

Bayesian statistics approach to chess engines optimization

Ivan Ivec, Ivana Vojnović

Abstract

We develop a new method for stochastic optimization using the Bayesian statistics approach. More precisely, we optimize parameters of chess engines as those data are available to us, but the method should apply to all situations where we want to optimize a certain gain/loss function which has no analytical form and thus cannot be measured directly but only by comparison of two parameter sets. We also experimentally compare the new method with the famous SPSA method.

Keywords: Bayesian statistics, stochastic optimization, chess engines

2010 Math. Subj. Class.: 62F15, 62L20, 65K10

1 Introduction

The simultaneous perturbation stochastic approximation (SPSA) method for multivariate optimization problems was introduced in [6], and it found numerous applications in engineering and the physical and social sciences. The method was more fully analyzed in [7], while in [8], it was given a simple step-by-step guide to the implementation of SPSA, with references to various applications.

More recently, the SPSA method was implemented in [10] as a tuner of parameters of the Stockfish chess engine. Stockfish is an open source project [11], and it is currently the strongest chess engine in the world

(winning the last five seasons of the famous TCEC competition [12] and being on the top of the CCRL rating lists [3]), much stronger than the engine used by the famous Deep Blue computer, which defeated a reigning chess world champion, Garry Kasparov, in 1997. Over the years, the SPSA method brought many successful patches to Stockfish and other chess engines.

However, tuning parameters of chess engines is computationally very demanding, and many tuning attempts are unsuccessful. In this paper, we try to use the Bayesian statistics approach to find a method that will be better, at least in some situations. We choose this approach because Bayesian inference is known for giving good posterior distributions from pretty bad prior distributions if the prior's support and shape are reasonably selected. For more details and references regarding Bayesian optimization, see [5].

In the next section, we provide some theoretical justifications for our method, which we call the BSPSA method because of the Bayesian approach and because of a similarity with the SPSA method. In fact, a difference from SPSA method will be in a way of updating parameters' approximations. We want to optimize parameters $\theta = (\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n)}) \in \mathbb{R}^n$, and because the game of chess is a stochastic process, that actually means that we are looking for a distribution of the optimal value of θ . In other words, we try to obtain θ that will work best in most situations, and the mean value of the distribution will be a candidate for that. SPSA updates parameters by formula

$$\theta_{k+1}^{(i)} = \theta_k^{(i)} + \frac{a_k^{(i)}}{\Delta_k^{(i)} c_k^{(i)}} \cdot w_k,$$

where $a_k^{(i)}$, $\Delta_k^{(i)}$ and $c_k^{(i)}$ are appropriate sequences and w_k is the result of an experiment (two-game match in our case) – see the beginning of the third section for more details. BSPSA will update parameters using the famous Bayes' theorem for the conditional probability, i.e. the formula (1) given below. We provide experimental results in the third section, comparing SPSA and BSPSA. Although we stick to tuning the Stockfish parameters, we believe the method could also be applied to other situations where a gain/loss function is not measured directly but only by comparing two parameter sets. We discuss such possibilities in the concluding section.

2 Theoretical development of a new method

For simplicity, we first look at a single parameter $\theta \in \mathbb{R}$ that we want to optimize by playing many chess games. The central limit theorem guarantees that the distribution of experimentally obtained optimal values using n games (for n large) will be approximately normal. Of course, it is always possible that we encounter an ill-behaved parameter whose optimal value will have an infinite mean or infinite variance and where the central limit theorem does not apply. Still, such parameters are not useful in the code of our engine, and we strongly try to avoid them. Therefore, we construct a sequence (θ_k) , hopefully converging to the optimal value θ_0 , by starting with $N(\theta_k, s_k^2)$ normal prior distribution and calculating $N(\theta_{k+1}, s_{k+1}^2)$ posterior distribution.

So, assuming that we have already obtained θ_k , we play two games between an engine E_1 that uses $\theta = \theta_k + c_k$ and an engine E_2 that uses $\theta = \theta_k - c_k$, where $c_k > 0$ will be chosen in the same way as in the SPSA method or in some other way, whichever proves to be more efficient. The result of this two-game match, from the perspective of the engine E_1 , we denote by w_k (with possible values 2, 1, 0, -1, -2, coming from one point for a win, zero points for a draw and -1 for a loss).

