Wang, Huiqiao; Weiß, Christian H.; Zhang, Mingming

Goodness-of-fit testing in bivariate count time series based on a bivariate dispersion index

AStA Advances in Statistical Analysis

Provided in Cooperation with:
Springer Nature

Suggested Citation: Wang, Huiqiao; Weiß, Christian H.; Zhang, Mingming (2024) : Goodness-of-fit testing in bivariate count time series based on a bivariate dispersion index, AStA Advances in Statistical Analysis, ISSN 1863-818X, Springer, Berlin, Heidelberg, Vol. 109, Iss. 2, pp. 241-279, https://doi.org/10.1007/s10182-024-00512-3

AStA Advances in Statistical Analysis (2025) 109:241–279
https://doi.org/10.1007/s10182-024-00512-3

ORIGINAL PAPER

Goodness-of-fit testing in bivariate count time series based on a bivariate dispersion index

Huiqiao Wang[1] · **Christian H. Weiß**[2] · **Mingming Zhang**[1]

Received: 3 December 2023 / Accepted: 24 August 2024 / Published online: 17 September 2024

Abstract
A common choice for the marginal distribution of a bivariate count time series is the bivariate Poisson distribution. In practice, however, when the count data exhibit zero inflation, overdispersion or non-stationarity features, such that a marginal bivariate Poisson distribution is not suitable. To test the discrepancy between the actual count data and the bivariate Poisson distribution, we propose a new goodness-of-fit test based on a bivariate dispersion index. The asymptotic distribution of the test statistic under the null hypothesis of a first-order bivariate integer-valued autoregressive model with marginal bivariate Poisson distribution is derived, and the finite-sample performance of the goodness-of-fit test is analyzed by simulations. A real-data example illustrate the application and usefulness of the test in practice.

Keywords Asymptotic distribution · Bivariate dispersion index · Bivariate INAR(1) model · Bivariate Poisson distribution · Count time series

1 Introduction

During the last years, bivariate count processes $(X_{t,1}, X_{t,2})_{t\in\mathbb{Z}}$ with $\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$, i.e., where the quantitative random variables $X_{t,i}$ have a range contained in $\mathbb{N}_0 = \{0, 1, \dots\}$, received considerable interest in the research literature. In particular, a broad variety of time series models have been proposed for such data, see the surveys

✉ Huiqiao Wang
wanghqjo@163.com

Christian H. Weiß
weissc@hsu-hh.de

Mingming Zhang
zhangmingming@cass.cn

[1] Biogas Institute of Ministry of Agriculture and Rural Affairs, Renmin Road, Chengdu 610041, China

[2] Department of Mathematics and Statistics, Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany

by Karlis (2016); Weiß (2018, 2021); Fokianos (2021) for details and references. In this research, however, the focus is not on time series modelling, but on time series series diagnostics. Like the (multivariate) normal distribution for real-valued time series or the ordinary Poisson (Poi) distribution for univariate count time series, the bivariate Poisson (BPoi) distribution constitutes the default choice for modelling the marginal distribution in bivariate count time series. Here, $\{X_1, X_2\}$ is said to follow $BPoi(\lambda_1, \lambda_2, \lambda_0)$ if it can be generated from independent random variables Y_0, Y_1, Y_2 with $Y_i \sim Poi(\lambda_i)$ by $X_i = Y_i + Y_0$ for $i = 1, 2$; see Johnson et al. (1997); Kocherlakota and Kocherlakota (2014) for further information. For the practice of time series modelling, this implies the need to test if such a BPoi-model is appropriate for the given bivariate count time series. In the case of univariate count time series, one of the most popular goodness-of-fit (GoF) tests regarding a Poi-null hypothesis relies on Fisher's dispersion index. Such a simple moment-based GoF-test is not only easy to interpret, it is also possible to obtain closed-form asymptotics that allow for a computationally feasible implementation in practice, see Schweer and Weiß (2014); Weiß and Schweer (2016) for details. Here, however, we are concerned with bivariate counts such that GoF-tests for the BPoi-null are required. There are considerably less contributions to the GoF-test of a bivariate rather than a univariate Poi-distribution. Bivariate extensions to the univariate Fisher's dispersion index have been proposed by Crockett (1979); Loukas and Kemp (1986); Rayner and Best (1995); Best and Rayner (1997), but the asymptotics of these moment-based tests are only available for the case of independent and identically distributed (i. i. d.) bivariate counts. Recently, also more sophisticated GoF-tests relying on the probability generating function have been developed, see Novoa-Muñoz (2021) and the references therein, but again only for i. i. d. data. In particular, due to complicated asymptotics, these tests require for computationally demanding bootstrap implementations, where an extension to time-dependent data appears even more challenging. Thus, in the present article, we develop a GoF-test based on an appropriate type of bivariate dispersion index, and we derive the asymptotics for a relevant time series case. We also analyze if the proposed approach can be successfully used for testing the null hypothesis of a BPoi-marginal distribution, where we compare it to the joint application of the univariate dispersion test. Note that this research question differs from the work by Weiß and Aleksandrov (2019), where bivariate dispersion indexes were applied to time-lagged pairs of observations from a univariate count process, whereas we focus on the marginal distribution of a bivariate count process.

If testing the null of a BPoi-marginal distribution in a time series context, a time series model allowing for such a BPoi-marginal is required as the starting point. The probably most widely used model for bivariate count time series with BPoi-marginal is the first-order bivariate integer-valued autoregressive model(BINAR(1)) with diagonal thinning matrix as proposed by Pedeli and Karlis (2011). It is defined by the recursive scheme

$$\begin{bmatrix} X_{t,1} \\ X_{t,2} \end{bmatrix} = \begin{bmatrix} \alpha_1 & \\ & \alpha_2 \end{bmatrix} \circ \begin{bmatrix} X_{t-1,1} \\ X_{t-1,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t,1} \\ \varepsilon_{t,2} \end{bmatrix} = \begin{bmatrix} \alpha_1 \circ X_{t-1,1} + \varepsilon_{t,1} \\ \alpha_2 \circ X_{t-1,2} + \varepsilon_{t,2} \end{bmatrix}, \tag{1}$$

where the bivariate (and cross-correlated) count innovations $(\varepsilon_{t,1}, \varepsilon_{t,2})_{\mathbb{Z}}$ are i. i. d. Furthermore, "$\circ$" denotes the binomial thinning operator, which is defined by requiring for a conditional binomial distribution, namely $\alpha \circ X \mid X \sim Bin(X, \alpha)$ for some

$\alpha \in (0,1)$, and which serves as an integer-valued substitute of the ordinary multiplication $\alpha \cdot X$. In (1), the thinning operators are performed independently of each other and of the innovation sequence. Because of the diagonal thinning matrix, the component processes $(X_{t,i})_{\mathbb{Z}}$, $i = 1, 2$, are univariate INAR(1) processes, which also satisfy the mixing properties being required for applying a central limit theorem (CLT), see Ibragimov (1962); Schweer and Weiß (2014). In particular, their autocorrelation function (acf) is given by $\rho_i(h) = Corr[X_{t,i}, X_{t-h,i}] = \alpha_i^h$ for $h \in \mathbb{N} = \{1, 2, \ldots\}$. Note that $\alpha_1 = \alpha_2 = 0$ leads to the case of i. i. d. bivariate counts.

Due to its widespread use in practice and its appropriate stochastic properties, we shall focus our GoF-derivations to the case of an underlying BINAR(1) process, i.e., we test the BPoi-null hypothesis within the BINAR(1) framework (BPoiINAR(1) model). Note that the marginal distribution $\{X_{t,1}, X_{t,2}\} \sim BPoi(\lambda_1, \lambda_2, \lambda_0)$ implies that $X_{t,1} \sim Poi(\lambda_1 + \lambda_0)$ and $X_{t,2} \sim Poi(\lambda_2 + \lambda_0)$ hold as well. This BPoiINAR(1) process is achieved if the i. i. d. innovations satisfy

$$\{\varepsilon_{t,1}, \varepsilon_{t,2}\} \sim BPoi\Big(\lambda_1^*(1-\alpha_1) - \lambda_0(1-\alpha_1\alpha_2),\ \lambda_2^*(1-\alpha_2) - \lambda_0(1-\alpha_1\alpha_2),\ \lambda_0(1-\alpha_1\alpha_2)\Big), \tag{2}$$

where $\lambda_1^* = \lambda_1 + \lambda_0$, $\lambda_2^* = \lambda_2 + \lambda_0$, see Pedeli and Karlis (2011). The detailed derivation of (2) is also summarized in Appendix A.1. Model (2) shall be used for illustration throughout this article. In particular, for the considered bivariate dispersion index (see Sect. 2) and the corresponding dispersion test, its asymptotic distribution is derived under the null of the BPoiINAR(1) model. But it would also be possible to apply our test to other types of data-generating process (DGP) with BPoi-marginal by using a corresponding parametric-bootstrap implementation. In addition, we conjecture that asymptotic derivations should also be possible for some other types of count process with BPoi-marginal. For example, in view of the derivations by Aleksandrov and Weiß (2020) for univariate first-order integer-valued moving average (INMA(1)) models, we expect that an extension of our approach to an appropriate type of bivariate INMA(1) model could be possible, see Quoreshi et al. (2020) for a recent survey, which appears to be an interesting task for future research.

The outline of our article is as follows. In Sect. 2, we define the type of bivariate dispersion index to be considered in this article. The main part of this article is presented in Sect. 3, where we derive an explicit expression for the asymptotic distribution of the bivariate dispersion index under the null hypothesis of a BPoiINAR(1) DGP. This distribution can be used for an efficient implementation of the corresponding hypothesis test. We also discuss the special case of i. i. d. BPoi-counts and demonstrate that our novel asymptotic approximation outperforms an existing one. Finally, the joint asymptotics of the univariate dispersion indexes are derived, too. The finite-sample performance of our novel GoF-test is examined with simulation experiments in Sect. 4, while an illustrative data example is presented in Sect. 5. We conclude the article in Sect. 6.

2 A bivariate dispersion index

Let us first introduce some notations that are used throughout the manuscript:

$$\begin{aligned}
&\mu_1 := E[X_{t,1}]; \qquad \mu_2 := E[X_{t,2}];\\
&\mu_{1^2}(0) := E[X_{t,1}^2] = E^2[X_{t,1}] + Var[X_{t,1}] = \mu_1^2 + \sigma_1^2;\\
&\mu_{2^2}(0) := E[X_{t,2}^2] = E^2[X_{t,2}] + Var[X_{t,2}] = \mu_2^2 + \sigma_2^2;\\
&\mu_{1,2}(0) := E[X_{t,1}X_{t,2}] = E[X_{t,1}] \cdot E[X_{t,2}] + cov(X_{t,1}, X_{t,2}) = \mu_1 \cdot \mu_2 + \gamma_{12};\\
&\mu_{\varepsilon_1} := E[\varepsilon_{t,1}]; \qquad \mu_{\varepsilon_2} := E[\varepsilon_{t,2}].
\end{aligned} \tag{3}$$

Note that under the null hypothesis H_0: $\{X_{t,1}, X_{t,2}\} \sim BPoi(\lambda_1, \lambda_2, \lambda_0)$, the above moments become

$$\begin{aligned}
&\mu_1 = \lambda_1 + \lambda_0 = \lambda_1^*; \qquad \mu_2 = \lambda_2 + \lambda_0 = \lambda_2^*;\\
&\mu_{1^2}(0) = (\lambda_1 + \lambda_0)^2 + \lambda_1 + \lambda_0 = \lambda_1^{2*} + \lambda_1^*;\\
&\mu_{2^2}(0) = (\lambda_2 + \lambda_0)^2 + \lambda_2 + \lambda_0 = \lambda_2^{2*} + \lambda_2^*;\\
&\mu_{1,2}(0) = (\lambda_1 + \lambda_0)(\lambda_2 + \lambda_0) + \lambda_0 = \lambda_1^*\lambda_2^* + \lambda_0;\\
&\mu_{\varepsilon_i} = (\lambda_i + \lambda_0)(1 - \alpha_i) = \lambda_i^*(1 - \alpha_i) \quad \text{for } i = 1, 2.
\end{aligned} \tag{4}$$

Given an i. i. d. sample of n bivariate counts, $n \in \mathbb{N}$, a couple of bivariate dispersion indexes for uncovering deviations from the BPoi-null have been proposed in the literature, such as those by Crockett (1979); Loukas and Kemp (1986); Rayner and Best (1995); Best and Rayner (1997), to mention a few. In this article, we focus on the bivariate index proposed by Loukas and Kemp (1986) because of its simple and intuitive structure. Its sample version is defined by

$$\hat{I}_{LK} = \frac{\bar{X}_2 S_1^2 - 2S_{12}^2 + \bar{X}_1 S_2^2}{\bar{X}_1 \bar{X}_2 - S_{12}^2}, \tag{5}$$

where $\bar{X}_1, \bar{X}_2$ are the sample means, S_1^2, S_2^2 are the sample variances, and S_{12} is the sample cross-variance. Note that in the original proposal of Loukas and Kemp (1986), an additional re-scaling by the factor n was considered, i.e., $n\hat{I}_{LK}$ instead of $\hat{I}_{LK}$, but we omit the factor n for simplicity.

If we substitute the sample moments by the population moments (3), i.e.,

$$\begin{aligned}
&\bar{X}_2 \mapsto \mu_2, \qquad \bar{X}_1 \mapsto \mu_1, \qquad S_{12} \mapsto \gamma_{12} = \mu_{1,2}(0) - \mu_1\mu_2,\\
&S_1^2 \mapsto \sigma_1^2 = \mu_{1^2}(0) - \mu_1^2, \qquad S_2^2 \mapsto \sigma_2^2 = \mu_{2^2}(0) - \mu_2^2,
\end{aligned}$$

then (5) is re-expressed as follows:

$$I_{LK} := \frac{\mu_2 \cdot \sigma_1^2 - 2 \cdot \gamma_{12}^2 + \mu_1\sigma_2^2}{\mu_1\mu_2 - \gamma_{12}^2} \tag{6}$$

$$= \frac{\mu_2[\mu_{1^2}(0) - \mu_1^2] - 2 \cdot [\mu_{1,2}(0) - \mu_1\mu_2]^2 + \mu_1 \cdot [\mu_{2^2}(0) - \mu_2^2]}{\mu_1\mu_2 - [\mu_{1,2}(0) - \mu_1\mu_2]^2}. \tag{7}$$

Under H_0, if $\{X_{t,1}, X_{t,2}\}$ follows the BPoi-distribution, we get $I_{LK} = 2$.

3 The goodness-of-fit test based on the bivariate dispersion index

In what follows, we propose $\hat{I}_{LK}$ from (5) to be used as a test statistic for the null hypothesis that the bivariate count time series $\{X_{t,1}, X_{t,2}\}_{t=1,\ldots,n}$ is generated by a stationary BPoiINAR(1) process (2). Note that assuming the null hypothesis to hold, also the existence of moments up to any order (due to the BPoi-distribution) as well as relevant mixing properties (due to the INAR structure, see Schweer and Weiß (2014)) are implied, as they are required for applying a CLT like in Ibragimov (1962). To enable an asymptotic implementation of this GoF-test, we have to derive the asymptotic distribution of $\sqrt{n}\,(\hat{I}_{\mathrm{LK}} - 2)$ under the null, which is done in two steps. As I_{LK} equals $g_{LK}(\boldsymbol{\mu})$ with the vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_{1^2}(0), \mu_{2^2}(0), \mu_{1,2}(0))$ and the function

$$g_{LK}(\boldsymbol{y}) = \frac{y_2 \cdot (y_3 - y_1^2) - 2 \cdot (y_5 - y_1y_2)^2 + y_1(y_4 - y_2^2)}{y_1y_2 - (y_5 - y_1y_2)^2}, \tag{8}$$

for $\boldsymbol{y} = (y_1, y_2, y_3, y_4, y_5)$, we first derive a CLT for the vector of sample moments corresponding to $\boldsymbol{\mu}$, see Sect. 3.1. Afterwards in Sect. 3.2, we apply the Delta method Serfling (1980) and conclude the asymptotics of $\sqrt{n}\,\big(\hat{I}_{LK} - 2\big)$.

