correlation between `x` and `y`. The choice of likelihood function is again done by setting `type="vn"` or `type="p"` in the `<expmu>` tag.

To ensure backwards compatibility, `type="n"` however still requires the tags `<a>`, `<b>`, `<c>` to give the inverse of the covariance matrix instead of `<uncertainty>` and `<correlation>`, see [?].

<sup>1</sup>The values in the `<uncertainty>` tag can be given with or without a sign.

## 117 Multi-dimensional data

118 For correlated signal strengths in more than 2 dimensions, a new format is introduced.  
 119 We here illustrate it by means of the CMS result [?], which has signal strengths for 24  
 120 production and decay mode combinations plus a  $24 \times 24$  correlation matrix. First, we set  
 121 `dim="24"` and label the various signal strengths as axes `d1`, `d2`, `d3`, ... `d24`:<sup>2</sup>

```
122 <expmu dim="24" type="vn">
123   <eff axis="d1" prod="ggH" decay="gammagamma">1.0</eff>
124   <eff axis="d2" prod="ggH" decay="ZZ">1.0</eff>
125   <eff axis="d3" prod="ggH" decay="WW">1.0</eff>
126   ...
127   <eff axis="d24" prod="ttH" decay="tautau">1.0</eff>
```

128 The best-fit values for each axis are specified as

```
129   <bestfit>
130     <d1>1.16</d1>
131     <d2>1.22</d2>
132     <d3>1.35</d3>
133     ...
134     <d24>0.23</d24>
135   </bestfit>
```

136 The `<param>` tag then contains the uncertainties and correlations in the form

```
137   <param>
138     <uncertainty axis="d1" side="left">-0.18</uncertainty>
139     <uncertainty axis="d1" side="right">+0.21</uncertainty>
140     <uncertainty axis="d2" side="left">-0.21</uncertainty>
141     <uncertainty axis="d2" side="right">+0.23</uncertainty>
142     ...
143     <uncertainty axis="d24" side="left">-0.88</uncertainty>
144     <uncertainty axis="d24" side="right">+1.03</uncertainty>
145
146     <correlation entry="d1d2">0.12</correlation>
147     <correlation entry="d1d3">0.16</correlation>
148     <correlation entry="d1d4">0.08</correlation>
149     ...
150     <correlation entry="d23d24">0</correlation>
151   </param>
152 </expmu>
```

153 This will also work for `type="n"`.

---

<sup>2</sup>The `<experiment>`, `<source>`, `<sqrts>`, etc. tags are omitted for brevity.

### 3 Likelihood calculation

For the variable Gaussian, the computation of the likelihood in `computelikelihood.py` follows Section 3.6 “Variable Gaussian (2)” of [?]. The Poisson distribution is implemented according to Section 3.4 “Generalised Poisson”, eq. (10a), of [?].

*..... give explicit equations and explain implementation i the code ....*

## 4 ATLAS and CMS results included in the database update

### 4.1 ATLAS Run 2 results for $36 \text{ fb}^{-1}$

The ATLAS Run 2 results included in this release are summarised in Table 1 and explained in more detail below.

mode	$\gamma\gamma$	$ZZ^*$	$WW^*$	$\tau\tau$	$b\bar{b}$	inv.
ggH	[?]	[?]	[?]	[?]	—	—
VBF	[?]	[?]	[?]	[?]	[?]	—
WH	[?]	[?]	—	—	[?]	—
ZH	[?]	[?]	—	—	[?]	[?]
ttH	[?, ?]	[?, ?]	[?]	[?]	[?, ?]	—

Table 1: Overview of ATLAS Run 2 results included in this release.

**$H \rightarrow \gamma\gamma$  (HIGG-2016-21):** The ATLAS analysis [?] provides in Fig. 12  $H \rightarrow \gamma\gamma$  signal strengths separated into ggH, VBF, VH and “top” (ttH+ $t\bar{t}H$ ) production modes. Since no correlations are given for the signal strengths, we use instead the correlations for the stage-0 simplified template cross sections (STXS) provided in Fig. 40a of the ATLAS paper, which should be a close enough match. It turns out that these data do not allow to reproduce very well the ATLAS coupling fits for  $(C_V, C_F)$  or  $(C_\gamma, C_g)$ . The reason seems to be that the  $\mu$  values rounded to one decimal are not precise enough. We have therefore extracted the best-fit points and uncertainties from fits to the 1D profile likelihoods, which are provided as Auxiliary Figures 23a–d on the analysis webpage. and used these together with the STXS correlations in the Lilith XML file. This gives a better coupling fit, as shown in the validation plots in Fig. 1.

