On the Derivation of the Progress Rate and Self-Adaptation Response for the \((\mu_1, \lambda)\)–\(\sigma\)SA-ES on the Noisy Ellipsoid Model
On the Derivation of the Progress Rate and Self-Adaptation Response for the \((\mu/\mu_I, \lambda)-\sigma\text{SA-ES}\) on the Noisy Ellipsoid Model

\textbf{Technical Report TR 2015/01}

Alexander Melkozerov† and Hans-Georg Beyer‡

March, 30, 2015

1 Noisy Progress Rate

The goal of this section is to determine the \((\mu/\mu_I, \lambda)\)-ES progress rate along the \(i\)th axis of the noisy ellipsoid model

\[
F_{\text{noisy}}(y) = F(y) + \sigma \epsilon(y) N(0,1), \quad F(y) = \sum_{i=1}^{N} a_i y_i^2,
\]

where \(N(0,1)\) is a standard normally distributed random variate. Note that in order to simplify the analysis, the noise strength \(\sigma\) is taken at the parental state \(y\) and not at the offspring state \(y+\mathbf{x}\). This assumption is exact in the case of constant noise variance. It also holds asymptotically for \(N \to \infty\) since the optimal mutation strength goes to zero.

Following the approach introduced in [5], the ES progress along each axis is considered separately. That is, the noisy progress rate along the \(i\)th axis of the ellipsoid model (1) is defined as [5]

\[
\varphi_i = \frac{1}{\mu} \sum_{m=1}^{\mu} \mathbb{E} \left[ y_i^{(g)} - \left( \hat{y}_i^{(g)} \right)_{m;\lambda} | y^{(g)} \right],
\]

where the subscript \(m;\lambda\) refers to the observed \(m\)th-best of the \(\lambda\) offspring. To simplify notation, indices \((g)\) are omitted in the following derivations. Introducing the mutation vector

\[
\mathbf{x}_i = \delta_i \hat{\mathbf{z}}_i
\]

(note that \(\sigma\) is fixed and \(\mathbf{x}_i = \sigma \hat{\mathbf{z}}_i\) in the progress rate analysis due to the assumption \(\tau \xrightarrow{N \to \infty} 0\)) yields [5]

\[
\varphi_i = -\frac{1}{\mu} \sum_{m=1}^{\mu} \int_{-\infty}^{\infty} x p_{m;\lambda} (x|y) \, dx.
\]
The density of induced order statistics \( p_{m;\lambda}(x|y) \) in (4) has been obtained in [5], it reads in the noisy case

\[
p_{m;\lambda}(x|y) = \frac{\lambda!}{(m-1)! (\lambda-m)!} p_x(x) \int_{-\infty}^{+\infty} p_Q(q|x,y) P_Q(q|y)^{m-1} \left[ 1 - P_Q(q|y) \right]^{\lambda-m} dq, \tag{5}
\]

where \( p_Q(q|x,y) \) is the conditional density and \( P_Q(q|y) \) is the cumulative distribution function of the noisy local quality change

\[
Q_{\text{noisy}}(x,y) = F_{\text{noisy}}(y+x) - F(y), \tag{6}
\]

see Eqs. (9) and (10) below. In the next section, \( P_Q(q|y) \) is calculated for the noisy ellipsoid model.

1.1 Calculation of \( P_Q(q|y) \)

To determine \( P_Q(q|y) \), the normal approximation is used

\[
P_Q(q|y) \simeq \Phi \left( \frac{q - E [Q_{\text{noisy}}(x,y)]}{D [Q_{\text{noisy}}(x,y)]} \right), \tag{7}
\]

where \( E [Q_{\text{noisy}}(x,y)] \) is the expectation and \( D [Q_{\text{noisy}}(x,y)] \) is the standard deviation of the noisy local quality change [3]

\[
Q_{\text{noisy}}(x,y) := Q_y(x) + \sigma \epsilon \mathcal{N}(0,1) \tag{8}
\]

with \( x \sim \mathcal{N}(0,\sigma^2 I) \) being the mutation vector.

One obtains by inserting the expansion [5]

\[
Q_y(x) = \sum_{j=1}^{N} a_j \left( y_j + x_j \right)^2 - \sum_{j=1}^{N} a_j y_j^2
= \sum_{j=1}^{N} a_j \left( 2y_j x_j + x_j^2 \right), \tag{9}
\]

into Eq. (8)

\[
Q_{\text{noisy}}(x,y) = \sum_{j=1}^{N} a_j \left( 2y_j x_j + x_j^2 \right) + \sigma \epsilon \mathcal{N}(0,1), \tag{10}
\]

where \( x_j \) are components of the mutation vector \( x \). The expectation of \( Q_{\text{noisy}}(x,y) \) is equal to the expectation of the noise free \( Q_y(x) \)

\[
E [Q_{\text{noisy}}(x,y)] = \sigma^2 A_0, \tag{11}
\]

where

\[
A_n := \sum_{j \neq n}^{N} a_j, \tag{12}
\]