3.1 A CLT for some sample moments

Denote the vector $\boldsymbol{X_t} = (X_{t,1}, X_{t,2}, X_{t,1}^2, X_{t,2}^2, X_{t,1}X_{t,2})$, the sample mean of which is an estimator of $\boldsymbol{\mu}$. Furthermore, let $\boldsymbol{X_t}' = (X_{t,1} - \mu_1, X_{t,2} - \mu_2, X_{t,1}^2 - \mu_{1^2}(0), X_{t,2}^2 - \mu_{2^2}(0), X_{t,1}X_{t,2} - \mu_{1,2}(0))$ be its centered version. According to the CLT in Ibragimov (1962), we have

$$\sqrt{n} \cdot \frac{1}{n} \sum_{t=1}^{n} \boldsymbol{X_t}' = \sqrt{n}\bar{\boldsymbol{X}}_{\boldsymbol{t}}' \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{X_t'}}), \tag{9}$$

where the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{X_t'}} = n\,Var(\bar{\boldsymbol{X}}_{\boldsymbol{t}}') = n\,E[\bar{\boldsymbol{X}}_{\boldsymbol{t}}'\bar{\boldsymbol{X}}_{\boldsymbol{t}}'^{T}]$ equals

$$\boldsymbol{\Sigma}_{\boldsymbol{X_t'}} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} & \sigma_{35} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} & \sigma_{45} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_{55} \end{bmatrix}. \tag{10}$$

Since $\mathbf{\Sigma}_{\boldsymbol{X}_t'}$ is a symmetric matrix, so we only need to focus on the lower triangle of the matrix. For σ_{11}, σ_{22}, σ_{33}, σ_{44}, σ_{31}, and σ_{42}, explicit expression can directly be concluded from Weiß and Schweer (2016), namely

$$\begin{aligned}
\sigma_{11} &= \frac{1+\alpha_1}{1-\alpha_1}\mu_1 = \frac{1+\alpha_1}{1-\alpha_1}\lambda_1^*, \qquad \sigma_{22} = \frac{1+\alpha_2}{1-\alpha_2}\mu_2 = \frac{1+\alpha_2}{1-\alpha_2}\lambda_2^*,\\
\sigma_{33} &= \frac{1+\alpha_1^2}{1-\alpha_1^2}2\mu_1^2 + \frac{1+\alpha_1}{1-\alpha_1}\mu_1(1+2\mu_1)^2 = \frac{1+\alpha_1^2}{1-\alpha_1^2}2\lambda_1^{*^2} + \frac{1+\alpha_1}{1-\alpha_1}\lambda_1^*(1+2\lambda_1^*)^2,\\
\sigma_{44} &= \frac{1+\alpha_2^2}{1-\alpha_2^2}2\mu_2^2 + \frac{1+\alpha_2}{1-\alpha_2}\mu_2(1+2\mu_2)^2 = \frac{1+\alpha_2^2}{1-\alpha_2^2}2\lambda_2^{*^2} + \frac{1+\alpha_2}{1-\alpha_2}\lambda_2^*(1+2\lambda_2^*)^2,\\
\sigma_{31} &= \frac{1+\alpha_1}{1-\alpha_1}\mu_1(1+2\mu_1) = \frac{1+\alpha_1}{1-\alpha_1}\lambda_1^*(1+2\lambda_1^*),\\
\sigma_{42} &= \frac{1+\alpha_2}{1-\alpha_2}\mu_2(1+2\mu_2) = \frac{1+\alpha_2}{1-\alpha_2}\lambda_2^*(1+2\lambda_2^*).
\end{aligned} \tag{11}$$

Thus, in what follows, we only need to focus on the remaining covariances σ_{21}, σ_{32}, σ_{41}, σ_{51}, σ_{52}, σ_{43}, σ_{53}, σ_{54} and σ_{55}. To express these covariances, let us introduce the following moment notations (which are consistent with the ones used in (3)):

$$\mu_{1^r,2^s}(k) = E[X_{0,1}^r X_{k,2}^s], \qquad \mu_{i^r}(0) = E[X_{0,i}^r] \quad \text{for } i = 1,2. \tag{12}$$

Furthermore, we use

$$\mu^c_{1^r,2^s}(k) = E\big[\big(X_{0,1}^r - \mu_{1^r}(0)\big)\big(X_{k,2}^s - \mu_{2^s}(0)\big)\big] = \mu_{1^r,2^s}(k) - \mu_{1^r}(0)\mu_{2^s}(0), \tag{13}$$

to simplify the expressions for the cross central moments. The following closed-form expressions, which assume that the BPoiINAR(1)-null holds, shall be used:

$$\mu_{i^3}(0) = \mu_i + 3\mu_i^2 + \mu_i^3, \qquad \mu_{i^4}(0) = \mu_i + 7\mu_i^2 + 6\mu_i^3 + \mu_i^4 \quad \text{for } i = 1,2, \tag{14}$$

and

$$\begin{aligned}
\mu_{1^2,2}(0) =& \lambda_2\mu_{1^2}(0) + \frac{\lambda_0}{\lambda_1+\lambda_0}(\mu_1 + 3\mu_1^2 + \mu_1^3),\\
\mu_{1,2^2}(0) =& \lambda_1\mu_{2^2}(0) + \frac{\lambda_0}{\lambda_2+\lambda_0}(\mu_2 + 3\mu_2^2 + \mu_2^3),\\
\mu_{1^2,2^2}(0) =& \lambda_2\mu_{1^2}(0) + \frac{\lambda_1\lambda_0}{\mu_1^2}\mu_{1^3}(0) + \lambda_2^2\mu_{1^2}(0) + \frac{\lambda_0^2}{\mu_1^2}\mu_{1^4}(0) + \frac{2\lambda_2\lambda_0}{\mu_1}\mu_{1^3}(0),\\
\mu_{1^3,2}(0) =& \lambda_2\mu_{1^3}(0) + \frac{\lambda_0}{\mu_1}\mu_{1^4}(0), \qquad \mu_{1,2^3}(0) = \lambda_1\mu_{2^3}(0) + \frac{\lambda_0}{\mu_2}\mu_{2^4}(0).
\end{aligned} \tag{15}$$

The proof of (15) is given in Appendix A.2. The following Proposition 1 provides the missing covariances for the matrix $\mathbf{\Sigma}_{\boldsymbol{X}_t'}$ in (10).

Proposition 1 *Let* $\{X_{t,1}, X_{t,2}\}_{t=1,\ldots,n}$ *be generated by the BPoiINAR*(1) *process* (2). *Recall the moment formulas* (12)–(15). *Then, the entries of the covariance matrix* $\mathbf{\Sigma}_{X_t'}$ *in* (10), *except those already summarized in* (11), *are given by*:

$$
\begin{aligned}
\sigma_{21} &= \mu^c_{1,2}(0) + \mu^c_{1,2}(0)\big(\tfrac{\alpha_2}{1-\alpha_2} + \tfrac{\alpha_1}{1-\alpha_1}\big),\\
\sigma_{32} &= \mu^c_{1^2,2}(0) + \big(\tfrac{\alpha_1^2}{1-\alpha_1^2} + \tfrac{\alpha_2}{1-\alpha_2}\big)\mu^c_{1^2,2}(0) + \mu^c_{1,2}(0)\cdot(1+2\mu_1)\big(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1^2}{1-\alpha_1^2}\big),\\
\sigma_{41} &= \mu^c_{1,2^2}(0) + \big(\tfrac{\alpha_2^2}{1-\alpha_2^2} + \tfrac{\alpha_1}{1-\alpha_1}\big)\mu^c_{1,2^2}(0) + \mu^c_{1,2}(0)\cdot(1+2\mu_2)\big(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_2^2}{1-\alpha_2^2}\big),\\
\sigma_{51} &= \mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0) + \big(\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0)\big)\big(\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \tfrac{\alpha_1}{1-\alpha_1}\big)\\
&\quad + \mu_2\big(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big)(\mu_{1^2}(0) - \mu_1^2) + \mu_1\big(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big)\mu^c_{1,2}(0),\\
\sigma_{52} &= \mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0) + \big(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0)\big)\big(\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \tfrac{\alpha_1}{1-\alpha_1}\big)\\
&\quad + \mu_2\big(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big)\mu^c_{1,2}(0) + \mu_1\big(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big)(\mu_{2^2}(0) - \mu_2^2),\\
\sigma_{43} &= \mu^c_{1^2,2^2}(0) + \tfrac{\alpha_2^2}{1-\alpha_2^2}\mu^c_{1^2,2^2}(0) + \mu^c_{1^2,2}(0)(1+2\mu_2)(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_2^2}{1-\alpha_2^2})\\
&\quad + \tfrac{\alpha_1^2}{1-\alpha_1^2}\mu^c_{1^2,2^2}(0) + \mu^c_{1,2^2}(0)(1+2\mu_1)\big(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1^2}{1-\alpha_1^2}\big),\\
\sigma_{53} &= \mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0) + \big(\mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0)\big)\big(\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \tfrac{\alpha_1^2}{1-\alpha_1^2}\big)\\
&\quad + \big(\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0)\big)\big(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1^2}{1-\alpha_1^2}\big)(1+2\mu_1)\\
&\quad + \mu_2\big(\mu_{1^3}(0) - \mu_1\mu_{1^2}(0)\big)\big(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big) + \mu_1\mu^c_{1^2,2}(0)\big(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big),\\
\sigma_{54} &= \mu_{1,2^3}(0) - \mu_{2^2}(0)\mu_{1,2}(0) + \big(\mu_{2^3,1}(0) - \mu_{2^2}(0)\mu_{1,2}(0)\big)\big(\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \tfrac{\alpha_2^2}{1-\alpha_2^2}\big)\\
&\quad + \big(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0)\big)\big(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_2^2}{1-\alpha_2^2}\big)(1+2\mu_2)\\
&\quad + \mu_1\big(\mu_{2^3}(0) - \mu_2\mu_{2^2}(0)\big)\big(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big) + \mu_2\mu^c_{1,2^2}(0)\big(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big),\\
\sigma_{55} &= \mu_{1^2,2^2}(0) - \mu_{1,2}^2(0) + 2\Big\{\mu_2\big(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big)\big(\mu_{1^2,2}(0) - \mu_{1,2}(0)\mu_1\big)\\
&\quad + \big(\mu_{1^2,2^2}(0) - \mu_{1,2}^2(0)\big)\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \mu_1\big(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\big)\big(\mu_{1,2^2}(0) - \mu_{1,2}(0)\mu_2\big)\Big\}.
\end{aligned}
$$

The detailed proof is given in Appendices A.3–A.8.

In the special case of i. i. d. BPoi-counts, which means that $\alpha_1 = 0$ and $\alpha_2 = 0$, Proposition 1 further simplifies as in the following corollary.

Corollary 1 *If* $\{X_{t,1}, X_{t,2}\}_{t=1,\ldots,n}$ *are i. i. d. BPoi-counts, then the entries of the covariance matrix* $\mathbf{\Sigma}_{X_t'}$ *in* (10) *are*

$$\sigma'_{11} = \mu_1,\ \sigma'_{21} = \mu^c_{1,2}(0),\ \sigma'_{22} = \mu_2,$$
$$\sigma'_{31} = \mu_1(1+2\mu_1),\ \sigma'_{32} = \mu^c_{1^2,2}(0),\ \sigma'_{33} = 2\mu_1^2 + \mu_1(1+2\mu_1)^2,$$
$$\sigma'_{41} = \mu^c_{1,2^2}(0),\ \sigma'_{42} = \mu_2(1+2\mu_2),\ \sigma'_{43} = \mu^c_{1^2,2^2}(0),\ \sigma'_{44} = 2\mu_2^2 + \mu_2(1+2\mu_2)^2,$$
$$\sigma'_{51} = \mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0),\ \sigma'_{52} = \mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0),\ \sigma'_{55} = \mu_{1^2,2^2}(0) - \mu^2_{1,2}(0),$$
$$\sigma'_{53} = \mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0),\ \sigma'_{54} = \mu_{1,2^3}(0) - \mu_{2^2}(0)\mu_{1,2}(0).$$

3.2 The limiting distribution of the test statistic

After having determined the asymptotic distribution of $\sqrt{n}\,\bar{\boldsymbol{X}}'_t$, we next use the Delta method (first-order Taylor expansion) to derive the asymptotic distribution of $\sqrt{n}\left(\hat{I}_{LK} - 2\right)$. Here, $\hat{I}_{LK}$ can be expressed as $g_{LK}(\bar{\boldsymbol{X}}_t)$, where function $g_{LK}(\cdot)$ is defined in (8). We have

$$\sqrt{n}\big(g_{LK}(\bar{\boldsymbol{X}}_t) - g_{LK}(\boldsymbol{\mu})\big) \xrightarrow{d} N\big(\mathbf{0}, \sigma^2_{LK}\big) \quad \text{with } \sigma^2_{LK} = \boldsymbol{D}_{g_{LK}} \boldsymbol{\Sigma}_{\boldsymbol{X}'_t} \boldsymbol{D}^T_{g_{LK}}, \tag{16}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{X}'_t}$ was derived in Sect. 3.1. The $\boldsymbol{D}_{g_{LK}} = (g'_{LK,1}, g'_{LK,2}, g'_{LK,3}, g'_{LK,4}, g'_{LK,5})$ is the gradient vector of g_{LK} evaluated at $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_{1^2}(0), \mu_{2^2}(0), \mu_{1,2}(0))$, where the explicit expressions are as follows:

$$g'_{LK,1} = \frac{1}{(\mu_{1,2}(0)^2 - 2\mu_1\mu_{1,2}(0)\mu_2 + \mu_1\mu_2(-1+\mu_1\mu_2))^2}\Big\{ -2\big(\mu_1^2 + \mu_{1^2}(0)\big)\mu_{1,2}(0)\mu_2^2$$
$$- \mu_2^2\Big(\mu_{1^2}(0) - 2\mu_1\mu_{1^2}(0)\mu_2 + \mu_1^2\big((1+\mu_2)^2 - \mu_{2^2}(0)\big)\Big) + \mu_{1,2}(0)^2(\mu_2(2+2\mu_1+\mu_2) - \mu_{2^2}(0))\Big\}, \tag{17}$$

$$g'_{LK,2} = \frac{1}{\big(\mu_{1,2}(0)^2 - 2\mu_1\mu_{1,2}(0)\mu_2 + \mu_1\mu_2(-1+\mu_1\mu_2)\big)^2}\Big\{ -2\mu_1^2\mu_{12}(0)\big(\mu_2^2 + \mu_{2^2}(0)\big)$$
$$+ \mu_1^2\big(\big(-(1+\mu_1)^2 + \mu_{1^2}(0)\big)\mu_2^2 - \mu_{2^2}(0) + 2\mu_1\mu_2\mu_{2^2}(0)\big) + \mu_{1,2}(0)^2(-\mu_{1^2}(0) + \mu_1(2+\mu_1+2\mu_2))\Big\}, \tag{18}$$

$$g'_{LK,3} = \frac{\mu_2}{\mu_1\mu_2 - (\mu_{1,2}(0) - \mu_1\mu_2)^2}, \tag{19}$$

$$g'_{LK,4} = \frac{\mu_1}{\mu_1\mu_2 - (\mu_{1,2}(0) - \mu_1\mu_2)^2}, \tag{20}$$

$$g'_{LK,5} = \frac{2(-\mu_{1,2}(0) + \mu_1\mu_2)(-\mu_{1^2}(0)\mu_2 + \mu_1\mu_2(2+\mu_1+\mu_2) - \mu_1\mu_{2^2}(0))}{\big(\mu_{1,2}(0)^2 - 2\mu_1\mu_{1,2}(0)\mu_2 + \mu_1\mu_2(-1+\mu_1\mu_2)\big)^2}. \tag{21}$$

The derivation of $\boldsymbol{D}_{g_{LK}}$ is presented in Appendix A.9. Thus, combining (9) and (16), and after tedious algebraic computations, we finally get the following closed-form expression for σ^2_{LK}:

$$\sigma^2_{LK} = \frac{4\big(1-\alpha_1^2\alpha_2^2\big)\big(\lambda_0+\lambda_1\big)\big(\lambda_0+\lambda_2\big)\big[\big(\lambda_0+\lambda_1\big)\big(\lambda_0+\lambda_2\big)+\lambda_0^2\big]}{\big(1-\alpha_1^2\big)\big(1-\alpha_2^2\big)\big[\big(\lambda_0+\lambda_1\big)\big(\lambda_0+\lambda_2\big)-\lambda_0^2\big]^2}. \tag{22}$$

In the i. i. d. case, i.e., if $\alpha_1 = 0 = \alpha_2$, expression (22) further simplifies to

$$\sigma^2_{LK}\,|_{\alpha_1=0=\alpha_2} = \frac{4\big(\lambda_0+\lambda_1\big)\big(\lambda_0+\lambda_2\big)\big[\big(\lambda_0+\lambda_1\big)\big(\lambda_0+\lambda_2\big)+\lambda_0^2\big]}{\big[\big(\lambda_0+\lambda_1\big)\big(\lambda_0+\lambda_2\big)-\lambda_0^2\big]^2}. \tag{23}$$

Using (22) or (23), respectively, we can perform the $\hat{I}_{LK}$-test based on the critical value of the asymptotic normal distribution (16). Recall that $\lambda_0, \lambda_1, \lambda_2$ can be traced back to marginal moments of $\{X_{t,1}, X_{t,2}\}$, see (4). The true values thereof (as well as those of α_1, α_2), however, are not known in practice, so a plug-in approach has to be used: the sample moments of $\{X_{t,1}, X_{t,2}\}$ as well as the sample acf at lag 1 are used as consistent estimators.

Remark 1 At this point, a reference back to the original article by Loukas and Kemp (1986) is necessary. Loukas and Kemp (1986) solely considered the i. i. d.-case, and by a different reasoning, they concluded on an approximate χ^2-distribution for $n \cdot \hat{I}_{LK}$. According to their approximation, the variance $n \cdot Var[\hat{I}_{LK}]$ should always be ≈ 4, independently of $\lambda_0, \lambda_1, \lambda_2$. This does not agree with (23), which is > 4 as long as $\lambda_0 > 0$. Only if $\lambda_0 \approx 0$ or $\lambda_0 \ll \lambda_1, \lambda_2$, we get $\sigma^2_{LK} \approx 4$. To judge which approximation is more suitable in practice, we provide simulated values of the n-fold variance for various scenarios in Table 1 (using $R = 10^5$ replications). In the first block for $\lambda_0 = 0.5$, we see that these values are only slightly larger than 4, but nevertheless, our novel normal approximation (23) in row "∞" provides a much better agreement than just the value 4. In the second block for $\lambda_0 = 3$, however, the value 4 is considerably exceeded. But still, our novel normal approximation (23) is able to explain this value rather well (except for the case (2, 3) with sample size $n = 100$, where the sample distribution of $n \cdot \hat{I}_{LK}$ showed heavier tails than expected under normality). So we conclude that even in the i. i. d.-case, our novel asymptotics should be preferred upon the existing ones.