Now, if $f(w_k|\theta)$ is a conditional observation distribution of w_k , assuming that θ is given, then the Bayesian inference starting from the prior distribution with the density

$$\pi(\theta) \sim \exp\left(-\frac{1}{2}\left(\frac{\theta-\theta_k}{s_k}\right)^2\right)$$

gives the posterior distribution with the density

$$\pi(\theta|w_k) \sim f(w_k|\theta)\pi(\theta), \tag{1}$$

where by $a \sim b$ we denote that $\frac{a}{b}$ is a constant.

In what follows, we show that, under reasonable assumptions, we can approximate $f(w_k|\theta)$ with the density of the normal distribution with the mean of the form $A\theta + B$, for some constants $A > 0$, $B \in \mathbb{R}$. Although w_k can obtain only 5 values, approximation with normal distribution is viable because of the known fact from the Bayesian inference that it makes no difference whether we analyze observations one at a time in sequence using the posterior after the previous step as the prior for the next step, or whether we analyze all observations together in a single step starting with our initial prior [1, Sections 6.1, 11.1]. More precisely,

that is why we decide to update our belief about the parameter after every two games instead of playing more games with the same parameter. In that way, the posterior distribution $\pi(\theta|w_k)$ will also be normal (at least approximately), and we obtain explicit formulae for its calculation, thus avoiding the numerical integration needed to get a constant in (1).

By WR we denote a win rate, i.e. a ratio between wins/points obtained by the engine E_1 and wins/points obtained by the engine E_2 . Now we take an assumption that the strength of an engine (expressed as the number of wins/points obtained) is proportional to the density of the normal distribution with mean θ (as we assume that θ is optimal) and standard deviation σ (there is no clear justification for this assumption, but ideally we would like to construct an engine which parameters are independent and identically distributed with a finite variance, in which case the central limit theorem would be a kind of justification again), thus obtaining

$$\begin{aligned} WR &= \exp\left(-\frac{1}{2}\left(\frac{\theta_k + c_k - \theta}{\sigma}\right)^2 + \frac{1}{2}\left(\frac{\theta_k - c_k - \theta}{\sigma}\right)^2\right) \\ &= \exp\left(\frac{2c_k(\theta - \theta_k)}{\sigma^2}\right). \end{aligned}$$

The win percentage is then

$$WP = \frac{WR}{WR + 1},$$

and thus we get

$$\begin{aligned} w_k &\approx 2WP - 2(1 - WP) = 2 \cdot \frac{WR - 1}{WR + 1} \\ &= 2 \cdot \frac{\exp\left(\frac{2c_k(\theta - \theta_k)}{\sigma^2}\right) - 1}{\exp\left(\frac{2c_k(\theta - \theta_k)}{\sigma^2}\right) + 1}. \end{aligned} \tag{2}$$

Here we approximate w_k with its expectation ($2WP$ comes from two games with the win percentage WP and one point for each win, while $-2(1 - WP)$ comes from two games with the loss percentage $1 - WP$ and -1 point for each loss).

As the most difficult task in practice is to optimize a parameter that is already pretty close to the optimum, we can take a linear approximation of the right-hand side in (2) around $\theta = \theta_k$ to obtain

$$w_k \approx \frac{2c_k(\theta - \theta_k)}{\sigma^2}. \tag{3}$$

Finally, we use (3) to assume that $f(w_k|\theta)$ is approximately equal to the density of the normal distribution with mean $\frac{2c_k(\theta-\theta_k)}{\sigma^2}$ and standard deviation τ . We also assume that standard deviations σ and τ are constants not depending on θ because we are satisfied with high precision only around $\theta = \theta_k$.