To determine \( D [Q_{\text{noisy}}(x,y)] \), \( Q_{\text{noisy}}(x,y) \) is written down as a sum of noise-free components and the noisy term

\[
Q_{\text{noisy}}(x,y) = \sum_{j=1}^{N} (Q_y)_j + \sigma \epsilon \mathcal{N}(0,1), \tag{13}
\]

where each component \((Q_y)_j\) is defined as [5]

\[
(Q_y)_j := a_j \left( 2y_j x_j + x_j^2 \right). \tag{13}
\]
\[ D [Q_{\text{noisy}} (x, y)] \] is calculated then by means of the variances

\[
D^2 [Q_{\text{noisy}} (x, y)] = \sum_{j=1}^{N} D^2 [(Q_y)_j] + D^2 [\sigma, \mathcal{N} (0, 1)]
\]

\[
= \sum_{j=1}^{N} D^2 [(Q_y)_j] + \sigma^2, \quad (14)
\]

where \( D^2 [(Q_y)_j] \) has been obtained in [5]

\[
D^2 [(Q_y)_j] = 2a^2_j \sigma^2 (2y_j^2 + \sigma^2). \quad (15)
\]

Therefore, the standard deviation of \( Q_{\text{noisy}} (x, y) \) is

\[
D [Q_{\text{noisy}} (x, y)] = \sqrt{D^2 [Q_{\text{noisy}} (x, y)]}
\]

\[
= \sqrt{\sum_{j=1}^{N} D^2 [(Q_y)_j] + D^2 [\sigma, \mathcal{N} (0, 1)]}
\]

\[
= \sqrt{\sum_{j=1}^{N} 2a^2_j \sigma^2 (2y_j^2 + \sigma^2) + \sigma^2}
\]

\[
= \sigma \sqrt{B_0 + \sigma^2 / \sigma^2}, \quad (16)
\]

where

\[
B_n := \sum_{j \neq n}^{N} 2a^2_j (2y_j^2 + \sigma^2). \quad (17)
\]

After inserting (11) and (16) into (7) the conditional probability distribution reads

\[
P_Q (q|y) \simeq \Phi \left( \frac{q - \sigma^2 A_0}{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2}} \right). \quad (18)
\]

### 1.2 Calculation of \( p_Q (q|x, y) \)

The conditional density in Eq. (5) is calculated analogously to the noise free case [5]

\[
p_Q (q|x, y) \simeq \frac{1}{\sqrt{2\pi D [Q_{\text{noisy}} (x, y)]}} \exp \left[ -\frac{1}{2} \left( \frac{q - E [Q_{\text{noisy}} (x, y)]}{D [Q_{\text{noisy}} (x, y)]} \right)^2 \right], \quad (19)
\]

except that the noisy local quality change \( Q_{\text{noisy}} \) is used. First, the \( i \)th summand of (10) is taken out and the substitution \( x_i = \sigma z_i \) is used

\[
Q_{\text{noisy}} (x, y) = a_i (2y_i \sigma z_i + \sigma^2 z_i^2) + \sum_{j \neq i}^{N} a_j (2y_j x_j + x_j^2) + \sigma, \mathcal{N} (0, 1). \quad (20)
\]

Next, under the assumption that \( |\sigma z_i| \ll |2y_i| \) does hold with sufficiently high probability that is ensured for sufficiently small \( \sigma \to 0 \), a \( Q_{\text{noisy}} (x, y) \) approximation is introduced

\[
Q_{\text{noisy}} (x, y) \simeq 2a_i y_i x_i + \sum_{j \neq i}^{N} a_j (2y_j x_j + x_j^2) + \sigma, \mathcal{N} (0, 1). \quad (21)
\]
The validity of (21) is verified experimentally. Keeping \( x_i = x \) fixed (since this is the condition, \( D^2 [x] = 0 \)) yields
\[
E [Q_{\text{noisy}} (x, y) | x] = 2a_i y_i x + \sigma^2 A_i.
\] (22)

and
\[
D^2 [Q_{\text{noisy}} (x, y) | x] = \sigma^2 \sum_{j \neq i} 2a_j^2 (2y_j^2 + \sigma^2) + \sigma^2,
\] (23)

which leads with (17) to \( D [Q_{\text{noisy}} (x, y) | x] = \sigma \sqrt{B_i + \sigma^2 / \sigma^2} \). Inserting (22) and \( D [Q_{\text{noisy}} (x, y) | x] \) into (19) results in
\[
p_Q (q | x, y) \approx \frac{1}{2 \pi \sigma \sqrt{B_i + \sigma^2 / \sigma^2}} \exp \left[ -\frac{1}{2} \left( \frac{q - 2a_i y_i x - \sigma^2 A_i}{\sigma \sqrt{B_i + \sigma^2 / \sigma^2}} \right)^2 \right].
\] (24)

Eq. (24) is used in the next section to calculate the integral (4).