3.3 Joint distribution of the univariate indexes

As a competitor to our proposed $\hat{I}_{LK}$-test, we jointly apply the univariate dispersion indexes

$$I_1 = \frac{\sigma_1^2}{\mu_1} = \frac{\mu_{1^2}(0) - \mu_1^2}{\mu_1}, \quad I_2 = \frac{\sigma_2^2}{\mu_2} = \frac{\mu_{2^2}(0) - \mu_2^2}{\mu_2}.$$

to the components of the bivariate time series, i.e., we simultaneously test the null hypothesis that both components follow a univariate PoiINAR(1) process, see Schweer and Weiß (2014); Weiß and Schweer (2016). To get the correct distribution of this joint test, we again have to consider the cross-correlation between both components, i.e., the joint distribution of both statistics has to be considered. Since the univariate indexes only depend on the marginal first- and second-order moments of the series, analogously to Sect. 3.1, we define the vectors $\boldsymbol{Y}_t = (X_{t,1}, X_{t,2}, X_{t,1}^2, X_{t,2}^2)$ and $\boldsymbol{Y}_t' = (X_{t,1} - \mu_1, X_{t,2} - \mu_2, X_{t,1}^2 - \mu_{1^2}(0), X_{t,2}^2 - \mu_{2^2}(0))$. Note that these vectors are equal to the first four components of $\boldsymbol{X}_t$ and $\boldsymbol{X}_t'$ in Sect. 3.1, respectively. According to the CLT (9), we have

$$\sqrt{n}\,\bar{\boldsymbol{Y}}_t' \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{Y}_t'}),$$

The covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{Y}_t'} = n\,Var(\bar{\boldsymbol{Y}}_t')$, and its explicit expression is

$$\boldsymbol{\Sigma}_{\boldsymbol{Y}_t'} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{bmatrix},$$

which equals the upper 4×4-block of (10). Actually, all the expression in the matrix $\boldsymbol{\Sigma}_{\boldsymbol{Y}_t'}$ have already been derived in Sect. 3.1.

As we use the sample index vector $\left[\hat{I}_1 - 1, \hat{I}_2 - 1\right]^T$ as the test statistic, we focus on the limiting joint distribution of $\sqrt{n}\left[\hat{I}_1 - 1, \hat{I}_2 - 1\right]^T$. Note that the indexes I_i, $i = 1, 2$, can be written as $I_i = g_i(\boldsymbol{\mu})$ with $\boldsymbol{\mu} = (\mu_1, \mu_{1^2}(0), \mu_2, \mu_{2^2}(0))$, where

$$g_1(y_1, y_2, y_3, y_4) = \frac{y_3 - y_1^2}{y_1}, \quad g_2(y_1, y_2, y_3, y_4) = \frac{y_4 - y_2^2}{y_2}.$$

We compute the partial derivative of g_1 and g_2, evaluated at the vector $\boldsymbol{\mu}$, as

Table 1 Simulated values of $n \cdot Var[\hat{I}_{LK}]$, asymptotic value according to (23) in row "∞". Italic values used trimmed variance (omitting the ten lowest and largest simulated $\hat{I}_{LK}$ values) due to outlier problems

n	$\lambda_0 = 0.5$, $(\lambda_1, \lambda_2) =$				$\lambda_0 = 3$, $(\lambda_1, \lambda_2) =$			
	(2, 3)	(2, 8)	(6, 3)	(6, 8)	(2, 3)	(2, 8)	(6, 3)	(6, 8)
100	4.477	4.199	4.216	4.127	*39.254*	*7.967*	8.938	5.678
250	4.390	4.156	4.141	4.086	12.788	7.119	7.126	5.375
500	4.444	4.177	4.134	4.065	11.519	6.860	6.910	5.364
1000	4.393	4.169	4.174	4.071	10.970	6.767	6.843	5.336
1500	4.364	4.113	4.141	4.046	10.881	6.771	6.757	5.285
∞	4.360	4.144	4.134	4.055	10.612	6.654	6.720	5.280

$$d_{11} = \frac{\partial g_1}{\partial y_1} = -\frac{\mu_{1^2}(0)}{\mu_1^2} - 1,\ d_{12} = \frac{\partial g_1}{\partial y_2} = 0,\ d_{13} = \frac{\partial g_1}{\partial y_3} = \frac{1}{\mu_1},\ d_{14} = \frac{\partial g_1}{\partial y_4} = 0,$$
$$d_{21} = \frac{\partial g_2}{\partial y_1} = 0,\ d_{22} = \frac{\partial g_2}{\partial y_2} = -\frac{\mu_{2^2}(0)}{\mu_2^2} - 1,\ d_{23} = \frac{\partial g_2}{\partial y_3} = 0,\ d_{24} = \frac{\partial g_2}{\partial y_4} = \frac{1}{\mu_2}.$$

Thus, the Jacobian matrix of $\left[I_1 - 1, I_2 - 1\right]^T$ has the following form

$$\boldsymbol{D} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} \\ d_{21} & d_{22} & d_{23} & d_{24} \end{bmatrix}.$$

According to the Delta method, the joint distribution of $\sqrt{n}\left[\hat{I}_1 - 1, \hat{I}_2 - 1\right]^T$ is

$$\sqrt{n} \begin{bmatrix} \hat{I}_1 - 1 \\ \hat{I}_2 - 1 \end{bmatrix} \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_{uni}) \quad \text{with } \boldsymbol{\Sigma}_{uni} = \boldsymbol{D}\boldsymbol{\Sigma}_{\boldsymbol{Y}_t'}\boldsymbol{D}^T. \tag{24}$$

Then, according to the definition of the Chi-square distribution χ_2^2, we have

$$n \cdot \begin{bmatrix} \hat{I}_1 - 1 \\ \hat{I}_2 - 1 \end{bmatrix}^T \cdot \boldsymbol{\Sigma}_{uni}^{-1} \cdot \begin{bmatrix} \hat{I}_1 - 1 \\ \hat{I}_2 - 1 \end{bmatrix} \sim \chi_2^2. \tag{25}$$

So the test implementation is based on the χ_2^2-distribution (25). To simplify the required computations, it is useful to know again a closed-form expression of $\boldsymbol{\Sigma}_{uni}$ can be derived, namely

$$\boldsymbol{\Sigma}_{uni} = \begin{bmatrix} 2\,\frac{1+\alpha_1^2}{1-\alpha_1^2} & \frac{2(1-\alpha_1^2\alpha_2^2)\lambda_0^2}{(1-\alpha_1^2)(1-\alpha_2^2)(\lambda_0+\lambda_1)(\lambda_0+\lambda_2)} \\ \frac{2(1-\alpha_1^2\alpha_2^2)\lambda_0^2}{(1-\alpha_1^2)(1-\alpha_2^2)(\lambda_0+\lambda_1)(\lambda_0+\lambda_2)} & 2\,\frac{1+\alpha_2^2}{1-\alpha_2^2} \end{bmatrix}. \tag{26}$$

In the i. i. d.-case, the matrix $\boldsymbol{\Sigma}_{uni}$ further simplifies to the following form:

$$\boldsymbol{\Sigma}_{uni}\,|_{\alpha_1=0=\alpha_2} = \begin{bmatrix} 2 & \frac{2\lambda_0^2}{(\lambda_0+\lambda_1)(\lambda_0+\lambda_2)} \\ \frac{2\lambda_0^2}{(\lambda_0+\lambda_1)(\lambda_0+\lambda_2)} & 2 \end{bmatrix}. \tag{27}$$

Note that (26) and (27) constitute a novel extension of the univariate asymptotics derived by Schweer and Weiß (2014).

4 Simulation experiments

In this section, we present some results from a simulation study, which was done to evaluate the performance of the $\hat{I}_{LK}$-test as well as that of its asymptotic approximation. As the competitor, we consider the joint univariate indexes as discussed in Sect. 3.3.

4.1 Peformance under null hypothesis

To evaluate the approximation quality of the asymptotics derived in Sect. 3.2, we first simulate time series from the BPoiINAR(1) model, i.e., if H_0 is true. We consider the following parameter groups:

(A): $(\lambda_{1,\varepsilon}, \lambda_{2,\varepsilon}, \lambda_{0,\varepsilon}, \alpha_1, \alpha_2) = (3, 5, 0.5, 0.3, 0.5)$,
(B): $(\lambda_{1,\varepsilon}, \lambda_{2,\varepsilon}, \lambda_{0,\varepsilon}, \alpha_1, \alpha_2) = (1, 3, 0.5, 0.2, 0.4)$,
(C): $(\lambda_{1,\varepsilon}, \lambda_{2,\varepsilon}, \lambda_{0,\varepsilon}, \alpha_1, \alpha_2) = (2, 6, 0.5, 0.2, 0.4)$,
(D): $(\lambda_{1,\varepsilon}, \lambda_{2,\varepsilon}, \lambda_{0,\varepsilon}, \alpha_1, \alpha_2) = (3, 6, 1, 0.3, 0.5)$,
(E): $(\lambda_{1,\varepsilon}, \lambda_{2,\varepsilon}, \lambda_{0,\varepsilon}, \alpha_1, \alpha_2) = (5, 4, 0.5, 0.2, 0.4)$.

The most important property for practice is the achieved size. For convenience, we execute the tests on the 5%-level, where the test decisions are done with respect to two-sided critical values: The test leads to a rejection of the null if either the lower critical value $a_{0.025}$ or the upper critical value $a_{0.975}$ is violated. The proportion of rejections among R replications (we use $R = 10^5$) expresses the empirical size of the test (columns "biv" in Table 2), which has to be compared to the nominal 5%-level. As the competitor, we jointly apply the univariate-indexes test. It is implemented using the limiting χ^2_2-distribution, i.e., the upper critical value is $\chi^2_{2,1-\alpha}$, and the achieved sizes are tabulated in the columns "uni" of Table 2. For both types of GoF-test, the sizes are quite close to each other and, in particular, close to the given nominal level of 5%. So we conclude that the asymptotic implementation of both GoF-tests works rather well in practice.

In addition, we computed the standard error (SE) of the $\hat{I}_{LK}$-statistic in three ways. In Table 3, the column SE_{emp} represents the empirical standard deviation among the R replicates of $\hat{I}_{LK}$, $SE_{a;t}$ is the asymptotic SE implied by (22) with true parameters, whereas $SE_{a;e}$ expresses the mean asymptotic SE with plugged-in parameter estimates, i.e., the mean value obtained after inserting the R replicates of the parameter estimators into the asymptotics (22). The results are summarized in Table 3, where we recognize a good agreement between the three types of SE.

4.2 Power analyses

To evaluate the $\hat{I}_{LK}$-test's performance in uncovering violations of the BPoi-null hypothesis, we calculate the empirical power from R replications of various alternative models (and using a pre-run of length 250 for burn-in). We also compare the power of the $\hat{I}_{LK}$-test ("biv") to that of the joint application of the univariate test statistics ("uni") according to Sect. 3.3. As alternatives, we consider the following three types of DGP:

BINARCH(1) model: The bivariate INARCH(1) model (CH=conditional heteroscedasticity), see Lee et al. (2018), is defined by the conditional distribution $\boldsymbol{X}_t \mid \boldsymbol{X}_{t-1} \ldots \sim BPoi(\lambda_{t,1}, \lambda_{t,2}, \lambda_0)$, where the conditional mean $\boldsymbol{M}_t := E\left[\boldsymbol{X}_t \mid \boldsymbol{X}_{t-1}, \ldots\right]$ with $M_{t,i} = \lambda_0 + \lambda_{i,t}$ has the following recursive expression:

$$\boldsymbol{M}_t = \boldsymbol{B}_0 + \mathbf{A}\boldsymbol{X}_{t-1} = \begin{pmatrix} b_1 & \\ & b_2 \end{pmatrix} + \begin{pmatrix} \alpha_1 & \\ & \alpha_2 \end{pmatrix} \cdot \begin{pmatrix} X_{t,1} \\ X_{t,2} \end{pmatrix}.$$

We set $\lambda_0 = 0.5$, $\alpha_1 = 0.3$, $\alpha_2 = 0.5$, $b_1 = 2.45$, and $b_2 = 2.75$.

BNBINAR(1) model: The bivariate INAR(1) model (1), but with the innovation term $\{\varepsilon_{t,1}, \varepsilon_{t,2}\}$ following the bivariate negative binomial (BNB) distribution with parameters $(r, p_1, p_2, p_0) = (2, 0.3, 0.5, 0.2)$, see Pedeli and Karlis (2011). So $\{X_{t,1}, X_{t,2}\}$ has the same marginal mean as model (A).

CuBINAR(1) model: The circumstance-driven bivariate (CuB)INAR(1) model of Wang et al. (2024) defined by

$$\begin{bmatrix} X_{1,t}(s_t) \\ X_{2,t}(s_t) \end{bmatrix} = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \circ \begin{bmatrix} X_{1,t-1}(s_{t-1}) \\ X_{2,t-1}(s_{t-1}) \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t}(s_t, s_{t-1}) \\ \varepsilon_{2,t}(s_t, s_{t-1}) \end{bmatrix},$$

where the states s_t, $t \in \{1, \ldots, n\}$, are generated from a stationary 2-state Markov chain with initial probability vector (0.5, 0.5) and transition probability matrix $\begin{pmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{pmatrix}$. Under the generated state sequence, we consider the following two different parameter groups:

(F): $\left(\lambda_1(s_1), \lambda_1(s_2), \lambda_2(s_1), \lambda_2(s_2), \lambda_0, \alpha_1, \alpha_2\right) = (1, 3, 3, 5, 0.5, 0.15, 0.5),$

(G): $\left(\lambda_1(s_1), \lambda_1(s_2), \lambda_2(s_1), \lambda_2(s_2), \lambda_0, \alpha_1, \alpha_2\right) = (1, 2, 1.5, 2.5, 0.5, 0.15, 0.3).$

The obtained power results are summarized in Table 4. Both GoF-tests show a similar performance. For the BNBINAR(1) process, the joint univariate indexes usually lead to higher power values, but also the $\hat{I}_{LK}$-test is very powerful throughout (between 89 % and 100 % power). By contrast, the $\hat{I}_{LK}$-test is more powerful in case of the BINARCH(1) process and the CuBINAR(1) process with parameter group (G). Altogether, the $\hat{I}_{LK}$-test, which also utilizes the information about the cross-correlation of the bivariate counts, has very good power values for various alternatives, and often performs better than if jointly applying the univariate indexes.

Table 2 Simulated sizes of univariate and bivariate dispersion index test

n	(A)		(B)		(C)		(D)		(E)	
	uni	biv	uni	biv	uni	biv	uni	biv	uni	biv
100	0.0462	0.0505	0.0469	0.0461	0.0449	0.0471	0.0461	0.0495	0.0450	0.0457
250	0.0493	0.0492	0.0485	0.0486	0.0483	0.0480	0.0479	0.0487	0.0486	0.0495
500	0.0494	0.0502	0.0494	0.0499	0.0489	0.0498	0.0493	0.0501	0.0491	0.0493
1000	0.0510	0.0499	0.0493	0.0489	0.0498	0.0504	0.0492	0.0498	0.0495	0.0502
1500	0.0510	0.0495	0.0489	0.0504	0.0490	0.0488	0.0501	0.0500	0.0502	0.0498

Table 3 Empirical SE vs. asymptotic SE with either true parameters or plugged-in estimates, under different parameter groups

n	(A)			(B)			(C)		
	SE_{emp}	$SE_{a;t}$	$SE_{a;e}$	SE_{emp}	$SE_{a;t}$	$SE_{a;e}$	SE_{emp}	$SE_{a;t}$	$SE_{a;e}$
100	0.2411	0.2416	0.2366	0.2342	0.2312	0.2284	0.2243	0.2249	0.2220
250	0.1526	0.1528	0.1515	0.1468	0.1462	0.1455	0.1417	0.1423	0.1415
500	0.1078	0.1081	0.1076	0.1038	0.1034	0.1031	0.1009	0.1006	0.1003
1000	0.0764	0.0764	0.0763	0.0729	0.0731	0.0730	0.0714	0.0711	0.0710
1500	0.0623	0.0624	0.0623	0.0597	0.0597	0.0597	0.0579	0.0581	0.0580

n	(D)			(E)		
	SE_{emp}	$SE_{a;t}$	$SE_{a;e}$	SE_{emp}	$SE_{a;t}$	$SE_{a;e}$
100	0.2445	0.2457	0.2405	0.2220	0.2239	0.2211
250	0.1543	0.1554	0.1541	0.1416	0.1416	0.1409
500	0.1098	0.1099	0.1094	0.1000	0.1001	0.0999
1000	0.0777	0.0775	0.0775	0.0709	0.0708	0.0707
1500	0.0634	0.0633	0.0633	0.0578	0.0578	0.0578

5 Real-data analysis

In this section, we analyse a bivariate data set of accident counts, which was initially presented by Pedeli and Karlis (2011) and further discussed by Weiß (2018). The accident data $\boldsymbol{x}_t$, collected for the days $t = 1, \ldots, 365$ during the year 2001, consist of two sub-series, where $x_{t,1}$ expresses the number of daytime accidents and $x_{t,2}$ the number of nighttime accidents in the Schiphol area (Netherlands). The distinction is necessary as these types of accidents are known to behave differently, see Pedeli and Karlis (2011): while people more often travel for entertainment during nighttimes, we have rush-hour traffic, delivery traffic, etc. at daytimes. Nevertheless, (mild) cross-dependence between the daytime and nighttime traffic is commonly observed as both types of traffic have common environmental conditions (weather conditions, road characteristics, etc.). The data were read from Fig. 1 in Pedeli and Karlis (2011), a corresponding plot of the data is presented in Fig. 1. We observe about 7.277 daytime accidents in the mean, and 1.504 nighttime accidents, where the cross-correlation equals about 0.142. According to Pedeli and Karlis (2011); Weiß (2018), the autocorrelation structure of the data is sufficiently well explained by a bivariate INAR(1) model. So it remains to discuss an appropriate choice of the marginal distribution.