For the brevity of calculations we denote $A = \frac{2c_k}{\sigma^2}$, $B = \frac{-2c_k\theta_k}{\sigma^2}$ and from (1) we obtain

$$\begin{aligned} \pi(\theta|w_k) &\sim \exp\left(-\frac{1}{2}\left(\frac{w_k - A\theta - B}{\tau}\right)^2\right) \cdot \exp\left(-\frac{1}{2}\left(\frac{\theta - \theta_k}{s_k}\right)^2\right) \\ &= \exp\left(-\frac{1}{2}\left(\frac{w_k^2 + A^2\theta^2 + B^2 - 2A\theta w_k - 2Bw_k + 2AB\theta}{\tau^2} + \frac{\theta^2 - 2\theta_k\theta + \theta_k^2}{s_k^2}\right)\right) \\ &= \exp\left(-\frac{1}{2} \cdot \frac{(A^2s_k^2 + \tau^2)\theta^2 - 2(Aw_k s_k^2 - ABs_k^2 + \theta_k\tau^2)\theta + \dots}{s_k^2\tau^2}\right) \\ &\sim \exp\left(-\frac{1}{2} \cdot \frac{A^2s_k^2 + \tau^2}{s_k^2\tau^2} \left(\theta - \frac{Aw_k s_k^2 - ABs_k^2 + \theta_k\tau^2}{A^2s_k^2 + \tau^2}\right)^2\right). \end{aligned}$$

We have thus obtained that the posterior distribution is also a normal distribution with the mean

$$\theta_{k+1} = \frac{Aw_k s_k^2 - ABs_k^2 + \theta_k\tau^2}{A^2s_k^2 + \tau^2}$$

and the variance

$$s_{k+1}^2 = \frac{s_k^2\tau^2}{A^2s_k^2 + \tau^2}.$$

After inserting $A = \frac{2c_k}{\sigma^2}$, $B = \frac{-2c_k\theta_k}{\sigma^2}$ we obtain final formulae which we use to update our belief about the parameter after each two-game match:

$$\theta_{k+1} = \theta_k + \frac{2c_k s_k^2 \sigma^2}{4c_k^2 s_k^2 + \tau^2 \sigma^4} \cdot w_k, \quad (4)$$

$$s_{k+1}^2 = \frac{s_k^2 \tau^2 \sigma^4}{4c_k^2 s_k^2 + \tau^2 \sigma^4}. \quad (5)$$

Remark 1. In the limit $\tau \rightarrow 0$ the formula (4) reads

$$\theta_{k+1} = \theta_k + \frac{\sigma^2}{2c_k} \cdot w_k,$$

which is exactly the formula used in the SPSA method, with the difference that SPSA uses $a_k = \frac{a}{(A+k)^\alpha}$ instead of $\frac{\sigma^2}{2}$. This means that uncertainty of the two-game match, contained in the parameter τ , is our replacement for the decreasing sequence a_k used in the SPSA method. Also, it is obvious from the formula (5) that $0 < s_{k+1} < s_k$, and thus $\lim_{k \rightarrow \infty} s_k$ exists. By passing to the limit in (5) we get $\lim_{k \rightarrow \infty} s_k = 0$ if $\lim_{k \rightarrow \infty} c_k \neq 0$, which is a good sign to expect the convergence of θ_k to the optimal value. At first sight, it seems that c_k used in the SPSA method (and so also in our method) does not satisfy this condition. However, we first choose the number of iterations N and then we can choose c_N to be big enough and we have $c_k \geq c_N > 0$ for all $k \in \{1, \dots, N\}$. In addition, this is another reason (besides those mentioned in the concluding section) for looking for a better choice of c_k in future work.

Remark 2. If we want to optimize more parameters $\theta = (\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n)})$, then assuming that we have already obtained θ_k , we play two games between an engine E_1 that uses $\theta = \theta_k + \Delta_k c_k$ and an engine E_2 that uses $\theta = \theta_k - \Delta_k c_k$, where each component of c_k is positive and each component of Δ_k is using a Bernoulli ± 1 distribution with probability $\frac{1}{2}$ for each ± 1 outcome. This choice satisfies conditions introduced in [7] and has already been used in [10]. In our prior distribution we assume that all parameters are independent:

$$\pi(\theta) \sim \exp \left(-\frac{1}{2} \sum_{i=1}^n \left(\frac{\theta^{(i)} - \theta_k^{(i)}}{s_k^{(i)}} \right)^2 \right),$$

and similarly as above we obtain

$$WR = \exp \left(\sum_{i=1}^n \frac{2\Delta_k^{(i)} c_k^{(i)} (\theta^{(i)} - \theta_k^{(i)})}{[\sigma^{(i)}]^2} \right)$$

and

$$w_k \approx \sum_{i=1}^n \frac{2\Delta_k^{(i)} c_k^{(i)} (\theta^{(i)} - \theta_k^{(i)})}{[\sigma^{(i)}]^2}.$$