### 1.3 Calculation of the Integral

Inserting (5) into (4), changing the order of integration and denoting the inner integral by
\[
I_i (q | y) := \int_{-\infty}^{+\infty} x p_x (x) p_Q (q | x, y) \, dx,
\] (25)

the progress rate formula (4) reads
\[
\varphi_i = -\frac{1}{\mu} \sum_{m=1}^{\mu} \frac{\lambda!}{(m-1)! (\lambda-m)!} \int_{-\infty}^{+\infty} I_i (q | y) P_Q (q | y)^{m-1} \left[ 1 - P_Q (q | y) \right]^{\lambda-m} \, dq.
\] (26)

Inserting (18) and (24) into (25) and substituting \( t = x / \sigma \) yields
\[
I_i (q | y) \approx \frac{1}{2 \pi \sqrt{B_i + \sigma^2 / \sigma^2}} \int_{-\infty}^{+\infty} te^{-\frac{1}{2} t^2} \exp \left[ -\frac{1}{2} \left( \frac{q - 2y_i a_i \sigma t - \sigma^2 A_i}{\sigma \sqrt{B_i + \sigma^2 / \sigma^2}} \right)^2 \right] \, dt.
\] (27)

Further the integral formula [3]
\[
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} te^{-\frac{1}{2} t^2} e^{-\frac{1}{2} (at+b)^2} \, dt = \frac{-ab}{(\sqrt{1+a^2})^3} e^{-\frac{1}{2} \left( \frac{a^2}{1+a^2} \right)}
\] (28)

is applied to (27). Provided that for (17)
\[
B_i = 4 \sum_{j \neq n} a_j^2 y_j^2 + 2 \sigma^2 \sum_{j \neq n} a_j^2 \gg 4a_i^2 y_i^2
\] (29)

holds (this basically means that there is not a dominating component in the objective function), \( I_i \) can be further simplified yielding
\[
I_i (q | y) \approx \frac{2a_i y_i}{\sqrt{2\pi (B_0 + \sigma^2 / \sigma^2)}} \frac{q - \sigma^2 A_0}{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2}} e^{-\frac{1}{2} \left( \frac{q - \sigma^2 A_0}{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2}} \right)^2}.
\] (30)

Inserting Eqs. (30) and (18) into Eq. (26), setting \( s = \frac{q - \sigma^2 A_0}{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2}} \) and rearranging the resulting equation leads to
\[
\varphi_i \approx -\frac{\lambda}{\mu} \frac{2a_i y_i}{\sqrt{2\pi \sqrt{B_0 + \sigma^2 / \sigma^2}}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2} t^2} \sum_{m=1}^{\mu} \frac{(\lambda - 1)!}{(m-1)! (\lambda-m)!} \left[ \Phi (s) \right]^{m-1} \left[ 1 - \Phi (s) \right]^{\lambda-m} \, ds.
\] (31)
The sum in Eq. (31) can be expressed by the integral [3]

\[ \sum_{m=1}^{\mu} \frac{(\lambda - 1)!}{(m - 1)! (\lambda - m)!} [\Phi(s)]^{m-1} (1 - \Phi(s))^\lambda - m \]

\[ = \frac{(\lambda - 1)!}{(\lambda - \mu - 1)! (\mu - 1)!} \int_0^{1-\Phi(s)} v^{\lambda-1} (1-v)^{\mu-1} \, dv. \quad (32) \]

Inserting (32) into Eq. (31), noting that \( \frac{(\lambda - 1)!}{(m - 1)! (\lambda - m)!} = (\lambda - \mu) \left( \frac{\lambda}{\mu} \right) \), using the substitution \( v = 1 - \Phi(t) \) and exchanging the order of integration leads to

\[ \varphi_1 \simeq -\frac{\lambda - \mu}{\sqrt{2\pi}} \left( \frac{\lambda}{\mu} \right) 2\sigma y_i a_i \int_{-\infty}^{t=\infty} (1 - \Phi(t))^{\lambda - \mu - 1} (\Phi(t))^{\mu - 1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \int_{s=-\infty}^{s=t} se^{-\frac{1}{2}s^2} \, ds \, dt. \quad (33) \]