In Weiß (2018), the univariate dispersion indexes were applied separately to the $x_{t,1}$ and $x_{t,2}$, respectively. Their values $\hat{I}_1 = 2.8773$ and $\hat{I}_2 = 1.2479$ are both significant at the 5%-level, which indicates that the BPoiINAR(1) model is probably not sufficient for the data. To test this conjecture more thoroughly, also taking account of the apparent (though mild) cross-correlation, we apply the novel $\hat{I}_{LK}$-test. The maximum likelihood estimates of the BPoiINAR(1) model are

$$\hat{\alpha}_1 = 0.0769,\ \hat{\alpha}_2 = 0.0867,\ \hat{\lambda}_0 = 0.2700,\ \hat{\lambda}_1 = 7.0187,\ \hat{\lambda}_2 = 1.2265.$$

which are all significantly different from zero, see Pedeli and Karlis (2011); Weiß (2018). Recall that α_1, α_2 are the autocorrelation parameters, whereas the BPoi-parameters $\lambda_0, \lambda_1, \lambda_2$ satisfy $cov(X_{t,1}, X_{t,2}) = \lambda_0$ and $E[X_{t,i}] = \lambda_0 + \lambda_i$ for $i = 1, 2$. We use the fitted BPoiINAR(1) model to calculate the lower and upper critical value for the $\hat{I}_{\mathrm{LK}}$-test at the 5% level under the null hypothesis of the BPoiINAR(1) model, leading to ± 0.2086, while the test statistic $\hat{I}_{\mathrm{LK}} - 2$ itself is computed as 2.2790. Thus, we clearly reject the BPoi-null, which confirms the analysis of Pedeli and Karlis (2011); Weiß (2018), who finally recommended to use the BNBINAR(1) model for the accidents data. While the fitted BNBINAR(1) model does similarly well as the BPoiINAR(1) model in explaining the accidents' means (7.289 and 1.496), it better captures the apparent overdispersion (2.504 and 1.267), see Weiß (2018). Incorporating the overdispersion into the fitted model is important in view of predicting or monitoring future accident counts, because we thus account for more frequent extreme counts than under BPoi-assumptions.

6 Conclusions

In this paper, we proposed a new goodness-of-fit test for bivariate count time series under a bivariate-Poisson null hypothesis, more precisely, the bivariate-Poisson INAR(1) model. The construction of test statistics is based on the bivariate dispersion index by Loukas and Kemp (1986), which was originally proposed for testing i. i. d. bivariate counts. The asymptotic distribution of the test statistic was derived under the bivariate-Poisson INAR(1) null hypothesis. As a special case, we considered i. i. d. bivariate Poisson counts and improved an existing asymptotic approximation. Furthermore, we also adopted the joint univariate index as competitor and derived its asymptotics. The simulation results showed that under the null hypothesis, the sizes computed for the bivariate index and the joint univariate indexes are all close to the given nominal level. We also showed that the bivariate index leads to an appealing power performance, which is often better than that of the joint univariate indexes. Finally, we applied our novel GoF-test to real data, a bivariate time series of accident counts. We confirmed that a marginal Poisson distribution is not sufficient for explaining these data.

Appendix Proofs

A.1 Proof of (2)

If $\{X_1, X_2\} \sim BPoi(\lambda_1, \lambda_2, \lambda_0)$, then $X_1 \sim Poi(\lambda_1 + \lambda_0) = Poi(\lambda_1^*)$, $X_2 \sim Poi(\lambda_2 + \lambda_0) = Poi(\lambda_2^*)$. There are two equivalent ways of expressing the probability generating function (PGF) of $\{X_1, X_2\}$, see Kocherlakota and Kocherlakota (2014), namely

$$\Pi_1(t_1, t_2) = \exp\left[\lambda_1^*(t_1 - 1) + \lambda_2^*(t_2 - 1) + \lambda_0(t_1 - 1)(t_2 - 1)\right],$$
$$\Pi_2(t_1, t_2) = \exp\left[\lambda_1(t_1 - 1) + \lambda_2(t_2 - 1) + \lambda_0(t_1 t_2 - 1)\right].$$

Using $\Pi_2(t_1, t_2)$, the bivariate PGF of $\{X_{t,1}, X_{t,2}\}$ equals

Table 4 Simulated powers of univariate and bivariate dispersion index test

n	BINARCH(1)		BNBINAR(1)		CuBINAR(1)$_{(F)}$		CuBINAR(1)$_{(G)}$	
	uni	biv	uni	biv	uni	biv	uni	biv
100	0.3244	0.3315	0.9398	0.8874	0.8423	0.8345	0.2284	0.2429
250	0.6864	0.7114	0.9762	0.9365	0.9973	0.9968	0.4793	0.5322
500	0.9412	0.9454	0.9928	0.9456	0.9928	1	0.7755	0.8278
1000	0.9989	0.9989	0.9711	0.9903	0.9990	1	0.9711	0.9843
1500	1	0.9999	1	0.9982	1	1	0.9977	1

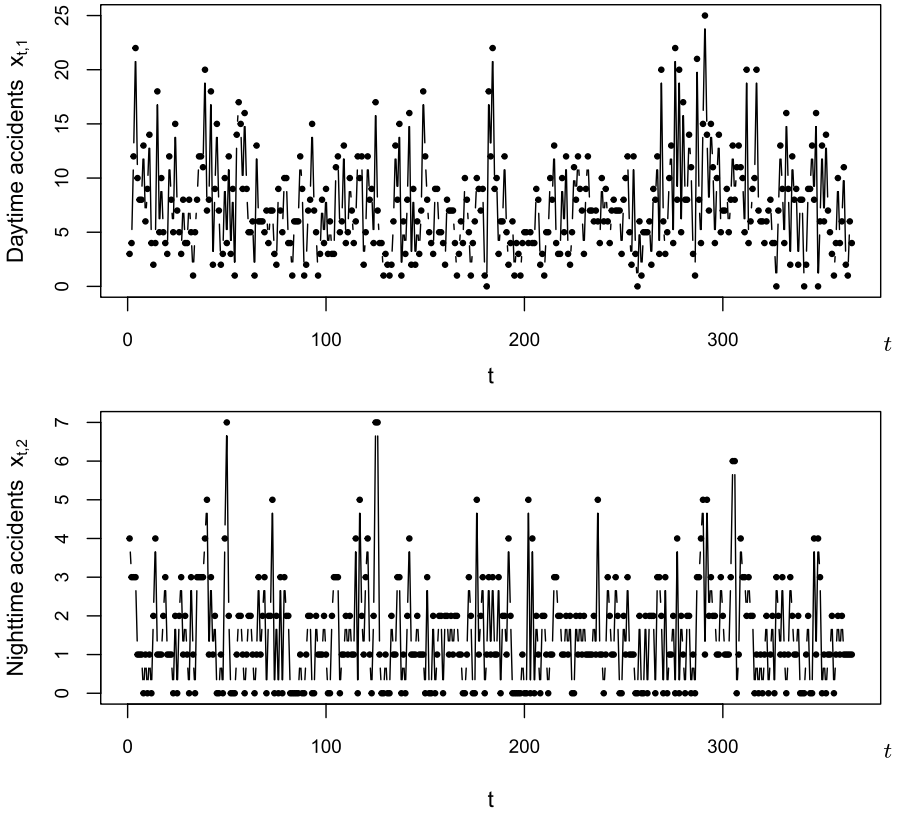

Fig. 1 Plots of the daily numbers of daytime and nighttime accidents data in the Schiphol area

$$
\begin{aligned}
E[a_1^{X_{t,1}} \cdot a_2^{X_{t,2}}] &= E[a_1^{\alpha_1 \circ X_{t-1,1} + \epsilon_{t,1}} \cdot a_2^{\alpha_2 \circ X_{t-1,2} + \epsilon_{t,2}}] \\
&= E[(1-\alpha_1+\alpha_1 a_1)^{X_{t-1,1}} \cdot (1-\alpha_2+\alpha_2 a_2)^{X_{t-1,2}}] \cdot E[a_1^{\epsilon_{t,1}} \cdot a_2^{\epsilon_{t,2}}] \\
&= \exp\Big\{ \lambda_1 \cdot (-\alpha_1+\alpha_1 a_1) + \lambda_2 \cdot (-\alpha_2+\alpha_2 a_2) \\
&\quad + \lambda_0 \big[(1-\alpha_1+\alpha_1 a_1)(1-\alpha_2+\alpha_2 a_2) - 1\big] \Big\} \cdot E[a_1^{\epsilon_{t,1}} \cdot a_2^{\epsilon_{t,2}}] \\
&= \exp\Big\{ \lambda_1 \alpha_1 (a_1 - 1) + \lambda_2 \alpha_2 (a_2 - 1) \\
&\quad + \lambda_0 \big[(a_1-1)\alpha_1 + (a_2-1)\alpha_2 + (a_1-1)(a_2-1)\alpha_1 \alpha_2\big] \Big\} \cdot E[a_1^{\epsilon_{t,1}} \cdot a_2^{\epsilon_{t,2}}],
\end{aligned}
$$

but we also have

$$
E[a_1^{X_{t,1}} \cdot a_2^{X_{t,2}}] = \exp\Big\{ \lambda_1 (a_1 - 1) + \lambda_2 (a_2 - 1) + \lambda_0 (a_1 a_2 - 1) \Big\}.
$$

Thus, the PGF of $\{\epsilon_{t,1}, \epsilon_{t,2}\}$ is derived as

$$E[a_1^{\varepsilon_{t,1}} \cdot a_2^{\varepsilon_{t,2}}] = \exp\Big\{ \lambda_1(a_1-1) + \lambda_2(a_2-1) + \lambda_0(a_1a_2-1) - \lambda_1\alpha_1(a_1-1) - \lambda_2\alpha_2(a_2-1) - \lambda_0\big[(a_1-1)\alpha_1 + (a_2-1)\alpha_2 + (a_1-1)(a_2-1)\alpha_1\alpha_2\big]\Big\}.$$

Using that

$$\begin{aligned} &-\lambda_0\big[(a_1-1)\alpha_1 + (a_2-1)\alpha_2 + (a_1-1)(a_2-1)\alpha_1\alpha_2\big] \\ &= -\lambda_0(a_1-1)\alpha_1 - \lambda_0(a_2-1)\alpha_2 - \lambda_0(a_1a_2 - a_1 - a_2 + 1)\alpha_1\alpha_2 \\ &= -\lambda_0(a_1-1)\alpha_1 - \lambda_0(a_2-1)\alpha_2 - \lambda_0[a_1a_2 - 1 - (a_1-1) - (a_2-1)]\alpha_1\alpha_2, \end{aligned}$$

we get

$$\begin{aligned} E[a_1^{\varepsilon_{t,1}} \cdot a_2^{\varepsilon_{t,2}}] &= \exp\Big\{ \lambda_1(a_1-1) + \lambda_2(a_2-1) + \lambda_0(a_1a_2-1) \\ &\quad - \lambda_1\alpha_1(a_1-1) - \lambda_2\alpha_2(a_2-1) \\ &\quad - \lambda_0(a_1-1)\alpha_1 - \lambda_0(a_2-1)\alpha_2 - \lambda_0[a_1a_2 - 1 - (a_1-1) - (a_2-1)]\alpha_1\alpha_2\Big\} \\ &= \exp\Big\{ (a_1-1)(\lambda_1 - \lambda_1\alpha_1 - \lambda_0\alpha_1 - \lambda_0\alpha_1\alpha_2) \\ &\quad + (a_2-1)(\lambda_2 - \lambda_2\alpha_2 - \lambda_0\alpha_2 - \lambda_0\alpha_1\alpha_2) \\ &\quad + (a_1a_2-1)(\lambda_0 - \lambda_0\alpha_1\alpha_2)\Big\} \\ &= \exp\Big\{ (a_1-1)[(\lambda_1+\lambda_0)(1-\alpha_1) - \lambda_0(1-\alpha_1\alpha_2)] \\ &\quad + (a_2-1)[(\lambda_2+\lambda_0)(1-\alpha_2) - \lambda_0(1-\alpha_1\alpha_2)] \\ &\quad + (a_1a_2-1)(\lambda_0 - \lambda_0\alpha_1\alpha_2)\Big\}. \end{aligned}$$

Using $\lambda_1^* = \lambda_1 + \lambda_0$, $\lambda_2^* = \lambda_2 + \lambda_0$, we get $\{\varepsilon_{t,1}, \varepsilon_{t,2}\} \sim BPoi(\lambda_1^*(1-\alpha_1) - \lambda_0(1-\alpha_1\alpha_2), \lambda_2^*(1-\alpha_2) - \lambda_0(1-\alpha_1\alpha_2), \lambda_0 - \lambda_0\alpha_1\alpha_2)$, which completes the proof of (2).

A.2 Proof of (15)

According to Kocherlakota and Kocherlakota (2014), the conditional moments of the BPoi-distribution are

$$E\big[X_1 \mid X_2 = x_2\big] = \lambda_1 + \frac{\lambda_0 x_2}{\lambda_2 + \lambda_0}, \qquad V\big[X_1 \mid X_2 = x_2\big] = \lambda_1 + \frac{\lambda_2\lambda_0 x_2}{(\lambda_2+\lambda_0)^2}.$$

Let us begin with $\mu_{1,2^2}(0) = E[X_{0,2}^2 X_{0,1}]$. We get

$$\begin{aligned} E\Big[X_{0,2}^2 \cdot X_{0,1}\Big] &= E\Big[X_{0,2}^2 \cdot E\big[X_{0,1} \mid X_{0,2}\big]\Big] = E\left[X_{0,2}^2 \cdot \left(\lambda_1 + \frac{\lambda_0 X_{0,2}}{\lambda_2+\lambda_0}\right)\right] \\ &= E\left[X_{0,2}^2\lambda_1 + X_{0,2}^3\frac{\lambda_0}{\lambda_2+\lambda_0}\right] = \lambda_1\mu_{2^2}(0) + \frac{\lambda_0}{\lambda_2+\lambda_0}(\mu_2 + 3\mu_2^2 + \mu_2^3). \end{aligned}$$

Similarly,

$$E\left[X_{0,1}^2 X_{0,2}\right] = \lambda_2 \mu_{1^2}(0) + \frac{\lambda_0}{\lambda_1+\lambda_0}(\mu_1 + 3\mu_1^2 + \mu_1^3).$$

Next, we turn to $\mu_{1^2,2^2}(0) = E\left[X_{0,1}^2 X_{0,2}^2\right]$, where

$$\begin{aligned}
E\left[X_{0,1}^2 X_{0,2}^2\right] &= E\left[E\left[X_{0,1}^2 \cdot X_{0,2}^2 \mid X_{0,1}\right]\right] = E\left[X_{0,1}^2 \cdot E\left[X_{0,2}^2 \mid X_{0,1}\right]\right] \\
&= E\left[X_{0,1}^2\left(\lambda_2 + \frac{\lambda_1\lambda_0 X_{0,1}}{(\lambda_1+\lambda_0)^2} + \lambda_2^2 + \frac{\lambda_0^2 X_{0,1}^2}{(\lambda_1+\lambda_0)^2} + 2\lambda_2\frac{\lambda_0 X_{0,1}}{\lambda_1+\lambda_0}\right)\right] \\
&= (\lambda_2+\lambda_2^2)E\left[X_{0,1}^2\right] + \frac{\lambda_1\lambda_0}{(\lambda_1+\lambda_0)^2}E\left[X_{0,1}^3\right] \\
&\quad + \frac{\lambda_0^2}{(\lambda_1+\lambda_0)^2}E\left[X_{0,1}^4\right] + \frac{2\lambda_2\lambda_0}{\lambda_1+\lambda_0}E\left[X_{0,1}^3\right] \\
&= (\lambda_2+\lambda_2^2)\mu_{1^2}(0) + \frac{\lambda_1\lambda_0}{\mu_1^2}\mu_{1^3}(0) + \frac{\lambda_0^2}{\mu_1^2}\mu_{1^4}(0) + \frac{2\lambda_2\lambda_0}{\mu_1}\mu_{1^3}(0).
\end{aligned}$$

Then, we get for $\mu_{1^3,2}(0) = E[X_{0,1}^3 X_{0,2}]$ that

$$\begin{aligned}
E[X_{0,1}^3 X_{0,2}] &= E\left[E\left[X_{0,1}^3 X_{0,2} \mid X_{0,1}\right]\right] = E\left[X_{0,1}^3 E\left[X_{0,2} \mid X_{0,1}\right]\right] \\
&= E\left[X_{0,1}^3\left(\lambda_2 + \frac{\lambda_0 X_{0,1}}{\lambda_1+\lambda_0}\right)\right] = E\left[X_{0,1}^3\lambda_2 + X_{0,1}^4\frac{\lambda_0}{\lambda_1+\lambda_0}\right] \\
&= \lambda_2 E\left[X_{0,1}^3\right] + \frac{\lambda_0}{\lambda_1+\lambda_0}E\left[X_{0,1}^4\right] = \lambda_2\mu_{1^3}(0) + \frac{\lambda_0}{\mu_1}\mu_{1^4}(0).
\end{aligned}$$

Similarly, $\mu_{1,2^3}(0) = E[X_{0,1}X_{0,2}^3]$ has the following expression:

$$\begin{aligned}
E[X_{0,2}^3 X_{0,1}] &= E\left[E\left[X_{0,2}^3 X_{0,1} \mid X_{0,2}\right]\right] = E\left[X_{0,2}^3 E\left[X_{0,1} \mid X_{0,2}\right]\right] \\
&= E\left[X_{0,2}^3\left(\lambda_1 + \frac{\lambda_0 X_{0,2}}{\lambda_2+\lambda_0}\right)\right] = E\left[X_{0,2}^3\lambda_1 + X_{0,2}^4\frac{\lambda_0}{\lambda_2+\lambda_0}\right] \\
&= \lambda_1 E\left[X_{0,2}^3\right] + \frac{\lambda_0}{\lambda_2+\lambda_0}E\left[X_{0,2}^4\right] = \lambda_1\mu_{2^3}(0) + \frac{\lambda_0}{\mu_2}\mu_{2^4}(0).
\end{aligned}$$

This completes the proof of (15).