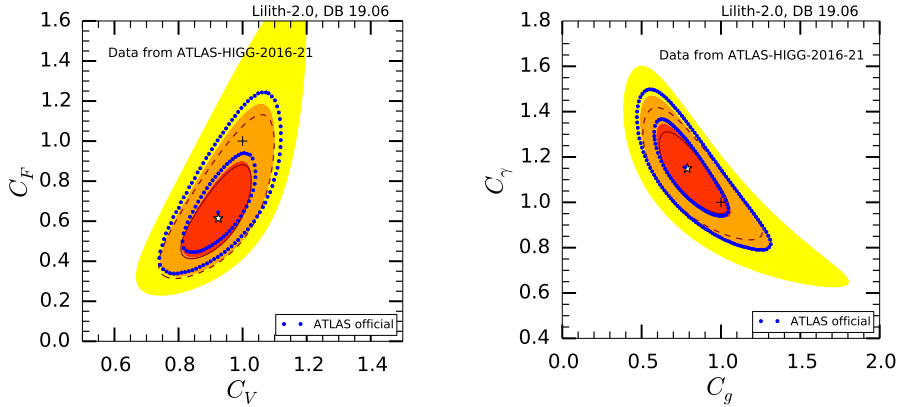


Figure 1: Fit of  $C_F$  vs.  $C_V$  (left) and  $C_\gamma$  vs.  $C_g$  (right) for data from the ATLAS  $H \rightarrow \gamma\gamma$  analysis [?]. The red, orange and yellow filled areas show the 68%, 95% and 99.7% CL regions obtained with Lilith using best-fit values and uncertainties for the signal strengths as extracted from Aux. Figs. 23a–d of the ATLAS analysis together with the correlation matrix for the stage-0 STXS. This can be compared to the 68%, 95% CL contours obtained using the rounded values from Fig. 12 of [?] (solid and dashed dark red lines) and to the official 68%, 95% CL contours from ATLAS (blue dots).

We note that the same fit quality is obtained when using only the correlation between ggH and VBF production modes (as 2D data) and treating VH and ttH as independent (as 1D data). The relevant XML files are all included in the database, so the user can choose

the preferred combination. This is relevant to avoid double counting when combining this with other measurements, concretely with the data from [?].

**$H \rightarrow ZZ^* \rightarrow 4l$  (HIGG-2016-22):** A similar issue as discussed for  $H \rightarrow \gamma\gamma$  above arises when using the signal strengths given in Table 9 of [?] together with the correlation matrix given in Aux. Fig. 4a. We therefore use  $\mu(\text{ggH}, ZZ^*)$  and  $\mu(\text{VBF}, ZZ^*)$  extracted from the 1D profile likelihoods Aux. Figs. 7a and 7b with a correlation  $\rho = -0.41$  according to Aux. Fig. 4c. For the VH and ttH production modes, we convert the given 95% CL limits into  $\mu(\text{VH}, ZZ^*) = 0_{-0.0}^{+1.85}$  and  $\mu(\text{ttH}, ZZ^*) = 0_{-0.0}^{+3.75}$ . For validation, we compare to the  $C_F$  vs.  $C_V$  fit from ATLAS, see Fig. 2.

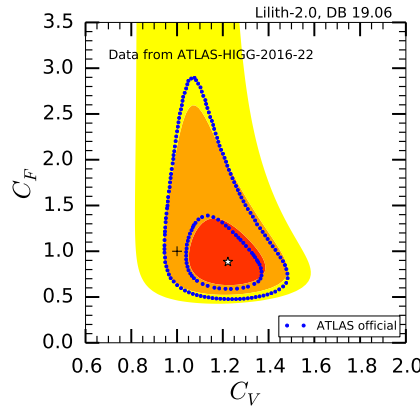


Figure 2: Fit of  $C_F$  vs.  $C_V$  for data from the ATLAS  $H \rightarrow ZZ^*$  analysis [?]. The 68%, 95% and 99.7% CL regions obtained with Lilith are shown as red, orange and yellow areas, and compared to the 68%, 95% CL contours from ATLAS (in blue).