Since \( \int_{s=-\infty}^{s=t} se^{-\frac{1}{2}s^2} \, ds = -e^{-\frac{1}{2}t^2} \), Eq. (33) simplifies to

\[ \varphi_1 \simeq -\frac{2\sigma y_i a_i}{\sqrt{B_0 + \sigma^2/\sigma^2}} \int_{-\infty}^{t=\infty} e^{-t^2} (1 - \Phi(t))^{\lambda - \mu - 1} (\Phi(t))^{\mu - 1} \, dt. \quad (34) \]

Comparing the resulting integral with

\[ e_{a,b}^{\mu,\lambda} = \frac{\lambda - \mu}{\sqrt{2\pi}^{\lambda+1}} \left( \frac{\lambda}{\mu} \right) \int_{-\infty}^{+\infty} (-t)^b e^{-\frac{1}{2}t^2} (1 - \Phi(t))^{\lambda - \mu - 1} (\Phi(t))^{\mu - a} \, dt, \quad (35) \]

the coefficient \( e_{\mu/\mu,\lambda} = e_{1,0}^{1,0} \) is recognized leading to the final formula

\[ \varphi_1 \simeq -\frac{2\sigma e_{\mu/\mu,\lambda} y_i a_i}{\sqrt{\sigma^2/\sigma^2 + \sum_{j=1}^{N} 2a_j^2 (2y_j^2 + \sigma^2)}}. \quad (36) \]

2 Noisy Quadratic Progress Rate \( \varphi^{II} \)

The quadratic progress rate \( \varphi^{II}_j \) is defined as

\[ \varphi^{II}_j = \left[ \left( y_j^{(g)} \right)^2 - \left( y_j^{(g+1)} \right)^2 \right] |y^{(g)}|. \quad (37) \]

As has been shown in [4], this can be expressed by the product moments \( E_1 \) and \( E_2 \) yielding

\[ \varphi^{II}_j = 2y_j \varphi_j - \frac{2}{\mu^2} E_1 - \frac{1}{\mu^2} E_2. \quad (38) \]

In the following, the coordinate index \( j \) has been dropped for the sake of brevity (already done for \( E_1 \) and \( E_2 \) above).

The product moments \( E_1 \) and \( E_2 \) are calculated in this section, which are defined as [4]

\[ E_1 = \sigma^2 E \sum_{l=2}^{\mu} \sum_{k=1}^{l-1} z_k z_l |y|, \quad (39) \]

\[ E_2 = \sigma^2 E \sum_{m=1}^{\mu} z_m^2 |y|. \quad (40) \]
The \( z_{k;\lambda} \) noisy order statistics refer to the \( j \)th components of the mutation vector \( x_{k;\lambda} \) producing the \( k \)th best offspring

\[
\tilde{y}_{k;\lambda} = y + x_{k;\lambda} = y + \tilde{\sigma}_{k;\lambda} z_{k;\lambda}.
\]

The \( k \)th best offspring is ranked according to its objective function value \( \tilde{F}(\tilde{y}_{k;\lambda}) \) which depends on the random vector \( z_{k;\lambda} = \mathcal{N}(\mathbf{0}, \mathbf{I}) \) and the fitness noise. For the calculations it is assumed that \( \tilde{\sigma}_{k;\lambda} = \sigma \), i.e. \( y_{k;\lambda} = y + \sigma z_{k;\lambda} \). This assumption is asymptotically exact provided that for \( N \to \infty \) it holds \( \tau \to 0 \). If \( \tau \to 0 \), the results are of approximate nature and their validity must be checked by experiments.

To compute \( E_1 \) and \( E_2 \), the local quality change is considered first.

### 2.1 Local Quality Change

Using the noisy local quality function expansion leads to

\[
Q_{\text{noisy}}(\mathbf{x}, y) = \frac{\sum_{i=1}^{N} a_i \left( y_i + \sigma (z_i)_{k;\lambda} \right)^2 - \sum_{i=1}^{N} a_i y_i^2 + \sigma \epsilon \mathcal{N}(0, 1)}{2 \sigma a_j y_j} =: \nu_{k;\lambda}.
\]

Dividing both sides by \( 2 \sigma a_j y_j \) and introducing the quotient

\[
\frac{Q_{\text{noisy}}(\mathbf{x}, y)}{2 \sigma a_j y_j} =: \nu_{k;\lambda}
\]

yields

\[
\nu_{k;\lambda} = (z_j)_{k;\lambda} + \frac{\sum_{i \neq j} a_i y_i}{a_j y_j} (z_i)_{k;\lambda} + \frac{\sigma}{2} \sum_{i=1}^{N} a_i (z_i)_{k;\lambda}^2 + \frac{\sigma \epsilon}{2 \sigma a_j y_j} \mathcal{N}(0, 1).
\]