A.3 Derivation of σ_{21}

We compute the explicit expression of σ_{21} as follows:

$$\sigma_{21} = E[(X_{0,1} - \mu_1)(X_{0,2} - \mu_2)] + \sum_{k=1}^{\infty} \Big(E[(X_{0,1} - \mu_1)(X_{k,2} - \mu_2)] + E[(X_{k,1} - \mu_1)(X_{0,2} - \mu_2)] \Big) = \mu_{1,2}^c(0) + \sum_{k=1}^{\infty} \Big(\mu_{1,2}^c(k) + \mu_{2,1}^c(k) \Big).$$

Here,

$$\begin{aligned}
\mu_{1,2}^c(k) &= E[(X_{0,1} - \mu_1)(X_{k,2} - \mu_2)] \\
&= E[(\alpha_2^k \circ X_{0,2} + \sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j} - \mu_2)(X_{0,1} - \mu_1)] \\
&= \alpha_2^k \cdot E[X_{0,1}X_{0,2}] + E[X_{0,1}]\left(\sum_{j=0}^{k-1} \alpha_2^j \cdot E[\varepsilon_{2,k-j}] \right) - \mu_2(E[X_{0,1}] - \mu_1) \\
&\quad - \mu_1 \cdot \alpha_2^k \cdot E[X_{0,2}] - \mu_1 \cdot \left(\sum_{j=0}^{k-1} \alpha_2^j \cdot E[\varepsilon_{2,k-j}] \right) \\
&= \alpha_2^k \cdot E[X_{0,1}X_{0,2}] - \mu_1 \cdot \alpha_2^k \cdot E[X_{0,2}] \\
&= \alpha_2^k \cdot \mu_{1,2}(0) - \mu_1\mu_2\alpha_2^k = \alpha_2^k \cdot \mu_{1,2}^c(0).
\end{aligned}$$

In a similar way, we get

$$\mu_{2,1}^c(k) = E[(X_{k,1} - \mu_1)(X_{0,2} - \mu_2)] = \alpha_1^k \cdot \mu_{1,2}(0) - \mu_1\mu_2\alpha_1^k = \alpha_1^k \cdot \mu_{1,2}^c(0).$$

Thus, the explicit expression of σ_{21} is

$$\begin{aligned}
\sigma_{21} &= \mu_{1,2}^c(0) + \sum_{k=1}^{\infty} \Big\{ \big[\alpha_2^k \cdot \mu_{1,2}^c(0)\big] + \big[\alpha_1^k \cdot \mu_{1,2}^c(0)\big] \Big\} \\
&= \mu_{1,2}^c(0) + \mu_{1,2}^c(0)\Big(\frac{\alpha_2}{1-\alpha_2} + \frac{\alpha_1}{1-\alpha_1} \Big).
\end{aligned}$$

A.4 Derivation of σ_{32} and σ_{41}

The expressions of σ_{32} and σ_{41} are as follows

$$\sigma_{32} = E[(X_{0,1}^2 - \mu_{1^2}(0))(X_{0,2} - \mu_2)] + \sum_{k=1}^{\infty} \Big(E[(X_{k,1}^2 - \mu_{1^2}(0))(X_{0,2} - \mu_2)] + E[(X_{0,1}^2 - \mu_{1^2}(0))(X_{k,2} - \mu_2)] \Big) = \mu^c_{1^2,2}(0) + \sum_{k=1}^{\infty} \Big(\mu^c_{2,1^2}(k) + \mu^c_{1^2,2}(k) \Big),$$

$$\sigma_{41} = E[(X_{0,2}^2 - \mu_{2^2}(0))(X_{0,1} - \mu_1)] + \sum_{k=1}^{\infty} \Big(E[(X_{0,2}^2 - \mu_{2^2}(0))(X_{k,1} - \mu_1)] + E[(X_{k,2}^2 - \mu_{2^2}(0))(X_{0,1} - \mu_1)] \Big) = \mu^c_{1,2^2}(0) + \sum_{k=1}^{\infty} \Big(\mu^c_{2^2,1}(k) + \mu^c_{1,2^2}(k) \Big).$$

Here,

$$\begin{aligned}
\mu^c_{2,1^2}(k) &= E[(X_{k,1}^2 - \mu_{1^2}(0))(X_{0,2} - \mu_2)] \\
&= E[\Big((\alpha_1^k \circ X_{0,1} + \sum_{j=0}^{k-1} \alpha_1^j \circ \varepsilon_{1,k-j})^2 - \mu_{1^2}(0) \Big)(X_{0,2} - \mu_2)] \\
&= E[(\alpha_1^k \circ X_{0,1})^2 X_{0,2}] - \mu_2 E[(\alpha_1^k \circ X_{0,1})^2] \\
&\quad + 2\Big(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}] \Big) E[(\alpha_1^k \circ X_{0,1})(X_{0,2} - \mu_2)] \\
&\quad + E\Big(\sum_{j=0}^{k-1} \alpha_1^j \circ \varepsilon_{1,k-j} \Big)^2 E[X_{0,2} - \mu_2] - \mu_{1^2}(0) E[X_{0,2} - \mu_2] \qquad (E[X_{0,2}] = \mu_2) \\
&= E[(\alpha_1^k \circ X_{0,1})^2 X_{0,2}] - \mu_2 E[(\alpha_1^k \circ X_{0,1})^2] \\
&\quad + 2\Big(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}] \Big) E[(\alpha_1^k \circ X_{0,1})(X_{0,2} - \mu_2)] \\
&= \alpha_1^{2k} E[X_{0,1}^2 X_{0,2}] + \alpha_1^k (1 - \alpha_1^k) E[X_{0,1} X_{0,2}] - \mu_2 \alpha_1^{2k} E[X_{0,1}^2] \\
&\quad - \alpha_1^k (1 - \alpha_1^k) E[X_{0,1}] \mu_2 + 2 \cdot \mu_{\varepsilon_1} \frac{1 - \alpha_1^k}{1 - \alpha_1} \alpha_1^k (E[X_{0,1} X_{0,2}] - \mu_2 E[X_{0,1}]) \\
&= \alpha_1^{2k} \mu_{1^2,2}(0) + \alpha_1^k (1 - \alpha_1^k) \mu_{1,2}(0) - \mu_2 \alpha_1^{2k} \mu_{1^2}(0) \\
&\quad - \alpha_1^k (1 - \alpha_1^k) \mu_1 \mu_2 + 2 \cdot \mu_{\varepsilon_1} \frac{1 - \alpha_1^k}{1 - \alpha_1} \alpha_1^k (\mu_{1,2}(0) - \mu_1 \mu_2) \\
&= \alpha_1^{2k} \mu^c_{1^2,2}(0) + \mu^c_{1,2}(0) \cdot \alpha_1^k (1 - \alpha_1^k) \cdot \Big(1 + \frac{2\mu_{\varepsilon_1}}{1 - \alpha_1} \Big) \\
&= \alpha_1^{2k} \mu^c_{1^2,2}(0) + \mu^c_{1,2}(0) \cdot \alpha_1^k (1 - \alpha_1^k) \cdot (1 + 2\mu_2).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mu^c_{1,2^2}(k) &= E[(X^2_{k,2} - \mu_{2^2}(0))(X_{0,1} - \mu_1)] \\
&= E[\Big((\alpha_2^k \circ X_{0,2} + \sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j})^2 - \mu_{2^2}(0)\Big)(X_{0,1} - \mu_1)] \\
&= E[(\alpha_2^k \circ X_{0,2})^2 X_{0,1}] - \mu_1 E[(\alpha_2^k \circ X_{0,2})^2] \\
&\quad + 2\left(\sum_{j=0}^{k-1} \alpha_2^j \cdot E[\varepsilon_{2,k-j}]\right) E[(\alpha_2^k \circ X_{0,2})(X_{0,1} - \mu_1)] \\
&= \alpha_2^{2k} E[X^2_{0,2} X_{0,1}] + \alpha_2^k (1 - \alpha_2^k) E[X_{0,1} X_{0,2}] - \mu_1 \alpha_2^{2k} E[X^2_{0,2}] \\
&\quad - \alpha_2^k (1 - \alpha_2^k) E[X_{0,2}]\mu_1 + 2 \cdot \mu_{\varepsilon_2} \frac{1 - \alpha_2^k}{1 - \alpha_2} \alpha_2^k (E[X_{0,1} X_{0,2}] - \mu_1 E[X_{0,2}]) \\
&= \alpha_2^{2k} \mu_{1,2^2}(0) + \alpha_2^k (1 - \alpha_2^k)\mu_{1,2}(0) - \mu_1 \alpha_2^{2k} \mu_{2^2}(0) - \alpha_2^k (1 - \alpha_2^k)\mu_2 \mu_1 \\
&\quad + 2 \cdot \mu_{\varepsilon_2} \frac{1 - \alpha_2^k}{1 - \alpha_2} \alpha_2^k (\mu_{1,2}(0) - \mu_1 \mu_2) \\
&= \alpha_2^{2k} \mu^c_{1,2^2}(0) + \mu^c_{1,2}(0) \cdot \alpha_2^k (1 - \alpha_2^k) \cdot \left(1 + \frac{2\mu_{\varepsilon_2}}{1 - \alpha_2}\right) \\
&= \alpha_2^{2k} \mu^c_{1,2^2}(0) + \mu^c_{1,2}(0) \cdot \alpha_2^k (1 - \alpha_2^k) \cdot (1 + 2\mu_2),
\end{aligned}$$

and

$$\begin{aligned}
\mu^c_{1^2,2}(k) &= E[(X^2_{0,1} - \mu_{1^2}(0))(X_{k,2} - \mu_2)] \\
&= E[(X^2_{0,1} - \mu_{1^2}(0))(\alpha_2^k \circ X_{0,2} + \sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j} - \mu_2)] \\
&= E[X^2_{0,1}(\alpha_2^k \circ X_{0,2})] - \mu_{1^2}(0) E[\alpha_2^k \circ X_{0,2}] \\
&\quad + (E[X^2_{0,1}] - \mu_{1^2}(0)) \left[\left(\sum_{j=0}^{k-1} \alpha_2^j \cdot E[\varepsilon_{2,k-j}]\right) - \mu_2\right] \\
&= \alpha_2^k E[X^2_{0,1} X_{0,2}] - \mu_{1^2}(0) \alpha_2^k E[X_{0,2}] \\
&= \alpha_2^k \mu_{1^2,2}(0) - \mu_{1^2}(0) \alpha_2^k \mu_2 = \alpha_2^k \cdot \mu^c_{1^2,2}(0).
\end{aligned}$$

Furthermore, we get

$$\begin{aligned}
\mu^c_{2^2,1}(k) &= E[(X^2_{0,2} - \mu_{2^2}(0))(X_{k,1} - \mu_1)] = \alpha_1^k E[X^2_{0,2} X_{0,1}] - \mu_{2^2}(0) \alpha_1^k E[X_{0,1}] \\
&= \alpha_1^k \mu_{1,2^2}(0) - \mu_{2^2}(0) \alpha_1^k \mu_1 = \alpha_1^k \cdot \mu^c_{1,2^2}(0).
\end{aligned}$$

Altogether, the explicit expressions of σ_{32} and σ_{41} are

$$\sigma_{32} = \mu^c_{1^2,2}(0) + \sum_{k=1}^{\infty} \left\{ \left[\alpha_1^{2k} \mu^c_{1^2,2}(0) + \mu^c_{1,2}(0) \cdot \alpha_1^k (1-\alpha_1^k) \cdot (1+2\mu_1) \right] + \left[\alpha_2^k \cdot \mu^c_{1^2,2}(0) \right] \right\}$$
$$= \mu^c_{1^2,2}(0) + \left(\frac{\alpha_1^2}{1-\alpha_1^2} + \frac{\alpha_2}{1-\alpha_2} \right) \mu^c_{1^2,2}(0) + \mu^c_{1,2}(0) \cdot (1+2\mu_1) \left(\frac{\alpha_1}{1-\alpha_1} - \frac{\alpha_1^2}{1-\alpha_1^2} \right),$$
$$\sigma_{41} = \mu^c_{1,2^2}(0) + \sum_{k=1}^{\infty} \left\{ \left[\alpha_1^k \cdot \mu^c_{1,2^2}(0) \right] + \left[\alpha_2^{2k} \mu^c_{1,2^2}(0) + \mu^c_{1,2}(0) \cdot \alpha_2^k (1-\alpha_2^k) \cdot (1+2\mu_2) \right] \right\}$$
$$= \mu^c_{1,2^2}(0) + \left(\frac{\alpha_2^2}{1-\alpha_2^2} + \frac{\alpha_1}{1-\alpha_1} \right) \mu^c_{1,2^2}(0) + \mu^c_{1,2}(0) \cdot (1+2\mu_2) \left(\frac{\alpha_2}{1-\alpha_2} - \frac{\alpha_2^2}{1-\alpha_2^2} \right).$$

A.5 Derivation of σ_{51} and σ_{52}

The expressions of σ_{51} and σ_{52} are as follows

$$\sigma_{51} = E[(X_{0,1}X_{0,2} - \mu_{1,2}(0))(X_{0,1} - \mu_1)]$$
$$+ \sum_{k=1}^{\infty} \Big(E[(X_{k,1}X_{k,2} - \mu_{1,2}(0))(X_{0,1} - \mu_1)] + E[(X_{0,1}X_{0,2} - \mu_{1,2}(0))(X_{k,1} - \mu_1)] \Big)$$
$$=: (\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0)) + \sum_{k=1}^{\infty} \Big(\mu^c_{1,12}(k) + \mu^c_{12,1}(k) \Big)$$
$$\sigma_{52} = E[(X_{0,1}X_{0,2} - \mu_{1,2}(0))(X_{0,2} - \mu_2)]$$
$$+ \sum_{k=1}^{\infty} \Big(E[(X_{k,1}X_{k,2} - \mu_{1,2}(0))(X_{0,2} - \mu_2)] + E[(X_{0,1}X_{0,2} - \mu_{1,2}(0))(X_{k,2} - \mu_2)] \Big)$$
$$=: \mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0) + \sum_{k=1}^{\infty} \Big(\mu^c_{2,12}(k) + \mu^c_{12,2}(k) \Big).$$

Here,

$$
\begin{aligned}
\mu^c_{1,12}(k) &= E[(X_{k,1}X_{k,2} - \mu_{1,2}(0))(X_{0,1} - \mu_1)] \\
&= E\left\{ \left[\left(\alpha_1^k \circ X_{0,1} + \sum_{j=0}^{k-1} \alpha_1^j \circ \varepsilon_{1,k-j} \right) \left(\alpha_2^k \circ X_{0,2} + \sum_{l=0}^{k-1} \alpha_2^l \circ \varepsilon_{2,k-l} \right) - \mu_{1,2}(0) \right] (X_{0,1} - \mu_1) \right\} \\
&= E[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2}) X_{0,1}] - \mu_1 E[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2})] \\
&\quad + \left(\sum_{l=0}^{k-1} \alpha_2^l \cdot E[\varepsilon_{2,k-l}] \right) E\left[(\alpha_1^k \circ X_{0,1})\, X_{0,1} \right] - \left(\sum_{j=0}^{k-1} \alpha_2^j \cdot E[\varepsilon_{2,k-j}] \right) \mu_1 E\left[(\alpha_1^k \circ X_{0,1}) \right] \\
&\quad + \left(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}] \right) E\left[(\alpha_2^k \circ X_{0,2}) X_{0,1} \right] - \left(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}] \right) \mu_1 E\left[(\alpha_2^k \circ X_{0,2}) \right] \\
&\quad + E\left[\left(\sum_{j=0}^{k-1} \alpha_1^j \circ \varepsilon_{1,k-j} \right) \left(\sum_{l=0}^{k-1} \alpha_2^l \circ \varepsilon_{2,k-l} \right) \right] \cdot E[X_{0,1} - \mu_1] - \mu_{1,2}(0) E[X_{0,1} - \mu_1] \\
&= E[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2}) X_{0,1}] - \mu_1 E[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2})] \\
&\quad + \left(\sum_{l=0}^{k-1} \alpha_2^l \cdot E[\varepsilon_{2,k-l}] \right) E\left[(\alpha_1^k \circ X_{0,1})\, X_{0,1} \right] - \left(\sum_{j=0}^{k-1} \alpha_2^j \cdot E[\varepsilon_{2,k-j}] \right) \mu_1 E\left[(\alpha_1^k \circ X_{0,1}) \right] \\
&\quad + \left(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}] \right) E\left[(\alpha_2^k \circ X_{0,2}) X_{0,1} \right] - \left(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}] \right) \mu_1 E\left[(\alpha_2^k \circ X_{0,2}) \right] \\
&= \alpha_1^k \alpha_2^k E[X_{0,1}^2 X_{0,2}] - \mu_1 \alpha_1^k \alpha_2^k E[X_{0,1} X_{0,2}] + \mu_{\varepsilon_2} \frac{1 - \alpha_2^k}{1 - \alpha_2} \alpha_1^k E[X_{0,1}^2] - \mu_{\varepsilon_2} \frac{1 - \alpha_2^k}{1 - \alpha_2} \alpha_1^k \mu_1^2 \\
&\quad + \mu_{\varepsilon_1} \frac{1 - \alpha_1^k}{1 - \alpha_1} \alpha_2^k E[X_{0,1} X_{0,2}] - \mu_{\varepsilon_1} \frac{1 - \alpha_1^k}{1 - \alpha_1} \alpha_2^k \mu_1 \mu_2 \\
&= \alpha_1^k \alpha_2^k \mu_{1^2,2}(0) - \mu_1 \alpha_1^k \alpha_2^k \mu_{1,2}(0) + \mu_{\varepsilon_2} \frac{1 - \alpha_2^k}{1 - \alpha_2} \alpha_1^k \mu_{1^2}(0) - \mu_{\varepsilon_2} \frac{1 - \alpha_2^k}{1 - \alpha_2} \alpha_1^k \mu_1^2 \\
&\quad + \mu_{\varepsilon_1} \frac{1 - \alpha_1^k}{1 - \alpha_1} \alpha_2^k \mu_{1,2}(0) - \mu_{\varepsilon_1} \frac{1 - \alpha_1^k}{1 - \alpha_1} \alpha_2^k \mu_1 \mu_2 \\
&= \alpha_1^k \alpha_2^k \cdot (\mu_{1^2,2}(0) - \mu_1 \mu_{1,2}(0)) \\
&\quad + \frac{\mu_{\varepsilon_2}}{1 - \alpha_2} (1 - \alpha_2^k) \alpha_1^k (\mu_{1^2}(0) - \mu_1^2) \\
&\quad + \frac{\mu_{\varepsilon_1}}{1 - \alpha_1} (1 - \alpha_1^k) \alpha_2^k (\mu_{1,2}(0) - \mu_1 \mu_2) \\
&= \alpha_1^k \alpha_2^k \cdot (\mu_{1^2,2}(0) - \mu_1 \mu_{1,2}(0)) + \tfrac{\mu_{\varepsilon_2}}{1 - \alpha_2} (1 - \alpha_2^k) \alpha_1^k (\mu_{1^2}(0) - \mu_1^2) + \tfrac{\mu_{\varepsilon_1}}{1 - \alpha_1} (1 - \alpha_1^k) \alpha_2^k \mu^c_{1,2}(0) \\
&= \alpha_1^k \alpha_2^k \cdot (\mu_{1^2,2}(0) - \mu_1 \mu_{1,2}(0)) + \mu_2 (1 - \alpha_2^k) \alpha_1^k (\mu_{1^2}(0) - \mu_1^2) + \mu_1 (1 - \alpha_1^k) \alpha_2^k \mu^c_{1,2}(0).
\end{aligned}
$$