This means that we can assume that $f(w_k|\theta)$ is approximately equal to the density of the normal distribution with mean

$$\sum_{i=1}^n \frac{2\Delta_k^{(i)} c_k^{(i)} (\theta^{(i)} - \theta_k^{(i)})}{[\sigma^{(i)}]^2}$$

and standard deviation τ . Using notation

$$A^{(i)} = \frac{2\Delta_k^{(i)} c_k^{(i)}}{[\sigma^{(i)}]^2}, \quad B^{(i)} = \frac{-2\Delta_k^{(i)} c_k^{(i)} \theta_k^{(i)}}{[\sigma^{(i)}]^2} \quad (6)$$

we finally obtain the posterior distribution

$$\begin{aligned} \pi(\theta|w_k) \sim \exp\left(-\frac{1}{2}\left(\frac{w_k - \sum_{i=1}^n (A^{(i)}\theta^{(i)} + B^{(i)})}{\tau}\right)^2\right) \times \\ \exp\left(-\frac{1}{2}\sum_{i=1}^n \left(\frac{\theta^{(i)} - \theta_k^{(i)}}{s_k^{(i)}}\right)^2\right). \end{aligned} \quad (7)$$

We see that this is again a multivariate normal distribution but not with independent variables. In what follows, we first try to approximate it with an independent case and then perform a full calculation using a covariance matrix.

Formula (7) could be rewritten as

$$\begin{aligned} \pi(\theta|w_k) \sim \exp\left(-\frac{1}{2}\sum_{i=1}^n \left([\theta^{(i)}]^2 \left(\frac{[A^{(i)}]^2}{\tau^2} + \frac{1}{[s_k^{(i)}]^2}\right) \right. \right. \\ \left. \left. - 2\theta^{(i)} \left(\frac{A^{(i)}w_k - A^{(i)}\sum_{j=1}^n B^{(j)} - \frac{1}{2}A^{(i)}\sum_{j \neq i} A^{(j)}\theta^{(j)}}{\tau^2} + \frac{\theta_k^{(i)}}{[s_k^{(i)}]^2}\right)\right)\right), \end{aligned}$$

and if we use an approximation

$$\sum_{j \neq i} A^{(j)}\theta^{(j)} \approx \sum_{j \neq i} A^{(j)}\theta_k^{(j)} = -\sum_{j \neq i} B^{(j)}, \quad (8)$$

we obtain

$$\begin{aligned} \pi(\theta|w_k) \sim \exp\left(-\frac{1}{2}\sum_{i=1}^n \left([\theta^{(i)}]^2 \left(\frac{[A^{(i)}]^2}{\tau^2} + \frac{1}{[s_k^{(i)}]^2}\right) \right. \right. \\ \left. \left. - 2\theta^{(i)} \left(\frac{A^{(i)}w_k - A^{(i)}B^{(i)} - \frac{1}{2}A^{(i)}\sum_{j \neq i} B^{(j)}}{\tau^2} + \frac{\theta_k^{(i)}}{[s_k^{(i)}]^2}\right)\right)\right) \\ \sim \exp\left(-\frac{1}{2}\sum_{i=1}^n \frac{[A^{(i)}]^2[s_k^{(i)}]^2 + \tau^2}{[s_k^{(i)}]^2\tau^2} \times \right. \\ \left. \left(\theta^{(i)} - \frac{[s_k^{(i)}]^2 A^{(i)}(w_k - B^{(i)} - \frac{1}{2}\sum_{j \neq i} B^{(j)}) + \theta_k^{(i)}\tau^2}{[A^{(i)}]^2[s_k^{(i)}]^2 + \tau^2}\right)^2\right). \end{aligned}$$