Equation (44) is a sum of the random variate \( (z_j)_{k;\lambda} \), two sum expressions and a noise term. For \( N \to \infty \), the central limit theorem can be applied to the second and third term in (44) yielding an approximate normal distribution [4]

\[
\mathcal{N} \left( \frac{\sigma \sum_{i=1}^{N} a_i}{2 \sigma y_j}, \frac{1}{\sigma y_j^2} \left( \sum_{i \neq j} a_i^2 y_i^2 + \frac{\sigma^2}{2} \sum_{i=1}^{N} a_i^2 \right) \right).
\]

Taking into account (45) and the variance of the noisy term, Eq. (44) transforms into

\[
\nu_{k;\lambda} = (z_j)_{k;\lambda} + \mathcal{N} \left( \frac{\sigma \sum_{i=1}^{N} a_i}{2 \sigma y_j}, \frac{1}{\sigma y_j^2} \left( \sum_{i \neq j} a_i^2 y_i^2 + \frac{\sigma^2}{2} \sum_{i=1}^{N} a_i^2 + \frac{\sigma^2 \epsilon}{4 \sigma^2} \right) \right).
\]

The \( k \)th random variate \( (z_j)_{k;\lambda} \) in Eq. (46) corresponds to the \( k \)th best \( Q_{\text{noisy}}(\mathbf{x}, y) \) value which is proportional to \( \nu_{k;\lambda} \) (cf. Eq. (43)). Considering the second term in Eq. (46) as a noise term, the variates \( (z_j)_{k;\lambda} \) can be identified as noisy order statistics or concomitants of \( \nu_{k;\lambda} \).
2.2 Correlation Coefficient

The sums of product moments of \((z_j)_{k,\lambda}\) have been calculated in [4]

\[ E_1 = \mu(\mu - 1) \frac{\sigma^2}{2} \rho^2 e_{\mu,\lambda}^2, \]  
\[ E_2 = \mu\sigma^2 \left(1 + \rho^2 e_{\mu,\lambda}^{1,1}\right), \]

where \(\rho\) is the correlation coefficient [2]

\[ \rho = \frac{1}{\sqrt{1 + \beta^2}}. \]

As follows from Eq. (46), the variance \(\beta^2\) in (49) is expressed as

\[ \beta^2 = \frac{1}{a_j y_j} \left( \sum_{i \neq j}^N \sigma_i^2 y_i^2 + \frac{\sigma_j^2}{2} \sum_{i=1}^N \sigma_i^2 + \frac{\sigma_j^2}{4\sigma^2} \right). \]

Thus the correlation coefficient \(\rho\) for the noisy case reads

\[ \rho = \frac{|a_j y_j|}{\sqrt{a_j y_j^2 + \sum_{i \neq j}^N \sigma_i^2 y_i^2 + \frac{\sigma_j^2}{2} \sum_{i=1}^N \sigma_i^2 + \frac{\sigma_j^2}{4\sigma^2}}}. \]

Inserting Eq. (51) into Eqs. (47) and (48), the final formulae for the expectations of product moments are obtained

\[ E_1 \simeq \mu(\mu - 1) \frac{\sigma^2}{2} \frac{a_j^2 y_j^2 e_{\mu,\lambda}^{2,0}}{\frac{\sigma_j^2}{4\sigma^2} + \sum_{i=1}^N a_i^2 \left(y_i^2 + \frac{\sigma_i^2}{2}\right)} \]

and

\[ E_2 \simeq \mu\sigma^2 \left(1 + \frac{a_j^2 y_j^2 e_{\mu,\lambda}^{1,1}}{\frac{\sigma_j^2}{4\sigma^2} + \sum_{i=1}^N a_i^2 \left(y_i^2 + \frac{\sigma_i^2}{2}\right)}\right). \]

3 The Noisy SAR Function

The derivation of the SAR formula for the \((\mu/\mu_1,\lambda,\sigma)-\sigma\text{SA-ES}\) on the noisy ellipsoid model follows the analysis steps for its noise-free counterpart described in [5]. The starting point is the integral representation of the SAR function [5]

\[ \psi(\sigma) = \frac{1}{\mu} \sum_{m=1}^{\infty} \int_0^{\infty} \left(\frac{\tilde{\sigma} - \sigma}{\sigma}\right) p_{\mu,\lambda}(\tilde{\sigma} | \sigma) \, d\tilde{\sigma}, \]
where $p_{m,\lambda}(\bar{\sigma} | \sigma)$ is the density of induced order statistics. The $p_{m,\lambda}(\bar{\sigma} | \sigma)$ formula reads [5]

$$p_{m,\lambda}(\bar{\sigma} | \sigma) = \frac{\lambda^m}{(m-1)! (\lambda - m)!} p_\sigma(\bar{\sigma} | \sigma) \int_{-\infty}^{\infty} p_Q(q | \bar{\sigma}) P_Q(q | \sigma)^{m-1} (1 - P_Q(q | \sigma))^\lambda - m \; dq.$$  \hspace{1cm} (55)

where $p_\sigma(\bar{\sigma} | \sigma)$ is the distribution density of the log-normal mutations [3]

$$p_\sigma(\bar{\sigma} | \sigma) = \frac{1}{\sqrt{2\pi \bar{\sigma}}} \exp \left[ - \frac{1}{2} \left( \frac{\ln (\bar{\sigma}/\sigma)}{\tau} \right)^2 \right].$$  \hspace{1cm} (56)