In a similar way, we get

$$
\begin{aligned}
\mu^c_{2,12}(k) &= E[(X_{k,1}X_{k,2} - \mu_{1,2}(0))(X_{0,2} - \mu_2)] \\
&= E\left\{\left[\left(\alpha_1^k \circ X_{0,1} + \sum_{j=0}^{k-1} \alpha_1^j \circ \varepsilon_{1,k-j}\right)\left(\alpha_2^k \circ X_{0,2} + \sum_{l=0}^{k-1} \alpha_2^l \circ \varepsilon_{2,k-l}\right) - \mu_{1,2}(0)\right](X_{0,2} - \mu_2)\right\} \\
&= E[(\alpha_1^k \circ X_{0,2})(\alpha_2^k \circ X_{0,2})X_{0,2}] - \mu_2 E[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2})] \\
&\quad + \left(\sum_{l=0}^{k-1} \alpha_2^l \cdot E[\varepsilon_{2,k-l}]\right) E\left[(\alpha_1^k \circ X_{0,1})X_{0,2}\right] - \left(\sum_{j=0}^{k-1} \alpha_2^j \cdot E[\varepsilon_{2,k-j}]\right)\mu_2 E\left[(\alpha_1^k \circ X_{0,1})\right] \\
&\quad + \left(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}]\right) E\left[(\alpha_2^k \circ X_{0,2})X_{0,2}\right] - \left(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}]\right)\mu_2 E\left[(\alpha_2^k \circ X_{0,2})\right] \\
&= \alpha_1^k \alpha_2^k E[X_{0,2}^2 X_{0,1}] - \mu_2 \alpha_1^k \alpha_2^k E[X_{0,1}X_{0,2}] + \mu_{\varepsilon_2} \frac{1-\alpha_2^k}{1-\alpha_2} \alpha_1^k E[X_{0,1}X_{0,2}] - \mu_{\varepsilon_2} \frac{1-\alpha_2^k}{1-\alpha_2} \alpha_1^k \mu_1 \mu_2 \\
&\quad + \mu_{\varepsilon_1} \frac{1-\alpha_1^k}{1-\alpha_1} \alpha_2^k E[X_{0,2}^2] - \mu_{\varepsilon_1} \frac{1-\alpha_1^k}{1-\alpha_1} \alpha_2^k \mu_2^2 \\
&= \alpha_1^k \alpha_2^k \mu_{1,2^2}(0) - \mu_2 \alpha_1^k \alpha_2^k \mu_{1,2}(0) + \mu_{\varepsilon_2} \frac{1-\alpha_2^k}{1-\alpha_2} \alpha_1^k \mu_{1,2}(0) - \mu_{\varepsilon_2} \frac{1-\alpha_2^k}{1-\alpha_2} \alpha_1^k \mu_1 \mu_2 \\
&\quad + \mu_{\varepsilon_1} \frac{1-\alpha_1^k}{1-\alpha_1} \alpha_2^k \mu_{2^2}(0) - \mu_{\varepsilon_1} \frac{1-\alpha_1^k}{1-\alpha_1} \alpha_2^k \mu_2^2 \\
&= \alpha_1^k \alpha_2^k \cdot (\mu_{1,2^2}(0) - \mu_2 \mu_{1,2}(0)) \\
&\quad + \frac{\mu_{\varepsilon_2}}{1-\alpha_2}(1-\alpha_2^k)\alpha_1^k(\mu_{1,2}(0) - \mu_1\mu_2) \\
&\quad + \frac{\mu_{\varepsilon_1}}{1-\alpha_1}(1-\alpha_1^k)\alpha_2^k(\mu_{2^2}(0) - \mu_2^2) \\
&= \alpha_1^k \alpha_2^k \cdot (\mu_{1,2^2}(0) - \mu_2 \mu_{1,2}(0)) + \tfrac{\mu_{\varepsilon_2}}{1-\alpha_2}(1-\alpha_2^k)\alpha_1^k \mu^c_{1,2}(0) + \tfrac{\mu_{\varepsilon_1}}{1-\alpha_1}(1-\alpha_1^k)\alpha_2^k(\mu_{2^2}(0) - \mu_2^2) \\
&= \alpha_1^k \alpha_2^k \cdot (\mu_{1,2^2}(0) - \mu_2 \mu_{1,2}(0)) + \mu_2(1-\alpha_2^k)\alpha_1^k \mu^c_{1,2}(0) + \mu_1(1-\alpha_1^k)\alpha_2^k(\mu_{2^2}(0) - \mu_2^2).
\end{aligned}
$$

Next,

$$
\begin{aligned}
\mu^c_{12,1}(k) &= E[(X_{0,1}X_{0,2} - \mu_{1,2}(0))(X_{k,1} - \mu_1)] \\
&= E\left[(X_{0,1}X_{0,2} - \mu_{1,2}(0))\left(\alpha_1^k \circ X_{0,1} + \sum_{j=0}^{k-1} \alpha_1^j \circ \varepsilon_{1,k-j} - \mu_1\right)\right] \\
&= \alpha_1^k E[X_{0,1}^2 X_{0,2}] - \mu_{1,2}(0)\alpha_1^k E[X_{0,1}] \\
&= \alpha_1^k \cdot \mu_{1^2,2}(0) - \mu_{1,2}(0)\alpha_1^k \mu_1 = \alpha_1^k \cdot (\mu_{1^2,2}(0) - \mu_1 \mu_{1,2}(0)).
\end{aligned}
$$

Similarly,

$$\mu^c_{12,2}(k) = E[(X_{0,1}X_{0,2} - \mu_{1,2}(0))(X_{k,2} - \mu_2)] = \alpha_2^k E[X_{0,2}^2 X_{0,1}] - \mu_{1,2}(0)\alpha_2^k E[X_{0,2}]$$
$$= \alpha_2^k \cdot \mu_{1,2^2}(0) - \mu_{1,2}(0)\alpha_2^k \mu_2 = \alpha_2^k \cdot (\mu_{1,2^2}(0) - \mu_2 \mu_{1,2}(0)).$$

Altogether, the explicit expressions of σ_{51} and σ_{52} are

$$\begin{aligned}
\sigma_{51} &= \mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0) + \sum_{k=1}^{\infty}\Big\{[\alpha_1^k\alpha_2^k \cdot (\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0)) \\
&\quad + \mu_2(1-\alpha_2^k)\alpha_1^k(\mu_{1^2}(0) - \mu_1^2) + \mu_1(1-\alpha_1^k)\alpha_2^k\mu^c_{1,2}(0)] + \alpha_1^k \cdot (\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0))\Big\} \\
&= (\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0))\left(1 + \frac{\alpha_1}{1-\alpha_1}\right) + (\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0))\left(\frac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \frac{\alpha_1}{1-\alpha_1}\right) \\
&\quad + \mu_2\left(\frac{\alpha_1}{1-\alpha_1} - \frac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\right)(\mu_{1^2}(0) - \mu_1^2) + \mu_1(\frac{\alpha_2}{1-\alpha_2} - \frac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2})\mu^c_{1,2}(0) \\
\sigma_{52} &= \mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0) + \sum_{k=1}^{\infty}\Big\{[\alpha_1^k\alpha_2^k \cdot (\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0)) \\
&\quad + \mu_2(1-\alpha_2^k)\alpha_1^k\mu^c_{1,2}(0) + \mu_1(1-\alpha_1^k)\alpha_2^k(\mu_{2^2}(0) - \mu_2^2)] + \alpha_2^k \cdot \mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0)\Big\} \\
&= (\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0))\left(1 + \frac{\alpha_2}{1-\alpha_2}\right) + (\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0))\left(\frac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \frac{\alpha_1}{1-\alpha_1}\right) \\
&\quad + \mu_2\left(\frac{\alpha_1}{1-\alpha_1} - \frac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\right)\mu^c_{1,2}(0) + \mu_1\left(\frac{\alpha_2}{1-\alpha_2} - \frac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\right)(\mu_{2^2}(0) - \mu_2^2).
\end{aligned}$$

A.6 Derivation of σ_{43}

σ_{43} satisfies

$$\sigma_{43} = E\Big[\Big(X_{0,2}^2 - \mu_{2^2}(0)\Big)\Big(X_{0,1}^2 - \mu_{1^2}(0)\Big)\Big] + \sum_{k=1}^{\infty}\Big\{E\Big[\Big(X_{k,2}^2 - \mu_{2^2}(0)\Big)\Big(X_{0,1}^2 - \mu_{1^2}(0)\Big)\Big]\Big\}$$
$$+ E\Big[\Big(X_{0,2}^2 - \mu_{2^2}(0)\Big)\Big(X_{k,1}^2 - \mu_{1^2}(0)\Big)\Big] = \mu^c_{1^2,2^2}(0) + \sum_{k=1}^{\infty}\Big(\mu^c_{1^2,2^2}(k) + \mu^c_{2^2,1^2}(k)\Big).$$

Here,

$$
\begin{aligned}
\mu^c_{1^2,2^2}(k) &= E[(X^2_{k,2} - \mu_{2^2}(0))(X^2_{0,1} - \mu_{1^2}(0))] \\
&= E\left\{\left[\left(\alpha_2^k \circ X_{0,2} + \sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j}\right)^2 - \mu_{2^2}(0)\right](X^2_{0,1} - \mu_{1^2}(0))\right\} \\
&= E\left\{\left[(\alpha_2^k \circ X_{0,2})^2 + \left(\sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j}\right)^2 + 2(\alpha_2^k \circ X_{0,2})\left(\sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j}\right) - \mu_{2^2}(0)\right]\right. \\
&\quad \left.\left(X^2_{0,1} - \mu_{1^2}(0)\right)\right\} \\
&= E\left[\left(\alpha_2^k \circ X_{0,2}\right)^2 X^2_{0,1} - \mu_{1^2}(0)\left(\alpha_2^k \circ X_{0,2}\right)^2\right] \\
&\quad + E\left[2\left(\alpha_2^k \circ X_{0,2}\right)\left(\sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j}\right)X^2_{0,1}\right] - \mu_{1^2}(0)E\left[2\left(\alpha_2^k \circ X_{0,2}\right)\left(\sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j}\right)\right] \\
&= \alpha_2^{2k} E\left[X^2_{0,2}X^2_{0,1}\right] + \alpha_2^k(1-\alpha_2^k)E\left[X_{0,2}X^2_{0,1}\right] - \mu_{1^2}(0)\left(\alpha_2^{2k}E[X^2_{0,2}] + \alpha_2^k(1-\alpha_2^k)E[X_{0,2}]\right) \\
&\quad + 2\left(\sum_{j=0}^{k-1} \alpha_2^j E[\varepsilon_{2,k-j}]\right)\alpha_2^k E\left[X_{0,2}X^2_{0,1}\right] - 2\left(\sum_{j=0}^{k-1} \alpha_2^j E[\varepsilon_{2,k-j}]\right)\mu_{1^2}(0)\alpha_2^k E[X_{0,2}] \\
&= \alpha_2^{2k}\mu_{1^2,2^2}(0) + \alpha_2^k(1-\alpha_2^k)\mu_{1^2,2}(0) - \mu_{1^2}(0)\left(\alpha_2^{2k}\mu_{2^2}(0) + \alpha_2^k(1-\alpha_2^k)\mu_2\right) \\
&\quad + 2\mu_{\varepsilon_2}\left(\sum_{j=0}^{k-1} \alpha_2^j\right)\alpha_2^k\mu_{1^2,2}(0) - 2\mu_{\varepsilon_2}\left(\sum_{j=0}^{k-1} \alpha_2^j\right)\mu_{1^2}(0)\alpha_2^k\mu_2 \\
&= \alpha_2^{2k}(\mu_{1^2,2^2}(0) - \mu_{1^2}(0)\mu_{2^2}(0)) + \alpha_2^k(1-\alpha_2^k)\left(\mu_{1^2,2}(0) - \mu_{1^2}(0)\mu_2\right) \\
&\quad + 2\mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_2^k\left(\mu_{1^2,2}(0) - \mu_{1^2}(0)\mu_2\right) \\
&= \alpha_2^{2k}\mu^c_{1^2,2^2}(0) + \alpha_2^k(1-\alpha_2^k)\mu^c_{1^2,2}(0) + 2\mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_2^k\mu^c_{1^2,2}(0) \\
&= \alpha_2^{2k}\mu^c_{1^2,2^2}(0) + \alpha_2^k\left(1-\alpha_2^k\right)\mu^c_{1^2,2}(0)(1+2\mu_2\Big).
\end{aligned}
$$

Thus,

$$\sum_{k=1}^{\infty}\mu^c_{1^2,2^2}(k)=\sum_{k=1}^{\infty}\left\{\alpha_2^{2k}\cdot\mu^c_{1^2,2^2}(0)+\alpha_2^k(1-\alpha_2^k)\mu^c_{1^2,2}(0)(1+2\mu_2)\right\}$$
$$=\frac{\alpha_2^2}{1-\alpha_2^2}\mu^c_{1^2,2^2}(0)+\mu^c_{1^2,2}(0)(1+2\mu_2)\left(\frac{\alpha_2}{1-\alpha_2}-\frac{\alpha_2^2}{1-\alpha_2^2}\right).$$

Similarly, we have

$$\begin{aligned}
\mu^c_{2^2,1^2}(k)&=E\left[\left(X^2_{0,2}-\mu_{2^2}(0)\right)\left(X^2_{k,1}-\mu_{1^2}(0)\right)\right]\\
&=E\left\{\left[\left(\alpha_1^k\circ X_{0,1}+\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\right)^2-\mu_{1^2}(0)\right]\left(X^2_{0,2}-\mu_{2^2}(0)\right)\right\}\\
&=E\left\{\left[(\alpha_1^k\circ X_{0,1})^2+\left(\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\right)^2+2(\alpha_1^k\circ X_{0,1})\left(\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\right)\right.\right.\\
&\quad\left.\left.-\mu_{1^2}(0)\right](X^2_{0,2}-\mu_{2^2}(0))\right\}\\
&=E\left[(\alpha_1^k\circ X_{0,1})^2X^2_{0,2}-\mu_{2^2}(0)(\alpha_1^k\circ X_{0,1})^2\right]\\
&\quad+E\left[2\cdot(\alpha_1^k\circ X_{0,1})\left(\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\right)X^2_{0,2}\right]-\mu_{2^2}(0)E\left[2(\alpha_1^k\circ X_{0,1})\left(\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\right)\right]\\
&=\alpha_1^{2k}E\left[X^2_{0,2}X^2_{0,1}\right]+\alpha_1^k(1-\alpha_1^k)E\left[X_{0,1}X^2_{0,2}\right]-\mu_{2^2}(0)\left[\alpha_1^{2k}E\left[X^2_{0,1}\right]+\alpha_1^k(1-\alpha_1^k)E[X_{0,1}]\right]\\
&\quad+2\left(\sum_{j=0}^{k-1}\alpha_1^jE[\varepsilon_{1,k-j}]\right)\alpha_1^kE[X_{0,1}X^2_{0,2}]-2\left(\sum_{j=0}^{k-1}\alpha_1^jE[\varepsilon_{1,k-j}]\right)\mu_{2^2}(0)\alpha_1^kE[X_{0,1}]\\
&=\alpha_1^{2k}\mu_{1^2,2^2}(0)+\alpha_1^k(1-\alpha_1^k)\mu_{1,2^2}(0)-\mu_{2^2}(0)(\alpha_1^{2k}\mu_{1^2}(0)+\alpha_1^k(1-\alpha_1^k)\mu_1)\\
&\quad+2\mu_{\varepsilon_1}\left(\sum_{j=0}^{k-1}\alpha_1^j\right)\alpha_1^k\mu_{1,2^2}(0)-2\mu_{\varepsilon_1}\left(\sum_{j=0}^{k-1}\alpha_1^j\right)\mu_{2^2}(0)\alpha_1^k\mu_1\\
&=\alpha_1^{2k}\left(\mu_{1^2,2^2}(0)-\mu_{1^2}(0)\mu_{2^2}(0)\right)+\alpha_1^k(1-\alpha_1^k)\left(\mu_{1,2^2}(0)-\mu_{2^2}(0)\mu_1\right)\\
&\quad+2\mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_1^k(\mu_{1,2^2}(0)-\mu_{2^2}(0)\mu_1)\\
&=\alpha_1^{2k}\mu^c_{1^2,2^2}(0)+\alpha_1^k\left(1-\alpha_1^k\right)\mu^c_{1,2^2}(0)+2\mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_1^k\mu^c_{1,2^2}(0)\\
&=\alpha_1^{2k}\mu^c_{1^2,2^2}(0)+\alpha_1^k(1-\alpha_1^k)\mu^c_{1,2^2}(0)\left(1+\frac{2\mu_{\varepsilon_1}}{1-\alpha_1}\right)\\
&=\alpha_1^{2k}\mu^c_{1^2,2^2}(0)+\alpha_1^k(1-\alpha_1^k)\mu^c_{1,2^2}(0)(1+2\mu_1).
\end{aligned}$$