We have thus obtained (for any $i = 1, 2, \dots, n$)

$$\theta_{k+1}^{(i)} = \frac{[s_k^{(i)}]^2 A^{(i)}(w_k - B^{(i)} - \frac{1}{2}\sum_{j \neq i} B^{(j)}) + \theta_k^{(i)}\tau^2}{[A^{(i)}]^2[s_k^{(i)}]^2 + \tau^2}$$

and

$$[s_{k+1}^{(i)}]^2 = \frac{[s_k^{(i)}]^2 \tau^2}{[A^{(i)}]^2 [s_k^{(i)}]^2 + \tau^2}.$$

Finally, using (6) we get

$$\begin{aligned} \theta_{k+1}^{(i)} &= \theta_k^{(i)} + \sum_{j \neq i} \frac{2\Delta_k^{(i)} \Delta_k^{(j)} c_k^{(i)} c_k^{(j)} [s_k^{(i)}]^2 [\sigma^{(i)}]^2}{[\sigma^{(j)}]^2 (4[c_k^{(i)}]^2 [s_k^{(i)}]^2 + \tau^2 [\sigma^{(i)}]^4)} \cdot \theta_k^{(j)} \\ &\quad + \frac{2\Delta_k^{(i)} c_k^{(i)} [s_k^{(i)}]^2 [\sigma^{(i)}]^2}{4[c_k^{(i)}]^2 [s_k^{(i)}]^2 + \tau^2 [\sigma^{(i)}]^4} \cdot w_k, \end{aligned} \quad (9)$$

$$[s_{k+1}^{(i)}]^2 = \frac{[s_k^{(i)}]^2 \tau^2 [\sigma^{(i)}]^4}{4[c_k^{(i)}]^2 [s_k^{(i)}]^2 + \tau^2 [\sigma^{(i)}]^4}. \quad (10)$$

Moreover, we expect better results if we do not use the approximation (8). To do so, we take a general multivariate normal prior

$$\pi(\theta) \sim \exp\left(-\frac{1}{2}(\theta - \theta_k)^\top S_k^{-1}(\theta - \theta_k)\right),$$

with a positive definite covariance matrix S_k to rewrite (7) in the form

$$\begin{aligned} \pi(\theta|w_k) &\sim \exp\left(-\frac{1}{2}\left(\frac{w_k - \sum_{i=1}^n (A^{(i)}\theta^{(i)} + B^{(i)})}{\tau}\right)^2\right) \times \\ &\quad \exp\left(-\frac{1}{2}(\theta - \theta_k)^\top S_k^{-1}(\theta - \theta_k)\right). \end{aligned}$$

To simplify we use a substitution $\tilde{\theta} = \theta - \theta_k$ to obtain

$$\pi(\tilde{\theta}) \sim \exp\left(-\frac{1}{2}\tilde{\theta}^\top S_k^{-1}\tilde{\theta}\right)$$

and

$$\pi(\tilde{\theta}|w_k) \sim \exp\left(-\frac{1}{2}\left(\frac{w_k - \sum_{i=1}^n A^{(i)}\tilde{\theta}^{(i)}}{\tau}\right)^2\right) \cdot \exp\left(-\frac{1}{2}\tilde{\theta}^\top S_k^{-1}\tilde{\theta}\right). \quad (11)$$

We want to show that $\pi(\tilde{\theta}|w_k)$ could be written in the form

$$\pi(\tilde{\theta}|w_k) \sim \exp\left(-\frac{1}{2}(\tilde{\theta} - b_k)^\top S_{k+1}^{-1}(\tilde{\theta} - b_k)\right), \quad (12)$$

for some vector b_k and some matrix S_{k+1}^{-1} . We obtain that by equating the coefficients of polynomials in the exponents of (11) and (12). We denote elements of S_k^{-1} as (s_{ij}) and elements of S_{k+1}^{-1} as (t_{ij}) , and after equating the coefficients in front of $[\tilde{\theta}^{(i)}]^2$ we get

$$t_{ii} = s_{ii} + \frac{[A^{(i)}]^2}{\tau^2}, \quad (13)$$

while after equating the coefficients in front of $\tilde{\theta}^{(i)}\tilde{\theta}^{(j)}$ (for $i \neq j$) we get

$$t_{ij} = s_{ij} + \frac{A^{(i)}A^{(j)}}{\tau^2}. \quad (14)$$

So, we obtain simple rules for updating elements of S_k^{-1} given by (13) and (14). Actually, (14) covers also (13) by taking $i = j$.