The conditional probability distribution $P_Q(q | \sigma)$ is calculated using the formula

$$P_Q(q | \sigma) = \int_0^\infty P_Q(q | \bar{\sigma}) p_\sigma(\bar{\sigma} | \sigma) d\bar{\sigma}.$$  \hspace{1cm} (57)

For sufficiently small $\tau$, the integral in (57) can be approximated by [3]

$$P_Q(q | \sigma) \simeq \Phi \left( \frac{q - E[\sigma]}{D[\sigma]} \right).$$  \hspace{1cm} (58)

where the expectation $E[\sigma] = \sigma^2 A_0$ and the standard deviation $D[\sigma] = \sigma \sqrt{B_0 + \sigma^2 / \tau^2}$ have been calculated in Section 1.1. Inserting these results into (58) yields

$$P_Q(q | \sigma) \simeq \Phi \left( \frac{q - \sigma^2 A_0}{\sigma \sqrt{B_0 + \sigma^2 / \tau^2}} \right).$$  \hspace{1cm} (59)

In order to keep (55) tractable, $p_Q(q | \bar{\sigma})$ is also approximated by a normal distribution

$$p_Q(q | \bar{\sigma}) \simeq \frac{1}{\sqrt{2\pi D[Q_{\text{noisy}}(x, y) | \bar{\sigma}]}} \exp \left[ - \frac{1}{2} \left( \frac{q - E[Q_{\text{noisy}}(x, y) | \bar{\sigma}]}{D[Q_{\text{noisy}}(x, y) | \bar{\sigma}]} \right)^2 \right].$$  \hspace{1cm} (60)

The expectation $E[Q_{\text{noisy}}(x, y) | \bar{\sigma}]$ and standard deviation $D[Q_{\text{noisy}}(x, y) | \bar{\sigma}]$ can be obtained similarly to results in Section 1.2. Taking into account these results, one obtains

$$p_Q(q | \bar{\sigma}) \simeq \frac{1}{\sqrt{2\pi \bar{\sigma}} \sqrt{\tilde{B}_0 + \sigma^2 / \bar{\sigma}^2}} \exp \left[ - \frac{1}{2} \left( \frac{q - \bar{\sigma}^2 A_0}{\bar{\sigma} \sqrt{\tilde{B}_0 + \sigma^2 / \bar{\sigma}^2}} \right)^2 \right],$$  \hspace{1cm} (61)

where

$$\tilde{B}_0 = \sum_{j \neq n}^{N} 2a_j^2 (2 \bar{y}_j^2 + \bar{\sigma}^2).$$  \hspace{1cm} (62)

### 3.1 Calculation of the Integral

Inserting Eqs. (61) and (59) into (55) and Eq. (55) into (54) yields

$$\psi \simeq \frac{1}{\mu} \sum_{m=1}^{\infty} \int_0^{\infty} \left( \frac{\bar{\sigma} - \sigma}{\sigma} \right)^m \frac{\lambda!}{(m-1)! (\lambda - m)!} p_\sigma(\bar{\sigma} | \sigma)$$

$$\times \frac{1}{\sqrt{2\pi \bar{\sigma}}} \sqrt{\tilde{B}_0 + \sigma^2 / \bar{\sigma}^2} \frac{\lambda!}{(m-1)! (\lambda - m)!} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \left( \frac{s \sqrt{\tilde{B}_0 + \sigma^2 / \bar{\sigma}^2} - A_0 (a^2 - \bar{\sigma}^2)^2}{s \sqrt{\tilde{B}_0 + \sigma^2 / \bar{\sigma}^2}} \right)^2} \Phi(s)^{m-1} (1 - \Phi(s))^\lambda - m \; ds d\bar{\sigma},$$  \hspace{1cm} (63)
where the substitution $s = (q - \sigma^2 A_0) / \left( \sigma \sqrt{B_0 + \sigma^2 / \sigma^2} \right)$ has been employed.