Thus,

$$\sum_{k=1}^{\infty}\mu^c_{2^2,1^2}(k)=\sum_{k=1}^{\infty}\{\alpha_1^{2k}\cdot\mu^c_{1^2,2^2}(0)+\alpha_1^k(1-\alpha_1^k)\mu^c_{1,2^2}(0)(1+2\mu_1)\}$$
$$=\frac{\alpha_1^2}{1-\alpha_1^2}\mu^c_{1^2,2^2}(0)+\mu^c_{1,2^2}(0)(1+2\mu_1)\left(\frac{\alpha_1}{1-\alpha_1}-\frac{\alpha_1^2}{1-\alpha_1^2}\right).$$

Altogether, the explicit expression of σ_{43} is

$$\sigma_{43}=\mu^c_{1^2,2^2}(0)+\frac{\alpha_2^2}{1-\alpha_2^2}\mu^c_{1^2,2^2}(0)+\mu^c_{1^2,2}(0)(1+2\mu_2)\left(\frac{\alpha_2}{1-\alpha_2}-\frac{\alpha_2^2}{1-\alpha_2^2}\right)$$
$$+\frac{\alpha_1^2}{1-\alpha_1^2}\mu^c_{1^2,2^2}(0)+\mu^c_{1,2^2}(0)(1+2\mu_1)\left(\frac{\alpha_1}{1-\alpha_1}-\frac{\alpha_1^2}{1-\alpha_1^2}\right).$$

A.7 Derivation of σ_{53} and σ_{54}

For σ_{53} and σ_{54}, we get

$$\begin{aligned}\sigma_{53}&=E\Big[(X_{0,1}X_{0,2}-\mu_{1,2}(0))\Big(X^2_{0,1}-\mu_{1^2}(0)\Big)\Big]\\&\quad+\sum_{k=1}^{\infty}\Big\{E\Big[(X_{0,1}X_{0,2}-\mu_{1,2}(0))\Big(X^2_{k,1}-\mu_{1^2}(0)\Big)\Big]\\&\quad+E\Big[(X_{k,1}X_{k,2}-\mu_{1,2}(0))(X^2_{0,1}-\mu_{1^2}(0))\Big]\Big\}\\&=\mu_{1^3,2}(0)-\mu_{1^2}(0)\mu_{1,2}(0)+\sum_{k=1}^{\infty}\Big\{\mu^c_{12,1^2}(k)+\mu^c_{1^2,12}(k)\Big\}\\\sigma_{54}&=E\Big[(X_{0,1}X_{0,2}-\mu_{1,2}(0))\Big(X^2_{0,2}-\mu_{2^2}(0)\Big)\Big]\\&\quad+\sum_{k=1}^{\infty}\Big\{E\Big[(X_{k,1}X_{k,2}-\mu_{1,2}(0))\Big(X^2_{0,2}-\mu_{2^2}(0)\Big)\Big]\\&\quad+E\Big[(X_{0,1}X_{0,2}-\mu_{1,2}(0))(X^2_{k,2}-\mu_{2^2}(0))\Big]\Big\}\\&=\mu_{1,2^3}(0)-\mu_{2^2}(0)\mu_{1,2}(0)+\sum_{k=1}^{\infty}\Big(\mu^c_{2^2,12}(k)+\mu^c_{12,2^2}(k)\Big).\end{aligned}$$

Here,

$$
\begin{aligned}
\mu^c_{1^2,12}(k) &= E\Big[(X_{k,1}X_{k,2} - \mu_{1,2}(0))\Big(X^2_{0,1} - \mu_{1^2}(0)\Big)\Big] \\
&= E\Bigg\{\Bigg[\Bigg(\alpha_1^k \circ X_{0,1} + \sum_{j=0}^{k-1}\alpha_1^j \circ \varepsilon_{1,k-j}\Bigg)\Bigg(\alpha_2^k \circ X_{0,2} + \sum_{l=0}^{k-1}\alpha_2^l \circ \varepsilon_{2,k-l}\Bigg) \\
&\quad -\mu_{1,2}(0)\Big]\Big(X^2_{0,1} - \mu_{1^2}(0)\Big)\Bigg\} \\
&= E\Big[\big(\alpha_1^k \circ X_{0,1}\big)\big(\alpha_2^k \circ X_{0,2}\big)X^2_{0,1}\Big] - \mu_{1^2}(0)E[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2})] \\
&\quad + \Bigg(\sum_{l=0}^{k-1}\alpha_2^l E[\varepsilon_{2,k-l}]\Bigg)E\Big[(\alpha_1^k \circ X_{0,1})X^2_{0,1}\Big] - \Bigg(\sum_{l=0}^{k-1}\alpha_2^l E[\varepsilon_{2,k-l}]\Bigg)\mu_{1^2}(0)E\big[(\alpha_1^k \circ X_{0,1})\big] \\
&\quad + \Bigg(\sum_{j=0}^{k-1}\alpha_1^j E[\varepsilon_{1,k-j}]\Bigg)E\Big[(\alpha_2^k \circ X_{0,2})X^2_{0,1}\Big] - \Bigg(\sum_{j=0}^{k-1}\alpha_1^j E[\varepsilon_{1,k-j}]\Bigg)\mu_{1^2}(0)E\big[(\alpha_2^k \circ X_{0,2})\big] \\
&\quad + E\Bigg[\Bigg(\sum_{j=0}^{k-1}\alpha_1^j \circ \varepsilon_{1,k-j}\Bigg)\Bigg(\sum_{l=0}^{k-1}\alpha_2^l \circ \varepsilon_{2,k-l}\Bigg)\Bigg]E[X^2_{0,1} - \mu_{1^2}(0)] - \mu_{1,2}(0)E[X^2_{0,1} - \mu_{1^2}(0)] \\
&= \alpha_1^k\alpha_2^k E\Big[X^3_{0,1}X_{0,2}\Big] - \mu_{1^2}(0)\alpha_1^k\alpha_2^k E[X_{0,1}X_{0,2}] + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k E\Big[X^3_{0,1}\Big] \\
&\quad - \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k\mu_1\mu_{1^2}(0) + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k E\Big[X^2_{0,1}X_{0,2}\Big] - \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k\mu_{1^2}(0)\mu_2 \\
&= \alpha_1^k\alpha_2^k\mu_{1^3,2}(0) - \mu_{1^2}(0)\alpha_1^k\alpha_2^k\mu_{1,2}(0) + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k\mu_{1^3}(0) - \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k\mu_1\mu_{1^2}(0) \\
&\quad + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k\mu_{1^2,2}(0) - \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k\mu_{1^2}(0)\mu_2 \\
&= \alpha_1^k\alpha_2^k\Big(\mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0)\Big) \\
&\quad + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k\Big(\mu_{1^3}(0) - \mu_1\mu_{1^2}(0)\Big) + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k\mu^c_{1^2,2}(0).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mu^c_{2^2,12}(k) &= E\Big[(X_{k,1}X_{k,2} - \mu_{1,2}(0))\Big(X^2_{0,2} - \mu_{2^2}(0)\Big)\Big] \\
&= E\Bigg\{\Bigg[\Bigg(\alpha_1^k \circ X_{0,1} + \sum_{j=0}^{k-1} \alpha_1^j \circ \varepsilon_{1,k-j}\Bigg)\Bigg(\alpha_2^k \circ X_{0,2} + \sum_{l=0}^{k-1} \alpha_2^l \circ \varepsilon_{2,k-l}\Bigg) \\
&\quad -\mu_{1,2}(0)\Big](X^2_{0,2} - \mu_{2^2}(0))\Bigg\} \\
&= E\Big[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2})X^2_{0,2}] - \mu_{2^2}(0)E[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2})\Big] \\
&\quad + \Bigg(\sum_{l=0}^{k-1} \alpha_2^l \cdot E[\varepsilon_{2,k-l}]\Bigg)E\Big[(\alpha_1^k \circ X_{0,1})X^2_{0,2}\Big] - \Bigg(\sum_{l=0}^{k-1} \alpha_2^l \cdot E[\varepsilon_{2,k-l}]\Bigg)\mu_{2^2}(0)E\Big[(\alpha_1^k \circ X_{0,1})\Big] \\
&\quad + \Bigg(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}]\Bigg)E\Big[(\alpha_2^k \circ X_{0,2})X^2_{0,2}\Big] - \Bigg(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}]\Bigg)\mu_{2^2}(0)E\Big[(\alpha_2^k \circ X_{0,2})\Big] \\
&= \alpha_1^k\alpha_2^k E\Big[X_{0,1}X^3_{0,2}\Big] - \mu_{2^2}(0)\alpha_1^k\alpha_2^k E\Big[X_{0,1}X_{0,2}\Big] + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k E\Big[X_{0,1}X^2_{0,2}\Big] \\
&\quad - \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k\mu_1\mu_{2^2}(0) + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k E\Big[X^3_{0,2}\Big] - \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k\mu_{2^2}(0)\mu_2 \\
&= \alpha_1^k\alpha_2^k\mu_{1,2^3}(0) - \mu_{2^2}(0)\alpha_1^k\alpha_2^k\mu_{1,2}(0) + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k\mu_{1,2^2}(0) \\
&\quad - \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k\mu_1\mu_{2^2}(0) + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k\mu_{2^3}(0) - \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k\mu_{2^2}(0)\mu_2 \\
&= \alpha_1^k\alpha_2^k\Big(\mu_{1,2^3}(0) - \mu_{2^2}(0)\mu_{1,2}(0)\Big) \\
&\quad + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k\Big(\mu_{2^3}(0) - \mu_2\mu_{2^2}(0)\Big) + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k\mu^c_{1,2^2}(0).
\end{aligned}
$$

Next,

$$
\begin{aligned}
\mu^{c}_{12,1^2}(k) &= E\Big[(X_{0,1}X_{0,2}-\mu_{1,2}(0))\Big(X^2_{k,1}-\mu_{1^2}(0)\Big)\Big]\\
&= E\Bigg[(X_{0,1}X_{0,2}-\mu_{1,2}(0))\Bigg(\bigg(\alpha_1^k\circ X_{0,1}+\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\bigg)^2-\mu_{1^2}(0)\Bigg)\Bigg]\\
&= E\Bigg[(X_{0,1}X_{0,2}-\mu_{1,2}(0))\Bigg((\alpha_1^k\circ X_{0,1})^2+\bigg(\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\bigg)^2\\
&\quad +2(\alpha_1^k\circ X_{0,1})\bigg(\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\bigg)-\mu_{1^2}(0)\Bigg)\Bigg]\\
&= E\Bigg[(X_{0,1}X_{0,2}-\mu_{1,2}(0))\Bigg((\alpha_1^k\circ X_{0,1})^2+E[X_{0,1}X_{0,2}-\mu_{1,2}(0)]E\bigg(\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\bigg)^2\Bigg)\Bigg]\\
&\quad +2E\Bigg[(X_{0,1}X_{0,2}-\mu_{1,2}(0))(\alpha_1^k\circ X_{0,1})\Big(\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\Big)\Bigg]-E[X_{0,1}X_{0,2}-\mu_{1,2}(0)]\cdot\mu_{1^2}(0)\\
&= E\Big[(X_{0,1}X_{0,2}-\mu_{1,2}(0))\big(\alpha_1^k\circ X_{0,1}\big)^2\Big]\\
&\quad +2E\Bigg[(X_{0,1}X_{0,2}-\mu_{1,2}(0))(\alpha_1^k\circ X_{0,1})\bigg(\sum_{j=0}^{k-1}\alpha_1^j\circ\varepsilon_{1,k-j}\bigg)\Bigg]\\
&= E\big[(X_{0,1}X_{0,2}(\alpha_1^k\circ X_{0,1})^2\big]-\mu_{1,2}(0)E\big[(\alpha_1^k\circ X_{0,1})^2\big]\\
&\quad +2\bigg(\sum_{j=0}^{k-1}\alpha_1^j\cdot E[\varepsilon_{1,k-j}]\bigg)\big(E[X_{0,1}X_{0,2}\big(\alpha_1^k\circ X_{0,1}\big)]-\mu_{1,2}(0)E\big[\alpha_1^k\circ X_{0,1}\big]\big)\\
&= \alpha_1^{2k}E\Big[X^3_{0,1}X_{0,2}\Big]+\alpha_1^k(1-\alpha_1^k)E\Big[X^2_{0,1}X_{0,2}\Big]-\mu_{1,2}(0)\Big(\alpha_1^{2k}E[X^2_{0,1}]+\alpha_1^k(1-\alpha_1^k)E[X_{0,1}]\Big)\\
&\quad +2\mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_1^kE\Big[X^2_{0,1}X_{0,2}\Big]-\mu_{1,2}(0)\alpha_1^kE[X_{0,1}]\\
&= \alpha_1^{2k}\mu_{1^3,2}(0)+\alpha_1^k(1-\alpha_1^k)\mu_{1^2,2}(0)-\mu_{1,2}(0)\alpha_1^{2k}\mu_{1^2}(0)-\mu_{1,2}(0)\alpha_1^k(1-\alpha_1^k)\mu_1\\
&\quad +2\mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\big(\alpha_1^k\mu_{1^2,2}(0)-\mu_{1,2}(0)\alpha_1^k\mu_1\big)\\
&= \alpha_1^{2k}\Big(\mu_{1^3,2}(0)-\mu_{1^2}(0)\mu_{1,2}(0)\Big)+\alpha_1^k(1-\alpha_1^k)\Big(\mu_{1^2,2}(0)-\mu_1\mu_{1,2}(0)\Big)\\
&\quad +2\mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_1^k\Big(\mu_{1^2,2}(0)-\mu_1\mu_{1,2}(0)\Big).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mu^c_{12,2^2}(k) &= E[(X_{0,1}X_{0,2} - \mu_{1,2}(0))\Big(X^2_{k,2} - \mu_{2^2}(0)\Big)] \\
&= E\Bigg[(X_{0,1}X_{0,2} - \mu_{1,2}(0))\Bigg(\Bigg(\alpha_2^k \circ X_{0,2} + \sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j}\Bigg)^2 - \mu_{2^2}(0)\Bigg)\Bigg] \\
&= E\Bigg[(X_{0,1}X_{0,2} - \mu_{1,2}(0))\Bigg((\alpha_2^k \circ X_{0,2})^2 + \Bigg(\sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j}\Bigg)^2 \\
&\quad + 2(\alpha_2^k \circ X_{0,2})\Bigg(\sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j}\Bigg) - \mu_{2^2}(0)\Bigg)\Bigg] \\
&= E\big[(X_{0,1}X_{0,2} - \mu_{1,2}(0))(\alpha_1^k \circ X_{0,2})^2\big] \\
&\quad + 2E\Bigg[(X_{0,1}X_{0,2} - \mu_{1,2}(0))(\alpha_2^k \circ X_{0,2})\Bigg(\sum_{j=0}^{k-1} \alpha_2^j \circ \varepsilon_{2,k-j}\Bigg)\Bigg] \\
&= E[(X_{0,1}X_{0,2}(\alpha_2^k \circ X_{0,2})^2] - \mu_{1,2}(0)E[(\alpha_2^k \circ X_{0,2})^2] \\
&\quad + 2\Bigg(\sum_{j=0}^{k-1} \alpha_2^j \cdot E[\varepsilon_{2,k-j}]\Bigg)\big(E[X_{0,1}X_{0,2}(\alpha_2^k \circ X_{0,2})] - \mu_{1,2}(0)E[\alpha_2^k \circ X_{0,2}]\big) \\
&= \alpha_2^{2k}E\Big[X^3_{0,2}X_{0,1}\Big] + \alpha_2^k\big(1-\alpha_2^k\big)E\Big[X^2_{0,2}X_{0,1}\Big] - \mu_{1,2}(0)\Big(\alpha_2^{2k}E[X^2_{0,2}] + \alpha_2^k\big(1-\alpha_2^k\big)E[X_{0,2}]\Big) \\
&\quad + 2\mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_2^k E[X_{0,1}X^2_{0,2}] - \mu_{1,2}(0)\alpha_2^k E[X_{0,2}] \\
&= \alpha_2^{2k}\mu_{1,2^3}(0) + \alpha_2^k(1-\alpha_2^k)\mu_{1,2^2}(0) - \mu_{1,2}(0)\alpha_2^{2k}\mu_{2^2}(0) - \mu_{1,2}(0)\alpha_2^k(1-\alpha_2^k)\mu_2 \\
&\quad + 2\mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\big(\alpha_2^k\mu_{1,2^2}(0) - \mu_{1,2}(0)\alpha_2^k\mu_2\big) \\
&= \alpha_2^{2k}\Big(\mu_{2^3,1}(0) - \mu_{2^2}(0)\mu_{1,2}(0)\Big) + \alpha_2^k(1-\alpha_2^k)\Big(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0)\Big) \\
&\quad + 2\mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_2^k\Big(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0)\Big).
\end{aligned}
$$