Obtaining b_k is not straightforward. Namely, by equating the coefficients in front of $\tilde{\theta}^{(i)}$ we get

$$t_{ii}b_k^{(i)} + \sum_{j \neq i} t_{ij}b_k^{(j)} = \frac{A^{(i)}w_k}{\tau^2},$$

which is a system of linear equations, which written in the matrix form reads

$$S_{k+1}^{-1} \cdot \begin{bmatrix} b_k^{(1)} \\ b_k^{(2)} \\ \vdots \\ b_k^{(n)} \end{bmatrix} = \frac{w_k}{\tau^2} \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{bmatrix}. \quad (15)$$

The matrix S_{k+1}^{-1} is almost diagonal. Namely,

$$S_1^{-1} = \text{diag} \left(\frac{1}{[s_1^{(1)}]^2}, \frac{1}{[s_1^{(2)}]^2}, \dots, \frac{1}{[s_1^{(n)}]^2} \right),$$

and by (14) we update its elements with

$$\frac{A^{(i)}A^{(j)}}{\tau^2} = \frac{4\Delta_k^{(i)}\Delta_k^{(j)}c_k^{(i)}c_k^{(j)}}{\tau^2[\sigma^{(i)}]^2[\sigma^{(j)}]^2}.$$

c_k is quickly converging and $\Delta_k^{(i)}\Delta_k^{(j)}$ (for $i \neq j$) has a Bernoulli $\frac{1}{2}$ distribution with probability $\frac{1}{2}$ for each ± 1 outcome, which means that non-diagonal elements remain pretty close to zero all the time if we take $\sigma^{(i)}, \sigma^{(j)}$ large enough (this can always be achieved by rescaling θ if

necessary). This is the reason why we used Gauss-Jordan elimination method without pivoting [2, Section 4.2] for solving the system (15) and never encountered neither the speed problems, neither the division by zero problems. Finally, we update our belief about parameters simply with

$$\theta_{k+1}^{(i)} = \theta_k^{(i)} + b_k^{(i)}. \tag{16}$$

3 Experimental results

As already announced, we now compare BSPSAS (BSPSA simple) method given by (9),(10) and BSPSA method given by (14)–(16) with SPSA method described in [8] and implemented in [10]. SPSA updates parameters by formula

$$\theta_{k+1}^{(i)} = \theta_k^{(i)} + \frac{a_k^{(i)}}{\Delta_k^{(i)} c_k^{(i)}} \cdot w_k,$$

where

$$a_k^{(i)} = \frac{a^{(i)}}{(A+k)^\alpha}, \quad c_k^{(i)} = \frac{c^{(i)}}{k^\gamma}, \tag{17}$$

and where we use $\alpha = 0.602$, $\gamma = 0.101$, $A = 0.1N$ ($N =$ total number of iterations) – values recommended in [8]. Actually, in [10] new variables are introduced by

$$R_k^{(i)} = \frac{a_k^{(i)}}{[c_k^{(i)}]^2}$$

and $R^{(i)} = R_N^{(i)}$ are used as input variables. We use this implementation, and so we discuss optimal values for $R^{(i)}$ later on.

In BSPSA(S), we have used the same formula for $c_k^{(i)}$ so far, but that is also a somewhat problematic choice, as it will be discussed later. We also use fixed value $\tau = 0.6$. Let us recall that τ is the standard deviation of the approximate conditional distribution of w_k (a result of the two-game match), which can be calculated to be 0.6 for two equally strong engines with 82% draw rate (which is measured to be the case in our experimental tests). So, for useful application of BSPSA(S) it remains to obtain optimal values for $s_1^{(i)}$ and $\sigma^{(i)}$, just as it remains to obtain optimal values for $R^{(i)}$ in SPSA method. Before we start, we recall the famous Elo method for measuring the relative strength of chess engines/players (see [4] for more details).