Using the integral representation (32) of the regularized incomplete beta-function [3], the sum in (63) can be substituted with that integral and the resulting equation is

$$
\psi \simeq (\lambda - \mu) \left( \frac{\lambda}{\mu} \right) \int_0^\infty \left( p \sigma (\sigma | \sigma) \frac{1}{\sqrt{2\pi}} \frac{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2}}{\tilde{\sigma}} \right) \\
\times \int_{-\infty}^{\infty} e^{-\frac{1}{2} \sigma^2 \left(1 - \Phi(p)\right)^{\lambda-\mu-1} \Phi(p)^{\mu-1}} \\
\times \frac{1}{\sqrt{2\pi}} \frac{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2}}{\tilde{\sigma} \sqrt{B_0 + \sigma^2 / \tilde{\sigma}^2}} \int_{-\infty}^{s=p} \left( e^{-\frac{1}{2} \sigma^2 (\sigma^2 / \sigma^2 - A_0(s^2 - \sigma^2))^2} \right) ds \sigma d\tilde{\sigma}, \tag{64}
$$

Changing the order of the integration in (64) and applying the substitution $p = \Phi^{-1}(1 - \nu)$ yields

$$
\psi \simeq \frac{\lambda - \mu}{\sqrt{2\pi}} \left( \frac{\lambda}{\mu} \right) \int_0^\infty \left( p \sigma (\sigma | \sigma) \right) \frac{1}{\sqrt{2\pi}} \frac{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2}}{\tilde{\sigma} \sqrt{B_0 + \sigma^2 / \tilde{\sigma}^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \sigma^2 \left(1 - \Phi(p)\right)^{\lambda-\mu-1} \Phi(p)^{\mu-1}} \\
\times \frac{1}{\sqrt{2\pi}} \frac{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2}}{\tilde{\sigma} \sqrt{B_0 + \sigma^2 / \tilde{\sigma}^2}} \int_{-\infty}^{s=p} \left( e^{-\frac{1}{2} \sigma^2 (\sigma^2 / \sigma^2 - A_0(s^2 - \sigma^2))^2} \right) ds \sigma d\tilde{\sigma}, \tag{65}
$$

where the limits of the integral over $p$ have been reversed. Note that the innermost integral over $s$ is basically the CDF of the normal distribution. Using simple variable substitution, one easily finds

$$
\frac{1}{\sqrt{2\pi}} \frac{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2}}{\tilde{\sigma} \sqrt{B_0 + \sigma^2 / \tilde{\sigma}^2}} \int_{-\infty}^{s=p} \exp \left[ -\frac{1}{2} \left( \frac{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2} s - A_0(\tilde{\sigma}^2 - \sigma^2)}{\tilde{\sigma} \sqrt{B_0 + \sigma^2 / \tilde{\sigma}^2}} \right)^2 \right] ds

= \Phi \left( \frac{\sigma \sqrt{B_0 + \sigma^2 / \sigma^2} p - A_0(\tilde{\sigma}^2 - \sigma^2)}{\tilde{\sigma} \sqrt{B_0 + \sigma^2 / \tilde{\sigma}^2}} \right)

= \Phi \left( f(\tilde{\sigma}) \right). \tag{66}
$$

Inserting this result into (65) still yields an intractable integral w.r.t. $p$. Recalling that, for sufficiently small $\tau$, $\tilde{\sigma}$ deviates only weakly from $\sigma$ with high probability, (66) can be expanded in a Taylor series about $\sigma$. Breaking off after the linear term in $(\tilde{\sigma} - \sigma)$ yields after a tedious calculation

$$
\Phi \left( f(\tilde{\sigma}) \right) = \Phi(p) - \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \sigma^2 \left( \frac{B_0}{B_0 + \sigma^2 / \sigma^2} p + \frac{2\sigma A_0}{\sqrt{B_0 + \sigma^2 / \sigma^2}} \left( \tilde{\sigma} - \sigma \right) / \sigma \right) + \mathcal{O} \left( (\tilde{\sigma} - \sigma)^2 \right) \right]. \tag{67}
$$
Inserting (67) into (65) yields

\[ \psi \simeq \frac{\lambda - \mu}{\sqrt{2\pi}} \left( \frac{\lambda}{\mu} \right)^{\frac{1}{2}} \int_{0}^{\infty} \left( \frac{\sigma - \sigma}{\sigma} \right) p_\sigma (\bar{\sigma} | \sigma) \left( e^{-\frac{1}{2}p^2} (1 - \Phi (p))^{\lambda - \mu - 1} \Phi (p)^{\mu - 1} \right) \]

\[ \times \left( \Phi (p) - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}p^2} B_0 \frac{p - B_0 + \sigma_z^2 / \sigma^2}{B_0 + \sigma_z^2 / \sigma^2} \left( \frac{\bar{\sigma} - \sigma}{\sigma} \right) \right) \right) dp d\bar{\sigma} \]

\[ = \frac{\lambda - \mu}{\sqrt{2\pi}} \left( \frac{\lambda}{\mu} \right)^{\frac{1}{2}} \int_{0}^{\infty} \left( \frac{\sigma - \sigma}{\sigma} \right) p_\sigma (\bar{\sigma} | \sigma) \right) \]