Altogether, the explicit expressions of σ_{53} and σ_{54} are

$$
\begin{aligned}
\sigma_{53} &= \mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0) + \sum_{k=1}^{\infty}\Big\{\alpha_1^k\alpha_2^k\cdot(\mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0)) \\
&\quad + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k(\mu_{1^3}(0) - \mu_1\mu_{1^2}(0)) + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k\cdot\mu^c_{1^2,2}(0) \\
&\quad + \alpha_1^{2k}\cdot(\mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0)) + \alpha_1^k(1-\alpha_1^k)(\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0)) \\
&\quad + 2\mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_1^k(\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0))\Big\} \\
&= \mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0) + (\mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0))(\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \tfrac{\alpha_1^2}{1-\alpha_1^2}) \\
&\quad + (\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0))(\tfrac{\alpha_1}{1-\alpha_1} + \tfrac{\alpha_1^2}{1-\alpha_1^2})(1 + \tfrac{2\mu_{\varepsilon_1}}{1-\alpha_1}) \\
&\quad + \tfrac{\mu_{\varepsilon_2}}{1-\alpha_2}(\mu_{1^3}(0) - \mu_1\mu_{1^2}(0))(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}) + \tfrac{\mu_{\varepsilon_1}}{1-\alpha_1}\mu^c_{1^2,2}(0)(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}) \\
&= \mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0) + (\mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0))(\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \tfrac{\alpha_1^2}{1-\alpha_1^2}) \\
&\quad + (\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0))(\tfrac{\alpha_1}{1-\alpha_1} + \tfrac{\alpha_1^2}{1-\alpha_1^2})(1 + 2\mu_1) \\
&\quad + \mu_2(\mu_{1^3}(0) - \mu_1\mu_{1^2}(0))(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}) + \mu_1\mu^c_{1^2,2}(0)(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}),
\end{aligned}
$$

$$
\begin{aligned}
\sigma_{54} &= \mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0) + \sum_{k=1}^{\infty}\Big\{\alpha_1^k\alpha_2^k\cdot(\mu_{1,2^3}(0) - \mu_{2^2}(0)\mu_{1,2}(0)) \\
&\quad + \mu_{\varepsilon_1}\tfrac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k(\mu_{2^3}(0) - \mu_2\mu_{2^2}(0)) + \mu_{\varepsilon_2}\tfrac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k\cdot\mu^c_{1,2^2}(0) \\
&\quad + \alpha_2^{2k}\cdot(\mu_{1,2^3}(0) - \mu_{2^2}(0)\mu_{1,2}(0)) + \alpha_2^k(1-\alpha_2^k)(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0)) \\
&\quad + 2\mu_{\varepsilon_2}\tfrac{1-\alpha_2^k}{1-\alpha_2}\alpha_2^k(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0))\Big\} \\
&= \mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0) + (\mu_{1,2^3}(0) - \mu_{2^2}(0)\mu_{1,2}(0))(\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \tfrac{\alpha_2^2}{1-\alpha_2^2}) \\
&\quad + (\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0))(\tfrac{\alpha_2}{1-\alpha_2} + \tfrac{\alpha_2^2}{1-\alpha_2^2})(1 + \tfrac{2\mu_{\varepsilon_2}}{1-\alpha_2}) \\
&\quad + \tfrac{\mu_{\varepsilon_1}}{1-\alpha_1}(\mu_{2^3}(0) - \mu_2\mu_{2^2}(0))(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}) + \tfrac{\mu_{\varepsilon_2}}{1-\alpha_2}\mu^c_{1,2^2}(0)(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}) \\
&= \mu_{1^3,2}(0) - \mu_{1^2}(0)\mu_{1,2}(0) + (\mu_{1,2^3}(0) - \mu_{2^2}(0)\mu_{1,2}(0))(\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} + \tfrac{\alpha_2^2}{1-\alpha_2^2}) \\
&\quad + (\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0))(\tfrac{\alpha_2}{1-\alpha_2} + \tfrac{\alpha_2^2}{1-\alpha_2^2})(1 + 2\mu_2) \\
&\quad + \mu_1(\mu_{2^3}(0) - \mu_2\mu_{2^2}(0))(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}) + \mu_2\mu^c_{1,2^2}(0)(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}).
\end{aligned}
$$

A.8 Derivation of σ_{55}

We compute σ_{55} as

$$\sigma_{55} = E[(X_{0,1}X_{0,2} - \mu_{1,2}(0))^2] + 2\sum_{k=1}^{\infty} E[(X_{k,1}X_{k,2} - \mu_{1,2}(0))(X_{0,1}X_{0,2} - \mu_{1,2}(0))]$$
$$= \mu_{1^2,2^2}(0) - \mu_{1,2}^2(0) + 2\sum_{k=1}^{\infty} \mu_{12,12}^c(k).$$

Here,

$$\begin{aligned}
\mu_{12,12}^c(k) &= E[(X_{k,1}X_{k,2} - \mu_{1,2}(0))(X_{0,1}X_{0,2} - \mu_{1,2}(0))] \\
&= E\left[\left(\alpha_1^k \circ X_{0,1} + \sum_{j=0}^{k-1} \alpha_1^j \circ \varepsilon_{1,k-j}\right)\left(\alpha_2^k \circ X_{0,2} + \sum_{l=0}^{k-1} \alpha_2^l \circ \varepsilon_{2,k-l}\right)\right. \\
&\quad \left.-\mu_{1,2}(0)\Big)(X_{0,1}X_{0,2} - \mu_{1,2}(0))\right] \\
&= E\big[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2})(X_{0,1}X_{0,2} - \mu_{1,2}(0))\big] \\
&\quad + E[(\alpha_1^k \circ X_{0,1})\left(\sum_{l=0}^{k-1} \alpha_2^l \circ \varepsilon_{2,k-l}\right)(X_{0,1}X_{0,2} - \mu_{1,2}(0))\Big] \\
&\quad + E\left[\left(\alpha_2^k \circ X_{0,2}\right)\left(\sum_{j=0}^{k-1} \alpha_1^j \circ E[\varepsilon_{1,k-j}]\right)(X_{0,1}X_{0,2} - \mu_{1,2}(0))\right] \\
&= E\big[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2})X_{0,1}X_{0,2}\big] - \mu_{1,2}(0)E\big[(\alpha_1^k \circ X_{0,1})(\alpha_2^k \circ X_{0,2})\big] \\
&\quad + \left(\sum_{l=0}^{k-1} \alpha_2^l \cdot E[\varepsilon_{2,k-l}]\right)\big(E[(\alpha_1^k \circ X_{0,1})X_{0,1}X_{0,2}] - \mu_{1,2}(0)E[\alpha_1^k \circ X_{0,1}]\big) \\
&\quad + \left(\sum_{j=0}^{k-1} \alpha_1^j \cdot E[\varepsilon_{1,k-j}]\right)\big(E[(\alpha_2^k \circ X_{0,2})X_{0,1}X_{0,2}] - \mu_{1,2}(0)E[\alpha_2^k \circ X_{0,2}]\big) \\
&= \alpha_1^k\alpha_2^k E\Big[X_{0,1}^2X_{0,2}^2\Big] - \mu_{1,2}(0)\alpha_1^k\alpha_2^k E[X_{0,1}X_{0,2}] \\
&\quad + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\Big(\alpha_1^k E[X_{0,2}X_{0,1}^2] - \mu_{1,2}(0)\alpha_1^k E[X_{0,1}]\Big) \\
&\quad + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\Big(\alpha_2^k E[X_{0,1}X_{0,2}^2] - \mu_{1,2}(0)\alpha_2^k E[X_{0,2}]\Big) \\
&= \alpha_1^k\alpha_2^k\Big(\mu_{1^2,2^2}(0) - \mu_{1,2}^2(0)\Big) + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\big(\alpha_1^k\mu_{1^2,2}(0) - \mu_{1,2}(0)\alpha_1^k\mu_1\big) \\
&\quad + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}(\alpha_2^k\mu_{1,2^2}(0) - \mu_{1,2}(0)\alpha_2^k\mu_2) \\
&= \alpha_1^k\alpha_2^k\Big(\mu_{1^2,2^2}(0) - \mu_{1,2}^2(0)\Big) + \mu_{\varepsilon_2}\tfrac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k(\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0)) \\
&\quad + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0)).
\end{aligned}$$

Thus, the explicit expression of σ_{55} is

$$\begin{aligned}
\sigma_{55} &= \mu_{1^2,2^2}(0) - \mu_{1,2}^2(0) + 2\sum_{k=1}^{\infty}\Big\{\alpha_1^k\alpha_2^k\Big(\mu_{1^2,2^2}(0) - \mu_{1,2}^2(0)\Big) \\
&\quad + \mu_{\varepsilon_2}\frac{1-\alpha_2^k}{1-\alpha_2}\alpha_1^k(\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0)) + \mu_{\varepsilon_1}\frac{1-\alpha_1^k}{1-\alpha_1}\alpha_2^k(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0))\Big\} \\
&= \mu_{1^2,2^2}(0) - \mu_{1,2}^2(0) + 2\Big\{(\mu_{1^2,2^2}(0) - \mu_{1,2}^2(0))\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} \\
&\quad + \tfrac{\mu_{\varepsilon_2}}{1-\alpha_2}\Big(\tfrac{\alpha_1}{1-\alpha_1} - \frac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\Big)(\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0)) \\
&\quad + \frac{\mu_{\varepsilon_1}}{1-\alpha_1}\Big(\tfrac{\alpha_2}{1-\alpha_2} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\Big)(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0))\Big\} \\
&= \mu_{1^2,2^2}(0) - \mu_{1,2}^2(0) + 2\Big\{(\mu_{1^2,2^2}(0) - \mu_{1,2}^2(0))\tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2} \\
&\quad + \mu_2\Big(\tfrac{\alpha_1}{1-\alpha_1} - \tfrac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\Big)(\mu_{1^2,2}(0) - \mu_1\mu_{1,2}(0)) \\
&\quad + \mu_1\Big(\tfrac{\alpha_2}{1-\alpha_2} - \frac{\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\Big)(\mu_{1,2^2}(0) - \mu_2\mu_{1,2}(0))\Big\}.
\end{aligned}$$

A.9 Proof of (17)–(21)

First, we denote

$$g_{\mathrm{LK}}(\mathbf{y}) = \frac{y_2\cdot\left(y_3 - y_1^2\right) - 2\cdot\left(y_5 - y_1y_2\right)^2 + y_1\left(y_4 - y_2^2\right)}{y_1y_2 - \left(y_5 - y_1y_2\right)^2} =: \frac{A}{B}.$$

According to the quotient rule, we get

$$\frac{\partial g_{\mathrm{LK}}}{y_1} = \frac{\frac{\partial A}{\partial y_1}\cdot B - \frac{\partial B}{\partial y_1}\cdot A}{B^2},$$

where

$$\begin{aligned}
\frac{\partial A}{\partial y_1} &= \frac{\partial\left[y_2\cdot\left(y_3 - y_1^2\right) - 2\cdot\left(y_5 - y_1y_2\right)^2 + y_1\left(y_4 - y_2^2\right)\right]}{\partial y_1} \\
&= -2y_1y_2 + 4(y_5 - y_1y_2)\cdot y_2 + \left(y_4 - y_2^2\right), \\
\frac{\partial B}{\partial y_1} &= \frac{\partial\left[y_1y_2 - \left(y_5 - y_1y_2\right)^2\right]}{\partial y_1} \\
&= y_2 + 2\left(y_5 - y_1y_2\right)y_2.
\end{aligned}$$

Thus, the explicit expression of $\frac{\partial g_{\mathrm{LK}}}{y_1}$ is

$$\frac{1}{\left[y_1y_2-(y_5-y_1y_2)^2\right]^2}\Big\{[-2y_1y_2+4(y_5-y_1y_2)\cdot y_2+(y_4-y_2^2)]\cdot\left[y_1y_2-(y_5-y_1y_2)^2\right]$$
$$-\left[y_2+2(y_5-y_1y_2)y_2\right]\cdot\left[y_2\cdot(y_3-y_1^2)-2\cdot(y_5-y_1y_2)^2+y_1(y_4-y_2^2)\right]\Big\}.$$

Inserting $\boldsymbol{\mu}=\left(\mu_1,\mu_2,\mu_{1^2}(0),\mu_{2^2}(0),\mu_{1,2}(0)\right)$ and after simplification, we get the explicit expression for $g'_{\mathrm{LK},1}$ given in (17). Similarly, we derive the Eq. (18).

Next, we focus on Eq. (19). Since

$$\frac{\partial A}{\partial y_3}=\frac{\partial\left[y_2\cdot(y_3-y_1^2)-2\cdot(y_5-y_1y_2)^2+y_1(y_4-y_2^2)\right]}{\partial y_3}=y_2,$$
$$\frac{\partial B}{\partial y_3}=\frac{\partial\left[y_1y_2-(y_5-y_1y_2)^2\right]}{\partial y_3}=0,$$

the derivative $\frac{\partial g_{\mathrm{LK}}}{y_3}$ can be written as

$$\frac{\partial g_{\mathrm{LK}}}{y_3}=\frac{\frac{\partial A}{\partial y_3}\cdot B-\frac{\partial B}{\partial y_3}\cdot A}{B^2}=\frac{y_2}{y_1y_2-(y_5-y_1y_2)^2}.$$

Again inserting $\boldsymbol{\mu}=\left(\mu_1,\mu_2,\mu_{1^2}(0),\mu_{2^2}(0),\mu_{1,2}(0)\right)$, we get (19). The derivation of Eq. (20) is analogous to that of $g'_{\mathrm{LK},3}$.

Finally, we have

$$\frac{\partial B}{\partial y_5}=\frac{\partial\left[y_1y_2-(y_5-y_1y_2)^2\right]}{\partial y_5}=-2(y_5-y_1y_2),$$
$$\frac{\partial A}{\partial y_5}=\frac{\partial\left[y_2\cdot(y_3-y_1^2)-2\cdot(y_5-y_1y_2)^2+y_1(y_4-y_2^2)\right]}{\partial y_5}=-4(y_5-y_1y_2)=2\frac{\partial B}{\partial y_5}.$$

So

$$\frac{\partial g_{\mathrm{LK}}}{y_5}=\frac{\frac{\partial A}{\partial y_5}\cdot B-\frac{\partial B}{\partial y_5}\cdot A}{B^2}=-2(y_5-y_1y_2)\,\frac{2\cdot B-A}{B^2}$$
$$=\frac{2(y_5-y_1y_2)}{\left[y_1y_2-(y_5-y_1y_2)^2\right]^2}\Big\{-2\left[y_1y_2-(y_5-y_1y_2)^2\right]$$
$$+\left[y_2\cdot(y_3-y_1^2)-2\cdot(y_5-y_1y_2)^2+y_1(y_4-y_2^2)\right]\Big\}$$
$$=\frac{2(y_5-y_1y_2)}{\left[y_1y_2-(y_5-y_1y_2)^2\right]^2}\Big\{-2y_1y_2+\left[y_2\cdot(y_3-y_1^2)+y_1(y_4-y_2^2)\right]\Big\}.$$

Inserting $\boldsymbol{\mu} = \big(\mu_1, \mu_2, \mu_{1^2}(0), \mu_{2^2}(0), \mu_{1,2}(0)\big)$, we get the explicit expression of $g'_{\mathrm{LK},5}$ in (21).

Acknowledgements The authors thank the two referees for their useful comments on an earlier draft of this article.

Author contributions Huiqiao Wang and Christian H.Weiß contributed to the theoretical analysis and simulation study. Mingming Zhang performed the validation. The first draft of the manuscript was written by Huiqiao Wang and Christian H.Weiß. Mingming Zhang has done the final check. All authors read and approved the final manuscript.

Funding Information This research was supported by The Central Public-interest Scientific Institution Basal Research Fund (No. 1610012024004).

Data availability Data and materials that support the findings of this study are available in the Supplementary Materials.

Declaration

Conflict of interest The authors declared no potential conflict of interest with respect to the publication of this article.

References

Aleksandrov, B., Weiß, C.H.: Parameter estimation and diagnostic tests for INMA(1) processes. TEST **29**(1), 196–232 (2020)

Best, D.J., Rayner, J.C.W.: Crockett's test of fit for the bivariate Poisson. Biom J **39**(4), 423–430 (1997)

Crockett, N.G.: A quick test of fit of a bivariate distribution. In: McNeil, D. (ed.) Interactive Statistics, pp. 185–191. North-Holland, Amsterdam (1979)

Fokianos, K.: Multivariate count time series modelling. Econom Stat **407**, 424 (2021)

Ibragimov, I.: Some limit theorems for stationary processes. Theory Probab Appl **7**(4), 349–382 (1962)

Johnson, N.L., Kotz, S., Balakrishnan, N.: Discrete Multivariate Distributions. John Wiley & Sons, New Jersey (1997)

Karlis, D.: Models for multivariate count time series. In: Davis, R.A., et al. (eds.) Handbook of Discrete-Valued Time Series, pp. 407–424. Chapman & Hall/CRC Press, Boca Raton (2016)

Kocherlakota, S., Kocherlakota, K.: Bivariate Discrete Distributions. Chapman & Hall/CRC Press, New York (2014)

Lee, Y., Lee, S., Tjøstheim, D.: Asymptotic normality and parameter change test for bivariate Poisson INGARCH models. TEST **27**(1), 52–96 (2018)

Loukas, S., Kemp, C.D.: The index of dispersion test for the bivariate Poisson distribution. Biometrics **42**(4), 941–948 (1986)

Novoa-Muñoz, F.: Goodness-of-fit tests for the bivariate Poisson distribution. Commun Stat Simul Comput **50**(7), 1998–2014 (2021)

Pedeli, X., Karlis, D.: A bivariate INAR(1) process with application. Stat Modell **11**(4), 325–349 (2011)

Quoreshi, A.M.M.S., Uddin, R., Mamode Khan, N.A.: A review of INMA integer-valued model class, application and further development. Filomat **34**(1), 143–152 (2020)
Rayner, J.C.W., Best, D.J.: Smooth tests for the bivariate Poisson distribution. Australian J Stat **37**(2), 233–245 (1995)
Schweer, S., Weiß, C.H.: Compound Poisson INAR(1) processes: stochastic properties and testing for overdispersion. Comput Stat Data Anal **77**, 267–284 (2014)
Serfling, R.J.: Approximation Theorems of Mathematical Statistics. John Wiley & Sons, New York (1980)
Wang, H., Weiß, C.H., Zhang, M.: The circumstance-driven bivariate integer-valued autoregressive model. Entropy **26**(2), 168 (2024)
Weiß, C.H.: An Introduction to Discrete-Valued Time Series. Wiley, Chichester (2018)
Weiß, C.H.: Stationary count time series models. WIREs Comput Stat **13**(1), 1502 (2021)
Weiß, C.H., Aleksandrov, B.: Model diagnostics for Poisson INARMA processes using bivariate dispersion indexes. J Stat Theory Pract **13**(2), 26 (2019)
Weiß, C.H., Schweer, S.: Bias corrections for moment estimators in Poisson INAR(1) and INARCH(1) processes. Stat Probab Letters **112**, 124–130 (2016)