Definition 1. *We say that an engine/player E_1 is x Elo points stronger than an engine/player E_2 if his/her expected score (win percentage WP*

in our earlier notation) in their match is

$$WP = \frac{1}{1 + 10^{-\frac{x}{400}}}.$$

We also use the notation

$$\text{Elo}(x) = |p_2 - p_1|$$

if an engine with parameter $p_2 \in \mathbb{R}$ is x Elo points stronger than an engine with parameter $p_1 \in \mathbb{R}$.

First, we did many experiments under the perfect conditions of the simulator. Namely, we assumed that all parameters have an optimal value 0 and that Elo loss is quadratic around the optimum, which in the notation of the previous definition means $x = a(p_1^2 - p_2^2)$ for some constant $a > 0$. The code written in Perl [13] for the BSPSA algorithm (simulator mode is the part of the code), and the simulator results are available at [9].

Our first goal was to find hyperparameters ($R^{(i)}$ for SPSA and $s_1^{(i)}, \sigma^{(i)}$ for BSPSA(S)) that give the best results (on average) in the simulator mode. For SPSA we got an approximation

$$R^{(i)} = \frac{19362 \cdot \ln \left(1 + \frac{\text{Elo}(100)}{11405} \right)}{N^{0.6} (c_N^{(i)})^{1.6}}, \quad (18)$$

where, as already noted above, N is the total number of iterations during the execution of the algorithm. For BSPSA(S) we got much simpler formulae:

$$s_1^{(i)} = \Delta\theta^{(i)}, \quad \sigma^{(i)} = \text{Elo}(100), \quad (19)$$

which do not depend on the number of iterations N , but which in turn require estimation of $\Delta\theta^{(i)}$ – the distance of the parameter $\theta^{(i)}$ from the optimum. Furthermore, in both cases, we also need to estimate $\text{Elo}(100)$, as defined in Definition 1. That can be done with high precision with a relatively low number of games, and during that calculation, we can also get some approximation for $\Delta\theta^{(i)}$.

Then we used formulae (18) and (19) to run simulations with 1, 4, 16 and 64 parameters and 200000 iterations, where in all cases, the total distance from the optimum was 2 Elo points at the beginning of the simulation. We discuss the choice for $c_N^{(i)}$ later in the concluding section. We repeated each run 50 times to reduce statistical errors. The obtained results are presented in Table 1.

Method	Elo gain mean / standard deviation			
	1 parameter	4 parameters	16 parameters	64 parameters
SPSA	1.99944 / 0.00071	1.9929 / 0.0044	1.9048 / 0.0331	1.2616 / 0.1159
BSPSAS	1.99970 / 0.00042	1.9951 / 0.0035	1.9333 / 0.0256	1.2525 / 0.1093
BSPSA	1.99968 / 0.00044	1.9953 / 0.0033	1.9303 / 0.0243	1.2683 / 0.1054

Table 1

We see that under perfect simulator conditions when parameters are independent and parabolically behaved around the optimum, there is basically no difference between BSPSAS given by (9),(10) and BSPSA given by (14)–(16). So, for tuning a large number of parameters (> 100), where speed could be crucial, BSPSAS seems like a great choice, but for tuning a lower number of parameters, we still prefer BSPSA. We see that BSPSA(S) comes closer to the optimum (up to more than 40%) than SPSA, and the closer to the optimum we are, the more significant the difference is. BSPSA(S) is also more stable, with a lower standard deviation of the measurements. The difference is low for 64 parameters, but it seems that 200000 iterations are simply not enough in that case.

Finally, we put our method to the real test. We modified 29 parameters in the Stockfish chess engine to obtain an engine that is around 180 Elo weaker than the original Stockfish. Then we tried to tune those parameters with 100000 iterations using both SPSA and BSPSA and using formulae (18) and (19) to calculate input values for each tuner. Now the real chess games were played with $5 + 0.05$ time control (5 seconds for the whole game plus 50 milliseconds increment after each move – the maximum we could afford with our hardware). After tuning with SPSA, an engine was around 64 Elo weaker than the original Stockfish, while tuning with BSPSA brought him to -50 Elo compared to the original Stockfish, which is more than 20% closer to the optimum than the result obtained with SPSA. Moreover, we noticed that one parameter got stuck in the local optimum. After modifying that single parameter to its original value, we were around -22 Elo points for SPSA and around -15 Elo points for BSPSA.