**$H \rightarrow WW^* \rightarrow 2l2\nu$  (HIGG-2016-07):**

**$H \rightarrow \tau\tau$  (HIGG-2017-07):**

**$H \rightarrow b\bar{b}$  (HIGG-2016-29 and HIGG-2016-30):**

**$H \rightarrow \text{invisible}$  (HIGG-2016-28):** Results from the search for invisibly decaying Higgs bosons produced in association with a  $Z$  boson are presented in [?]. Assuming the Standard Model  $ZH$  production cross-section, an observed (expected) upper limit of 67% (39%) at the 95% confidence level is set on  $\text{BR}(H \rightarrow \text{inv})$  for  $m_H = 125$  GeV. We use  $1 - \text{CLs}$  as function  $\text{BR}(H \rightarrow \text{inv})$  extracted from auxiliary Figure 1c on the analysis' webpage.

## 4.2 CMS Run 2 results for $36 \text{ fb}^{-1}$

The CMS Run 2 results included in this release are summarised in Table 2 and explained in more detail below.

mode	$\gamma\gamma$	$ZZ^*$	$WW^*$	$\tau\tau$	$b\bar{b}$	$\mu\mu$	inv.
ggH	[?]	[?]	[?]	[?]	[?]	[?]	[?]
VBF	[?]	[?]	[?]	[?]	–	[?]	[?]
WH	[?]	[?]	[?]	[?]	[?]	–	[?]
ZH	[?]	[?]	[?]	[?]	[?]	–	[?]
ttH	[?]	[?]	[?]	[?]	[?]	–	–

Table 2: Overview of CMS Run 2 results included in this release. Note that we use the full  $24 \times 24$  correlation matrix for the signal strengths for each production and decay mode combination provided in [?].

**Combined measurements (HIG-17-031):** CMS presented in [?] a combination of the individual measurements for the  $H \rightarrow \gamma\gamma$  [?],  $ZZ$  [?],  $WW$  [?],  $\tau\tau$  [?],  $b\bar{b}$  [?, ?] and  $\mu\mu$  [?] decay modes as well as the  $t\bar{t}H$  analyses [?, ?, ?]. We use the best fit values and uncertainties for the signal strengths for each production and decay mode combination presented in Table 3 of [?] together with the  $24 \times 24$  correlation matrix provided as “Additional Figure 1” on the analysis webpage. As shown in Fig. 3, this allows to reproduce well the coupling fits of the CMS paper.

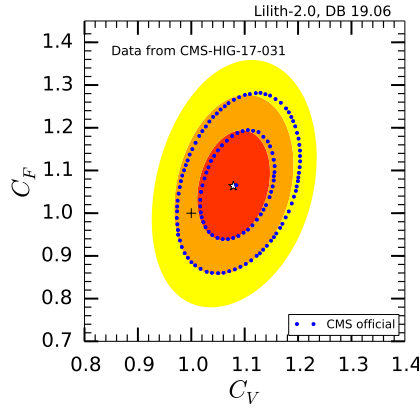


Figure 3: Fit of  $C_F$  vs.  $C_V$  using best-fit values and uncertainties for the signal strengths for each production (ggH, VBF, WH, ZH, ttH) and decay ( $\gamma\gamma$ ,  $ZZ$ ,  $WW$ ,  $\tau\tau$ ,  $b\bar{b}$ ,  $\mu\mu$ ) mode combination together with the  $24 \times 24$  correlation matrix from the CMS combination paper [?]. The  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  regions obtained with Lilith are shown as red, orange and yellow areas, and compared to the  $1\sigma$  and  $2\sigma$  contours from CMS (blue dots).

**$VH$ ,  $H \rightarrow \tau\tau$  (HIG-18-007):** The above data from [?] is supplemented by the results for the  $\tau\tau$  decay mode from the  $WH$  and  $ZH$  targeted analysis [?]. These are implemented in the form of 1D intervals for  $\mu(ZH, H \rightarrow \tau\tau)$  and  $\mu(WH, H \rightarrow \tau\tau)$  taken from Fig. 6 of [?].