\[ \times \left[ \int_{p=-\infty}^{p=+\infty} e^{-\frac{1}{2}p^2} (1 - \Phi (p))^{\lambda - \mu - 1} \Phi (p)^{\mu - 1} \left( e^{-\frac{1}{2}p^2} B_0 \frac{p - B_0 + \sigma_z^2 / \sigma^2}{B_0 + \sigma_z^2 / \sigma^2} \left( \frac{\bar{\sigma} - \sigma}{\sigma} \right) \right) dp \]

\[ + \int_{p=-\infty}^{p=+\infty} e^{-\frac{1}{2}p^2} (1 - \Phi (p))^{\lambda - \mu - 1} \Phi (p)^{\mu - 1} \left( e^{-\frac{1}{2}p^2} B_0 \frac{p - B_0 + \sigma_z^2 / \sigma^2}{B_0 + \sigma_z^2 / \sigma^2} \left( \frac{\bar{\sigma} - \sigma}{\sigma} \right) \right) \right) dp \]

\[ = \frac{\lambda - \mu}{\sqrt{2\pi}} \left( \frac{\lambda}{\mu} \right)^{\frac{1}{2}} \int_{0}^{\infty} \left( \frac{\sigma - \sigma}{\sigma} \right) p_\sigma (\bar{\sigma} | \sigma) \right) \]

\[ \times \left[ \int_{p=-\infty}^{p=+\infty} e^{-\frac{1}{2}p^2} (1 - \Phi (p))^{\lambda - \mu - 1} \Phi (p)^{\mu - 1} \right) dp \int_{0}^{\infty} \left( \frac{\bar{\sigma} - \sigma}{\sigma} \right) p_\sigma (\bar{\sigma} | \sigma) d\bar{\sigma} \]

\[ + \frac{B_0}{B_0 + \sigma_z^2 / \sigma^2} \lambda - \mu \mu (\lambda) \int_{p=-\infty}^{p=+\infty} e^{-\frac{1}{2}p^2} (1 - \Phi (p))^{\lambda - \mu - 1} \Phi (p)^{\mu - 1} dp \]

\[ \times \int_{0}^{\infty} \left( \frac{\bar{\sigma} - \sigma}{\sigma} \right)^2 p_\sigma (\bar{\sigma} | \sigma) d\bar{\sigma} \]

\[ - \frac{2A_0 \sigma}{\sqrt{B_0 + \sigma_z^2 / \sigma^2}} \lambda - \mu \mu (\lambda) \int_{p=-\infty}^{p=+\infty} e^{-\frac{1}{2}p^2} (1 - \Phi (p))^{\lambda - \mu - 1} \Phi (p)^{\mu - 1} dp \]

\[ \times \int_{0}^{\infty} \left( \frac{\bar{\sigma} - \sigma}{\sigma} \right)^2 p_\sigma (\bar{\sigma} | \sigma) d\bar{\sigma}. \]
Using the resulting expressions, Eq. (69) transforms into
\[
\psi \simeq \left( \frac{\tau^2}{2} + O(\tau^4) \right) + e^{1,1}_{\mu,\lambda} \frac{B_0}{B_0 + \sigma^2/\sigma^2} (\tau^2 + O(\tau^4)) - c_{\mu/\mu,\lambda} \frac{2A_0\sigma}{\sqrt{B_0 + \sigma^2/\sigma^2}} (\tau^2 + O(\tau^4)) \tag{72}
\]
and the final SAR formula reads
\[
\psi (\sigma) \simeq \tau^2 \left( \frac{1}{2} + e^{1,1}_{\mu,\lambda} \frac{B_0}{B_0 + \sigma^2/\sigma^2} - c_{\mu/\mu,\lambda} \frac{2A_0}{\sqrt{B_0 + \sigma^2/\sigma^2}} \right) \tag{73}
\]
\[
= \tau^2 \left( \frac{1}{2} + e^{1,1}_{\mu,\lambda} \frac{1}{1 + \frac{\sigma^2}{\sigma^2}} - c_{\mu/\mu,\lambda} \frac{2 \sum_{i=1}^{N} a_i}{\sigma^2 + \sum_{i=1}^{N} 2a_i^2 (2y_i^2 + \sigma^2)} \right).
\]

Acknowledgments

This work was supported by the Austrian Science Fund FWF under grants P19069-N18 and P22649-N23. The work of A. Melkozerov was supported by the state contract 8.1802.2014/K of the Russian Ministry of Education and Science (software development), the RFBR grant 14-29-0925 (analytic derivations), and the RSF grant 14-19-01232 (numerical experiments) in the TUSUR University.

References