4 Conclusion

An improvement was found over the classical SPSA method. Moreover, we believe that further improvements are possible because we used the same rule (17) for updating $c_k^{(i)}$ which has some side effects. Namely, if $c_k^{(i)}$ values are too large and parameter behavior around an optimum is

asymmetrical, we shall miss the optimum no matter how many iterations our algorithm takes. On the other hand, if $c_k^{(i)}$ values are too small, an algorithm will have slow progress and likely get stuck in the local optimum. The following picture shows an Elo dependence of one parameter that got stuck in the red region during our tuning tests.

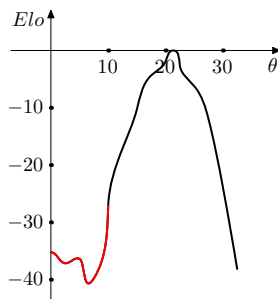


Figure 1: Behaviour of one critical parameter

Here is an idea for future work: to develop an algorithm that will change $c_k^{(i)}$ values intelligently, trying to distinguish between local and global optima. We believe that such algorithms could then be applied in many research areas, including training of neural networks and some strictly deterministic problems like finding optima of functions of several (many) variables.

Acknowledgements. This work has been supported in part by the Croatian Science Foundation under the project 2449 MiTPDE. The second author acknowledges the financial support of the Ministry of Education, Science and Technological Development of the Republic of Serbia (Grant No. 451-03-68/2020-14/ 200125). We are thankful to Professor Nataša Krklec–Jerinkić for her feedback and valuable discussions.

References

- [1] Bolstad W.M., Introduction to Bayesian Statistics, 2nd edition, John Wiley & Sons, 2007.
- [2] Ciarlet P.G., Introduction to numerical linear algebra and optimization, Cambridge University Press, 1989.
- [3] *Computer Chess Rating Lists*, available at <http://ccrl.chessdom.com/ccrl/4040>.
- [4] *Elo rating system*, available at https://en.wikipedia.org/wiki/Elo_rating_system.

- [5] Frazier P. I., *A Tutorial on Bayesian Optimization*, arXiv: 1807.02811v1, (2018).
- [6] Spall J.C., *A stochastic approximation technique for generating maximum likelihood parameter estimates*. In: Proceedings of the American Control Conference, 1987, pp. 1161–1167.
- [7] Spall J.C., *Multivariate stochastic approximation using a simultaneous perturbation gradient approximation*, IEEE Trans. Automat. Contr. **37** (1992), 332–341.
- [8] Spall J.C., *Implementation of the Simultaneous Perturbation Algorithm for Stochastic Optimization*, IEEE Trans. Aerosp. Electron. Syst. **34** (3) (1998), 817–823.
- [9] *SPSA Tuner and its variants*, available at <https://github.com/Iivec/spsa-1/tree/bspsa>.
- [10] *SPSA Tuner for Stockfish Chess Engine*, available at <https://github.com/zamar/spsa>.
- [11] *Stockfish Testing Framework*, available at <https://tests.stockfishchess.org/tests>.
- [12] *Top Chess Engine Championship*, available at <https://tcec-chess.com>.
- [13] L. Wall, T. Christiansen T and J. Orwant, *Programming Perl*, 3rd edition, O'Reilly, 2000.

Ivan Ivec

University of Zagreb, Faculty of Metallurgy, Aleja narodnih heroja 3,
Sisak 44000, Croatia

E-mail address: `iivec@simet.unizg.hr`

Ivana Vojnović

University of Novi Sad, Faculty of Sciences, Department of Mathematics
and Informatics, Trg Dositeja Obradovića 4, Novi Sad 21000, Serbia

E-mail address: `ivana.vojnovic@dmi.uns.ac.rs`