**$H \rightarrow$  invisible (HIG-17-023):** In [?], CMS performed a search for invisible decays of a Higgs boson produced through vector boson fusion. We use the profile likelihood ratios for the qqH-tag, Z(ll)H-, V(qq')H- and ggH-tag categories extracted from their Fig. 8b



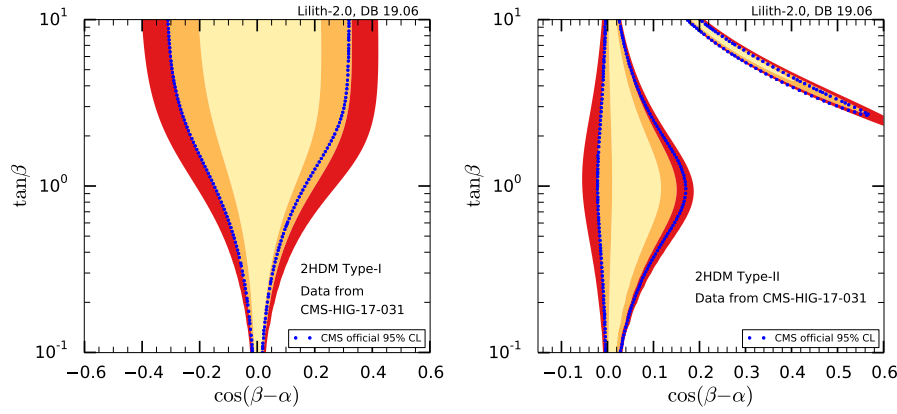


Figure 4: Fit of  $\tan \beta$  vs.  $\cos(\beta - \alpha)$  for the Two-Higgs-Doublet models of Type I (left) and Type II (right) using the data from the combined CMS measurement [?]. The beige, orange and red filled areas show the 68%, 95% and 99.7% CL regions obtained with *Lilith*, while the blue dots mark the 95% CL contours from CMS.

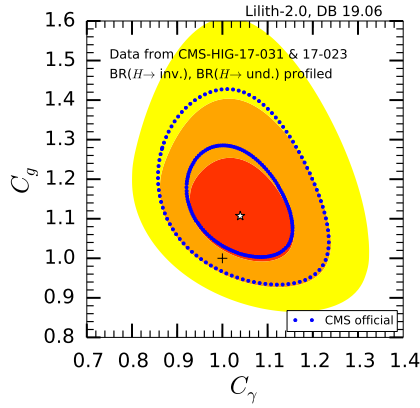


Figure 5: Fit of  $C_g$  vs.  $C_\gamma$  using the data from the combined CMS measurement [?] and the search for invisible decays of a Higgs boson [?]. The branching ratios of invisible and undetected decays are treated as free parameters in the fit. The  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  regions obtained with *Lilith* are shown as red, orange and yellow areas, and compared to the  $1\sigma$  and  $2\sigma$  contours from CMS (in blue).

220 together with the relative contributions from the different Higgs production mechanisms  
 221 given in Table 6 of that paper. This assumes that the relative signal contributions stay  
 222 roughly the same as for SM production cross sections. For validation, we reproduce in  
 223 Fig. 5 the  $C_g$  vs.  $C_\gamma$  fit of [?], where the branching ratios of invisible and undetected decays  
 224 are treated as free parameters.<sup>3</sup>

<sup>3</sup>The profiling in Fig. 5 was done with *Minuit*. Since *Minuit* does not allow conditional limits, in this case  $\text{BR}(H \rightarrow \text{inv.}) + \text{BR}(H \rightarrow \text{undetected}) < 1$ , we demanded that both  $\text{BR}(H \rightarrow \text{inv.})$  and  $\text{BR}(H \rightarrow \text{undetected})$  be less than 50%.

## 5 Status of Higgs coupling fits

## 6 Conclusion

must include a conclusion.

## Acknowledgements

S.K. thanks W. Adam, R. Schöfbeck, W. Waltenberger and N. Wardle for helpful discussions. This work was supported in part by the IN2P3 theory project “LHC-itools: methods and tools for the interpretation of the LHC Run 2 results for new physics”. D.T.N. thanks the LPSC Grenoble for hospitality and financial support for a research visit within the LHC-itools project. L.T.Q. thanks the ICISE ...

## A Overview of XML data files

## B Implementation of 2D Poisson likelihood with correlation

### B.1 Log-likelihood for Poisson distribution with continuous variable

The probability mass function of Poisson distribution, with parameter  $\lambda > 0$ , and variable  $k = 0, 1, 2, 3, \dots$ :

$$f(k; \lambda) = \frac{e^{-\lambda} \lambda^k}{k!}. \quad (1)$$

The log-likelihood function for Poisson distribution:

$$l(\lambda; k) = \log[f(k; \lambda)] = -\lambda + k \log \lambda - \log k!. \quad (2)$$

Here, since  $k$  is a discrete variable, we would redefine a log-likelihood function of parameter  $\lambda$  that fix continuous variable, denoted as “ $\nu$ ”. We re-define the parameter as  $\lambda \equiv \rho(\lambda - c) + \tau$ , with  $\eta = \tau/\rho$ , and  $c$  is the expected value at which the log-likelihood function reaches extrema. The new log-likelihood function reads:

$$l(\lambda; \nu) = -\rho(\lambda - c + \eta) + \nu \log \rho(\lambda - c + \eta) + \text{const.} \quad (3)$$

Here we want to set the extrema at  $\lambda = c$ . Since the Poisson distribution has expected value equal to parameter, we set  $\nu = \rho\eta$  so that the log-likelihood function reaches extrema at  $f(c)$ . For Poisson distribution, the extrema of log-likelihood function is not equal 0, we will consider its  $\Delta$  log-likelihood function:

$$\Delta l(\lambda; \nu) = l(\lambda; \nu) - l(c; \nu) = -\rho(\lambda - c) + \nu \ln \left[ 1 + \frac{\rho}{\nu}(\lambda - c) \right]. \quad (4)$$

$\Delta l(\sigma_-; \nu)$  and  $\Delta l(\sigma_+; \nu)$  yields  $-1/2$ , dividing their r.h.s terms of by  $\nu$  followed by taking the exponentials, we get the relation:

$$\frac{1 - \gamma\sigma_-}{1 - \gamma\sigma_+} = e^{-\gamma(\sigma_- + \sigma_+)}. \quad (5)$$

From  $\Delta l(\sigma_+; \nu) = -1/2$ , we could derive  $\nu$ :

$$\nu = \frac{1}{2(\gamma\sigma_+ - \ln(1 + \gamma\sigma_+))}. \quad (6)$$

## 251 B.2 A model for bivariate Poisson distribution with negative correlation

252 For references, see: Berkhout, Plug's.

253 The probability mass function (pmf) of Poisson distribution, with parameter  $\lambda > 0$ , and  
 254 variable  $k = 0, 1, 2, 3, \dots$ :

$$f(k; \lambda) = \frac{e^{-\lambda} \lambda^k}{k!}. \quad (7)$$

255 We would like to apply a model of bivariate Poisson distribution which allows negative  
 256 correlation. We define marginal and dependent pmf respectively as:

$$g_1(k_1; \lambda_1) = \frac{e^{-\lambda_1} \lambda_1^{k_1}}{k_1!}, \quad (8)$$

$$g_2(k_2|k_1; \lambda_2) = \frac{e^{-\lambda_2} \lambda_2^{k_2}}{k_2!}. \quad (9)$$

257 By convention, the parameters read:

$$\lambda_1 = e^{x' \beta_1}, \quad (10)$$

$$\lambda_2 = e^{x' \beta_2 + \alpha k_1}. \quad (11)$$

258 The parameter  $\alpha$  is added as a small correction that correspondent for the correlation of  
 259  $k_1, k_2$ . The joint pmf:

$$f(k_1, k_2; \lambda_1, \lambda_2) = g_2(k_2|k_1; \lambda_2) g_1(k_1; \lambda_1). \quad (12)$$

260 The marginal pmf remains the form of Poisson distribution, its expected value and variance  
 261 reads:

$$E(k_1) = \text{Var}(k_1) = \lambda_1. \quad (13)$$

262 The  $(r, s)$ th factorial moment of the joint distribution is derived as:

$$E[k_1(k_1 - 1) \dots (k_1 - r + 1) k_2(k_2 - 1) \dots (k_2 - s + 1)] \quad (14)$$

$$= \sum_{k_1=r}^{\infty} \sum_{k_2=s}^{\infty} \frac{e^{k_1(x' \beta_1) - \exp(x' \beta_1) + k_2(x' \beta_2 + \alpha k_1) - \exp(x' \beta_2 + \alpha k_1)}}{(k_1 - r)!(k_2 - s)!} \quad (15)$$

$$= \lambda_1^r e^{\lambda_1 [\exp(s\alpha) - 1] + s x' \beta_2 + r s \alpha}. \quad (16)$$

263 The expected value, variance of  $k_2$  and expected value of  $k_1 k_2$  is derived from the factorial  
 264 moment 16 as:

$$E(k_2) = e^{x' \beta_2 + (\exp(\alpha) - 1) \lambda_1}, \quad (17)$$

$$\text{Var}(k_2) = E(k_2) + [E(k_2)]^2 \{e^{\lambda_1 [\exp(\alpha) - 1]} - 1\}, \quad (18)$$

$$E(k_1 k_2) = \lambda_1 E(k_2) \exp(\alpha). \quad (19)$$

265 The covariance of  $k_1$  and  $k_2$  reads:

$$\text{Cov}(k_1, k_2) = E(k_1 k_2) - E(k_1) E(k_2) = \lambda_1 E(k_2) [\exp(\alpha) - 1], \quad (20)$$

266 The correlation reads:

$$\text{Corr}(k_1, k_2) = \frac{\lambda_1 E(k_2) [\exp(\alpha) - 1]}{\sqrt{\lambda_1 E(k_2) \{1 + E(k_2) \{e^{\lambda_1 [\exp(\alpha) - 1]} - 1\}}}}. \quad (21)$$

267 For later convenience, we derive the relation between  $\lambda_2$  and  $E(k_2)$ :

$$\lambda_2 = E(k_2) e^{\alpha k_1 - [\exp(\alpha) - 1] \lambda_1} = E(k_2) e^{\alpha k_1 - [\exp(\alpha) - 1] E(k_1)}. \quad (22)$$

### B.3 Log-likelihood for bivariate Poisson distribution with continuous variable and negative correlation

We derive a log-likelihood function that fixes continuous variables for the bivariate Poisson distribution which allows negative correlation. Now for the marginal pmf:

$$g_1(k_1; \lambda_1) = \frac{e^{-\lambda_1} \lambda_1^{k_1}}{k_1!}. \quad (23)$$

It remains the form of ordinary Poisson distribution. We apply 4 to derive the  $\Delta$  log-likelihood function:

$$\Delta l_1(\lambda_1; \nu_1) = -\rho_1(\lambda_1 - c_1) + \nu_1 \log \left[ 1 + \frac{\rho_1}{\nu_1}(\lambda_1 - c_1) \right]. \quad (24)$$

Considering the dependent pmf:

$$g_2(k_2|k_1; \lambda_2) = \frac{e^{-\lambda_2} \lambda_2^{k_2}}{k_2!}. \quad (25)$$

Its log-likelihood function fixing discrete variable  $k_2$ :

$$l_2(\lambda_2; k_2) = \log[g_2(k_2|k_1; \lambda_2)] = -\lambda_2 + k_2 \log \lambda_2 - \log k_2!, \quad (26)$$

with the parameter  $\lambda_2$  reads:

$$\lambda_2 = E(k_2) e^{\alpha k_1 - [\exp(\alpha) - 1] E(k_1)}. \quad (27)$$

For fixing continuous variable  $\nu_2$ , we redefine the parameter  $\lambda_2$  as:

$$\lambda_2 \equiv \rho_2(\lambda_2 - c_2 + \eta_2) e^{\alpha \nu_1 - [\exp(\alpha) - 1] \rho_1(\lambda_1 - c_1 + \eta_1)}. \quad (28)$$

The log-likelihood function  $l_2$  now become a function of two parameters  $\lambda_1, \lambda_2$  fixing two variables  $\nu_1, \nu_2$ :

$$l_2(\lambda_1, \lambda_2; \nu_1, \nu_2) = -\lambda_2 + \nu_2 \log \lambda_2 + \text{const.} \quad (29)$$

Its  $\Delta$  log-likelihood function reads:

$$\Delta l_2(\lambda_1, \lambda_2; \nu_1, \nu_2) = l_2(\lambda_1, \lambda_2; \nu_1, \nu_2) - l_2(c_1, c_2; \nu_1, \nu_2). \quad (30)$$

The correlation in 21 becomes:

$$\text{Corr}(\nu_1, \nu_2) = \frac{\nu_1 \nu_2 [\exp(\alpha) - 1]}{\sqrt{\nu_1 \nu_2 \{1 + \nu_2 \{e^{\nu_1 [\exp(\alpha) - 1]^2} - 1\}\}}}. \quad (31)$$

Finally, the  $\Delta$  log-likelihood for joint pmf:

$$\Delta l_{\text{join}}(\lambda_1, \lambda_2; \nu_1, \nu_2) = \Delta l_1(\lambda_1; \nu_1) + \Delta l_2(\lambda_1, \lambda_2; \nu_1, \nu_2). \quad (32)$$

## References