SABR Enhancements

- Fast Calibration, Arbitrage-Free Extrapolation, Arbitrage-Free SABR Implementation

Joshua Xingzhi Zhang
Bloomberg L. P. xzhang1@bloomberg.net
First version 11/15/2013

Abstract

In this article, we will discuss a few enhancements for SABR model implementation, first we will introduce a fast SABR calibration with the standard Hagan's formula by reducing the number of model parameters; then we will address the negative probability at the low strike wing with the Hagan's formula, and propose an arbitrage free patching to fix it by using shifted-Black model with Dirichlet boundary condition; we will also discuss an enhancement mapping from regular SABR model to solvable zero-correlation SABR; in the end we will introduce an efficient and accurate arbitrage free SABR implementation.

Key Words: SABR, arbitrage-free, shifted-black.

Introduction

SABR model was introduced about 10 years ago[1]

$$dF = \sigma C(F) dW_1, where\ C(F) = F^{\beta}$$
$$d\sigma = \nu\sigma dW_2$$
$$\langle dW_1 dW_2 \rangle = \rho dt$$

It becomes the market standard for interest rate derivative in recent years. There is no exact analytic solution except some special cases, such as $\beta = 0$[2], $\beta = 1$[4] , or $\rho = 0$[5]0. In practice, most people use the analytic approximation result proposed by Hagan et al [9].

A standard swaption volatility cube has very large number of grids need to be calculated. An efficient SABR calibration is still important even with the analytic approximation formula. In this article we will reduce the 4 parameter SABR model ($\alpha = \sigma(0), \beta, \nu, \rho$) into 3 parameters by matching α to ATM volatility (In practice, people usually fixe ρ orβ, then the model will be reduce from 3 parameters to 2 parameters). We find this method will speed up the total calibration about 8 times. This will be described on section 2.

Hagan's SABR formula will introduce negative probability for low strike wing for long term options. This becomes an issue in recent years due to the low rate environment. There are a lot researches trying the fix this problem.

Benaim et al [12] used a convex call option pricer to extrapolate the low strike wing directly, if this pricer can be easily find, this method will guarantee the positive probability because the convex shape. The problem with this method is that the convex call option pricer is not easy to find. Johnson and Nonas [13] modified Hagan's formula directly to satisfy Lee's low-strike moment formula [14] for SABR model with $\beta = 1$. This method can be practically arbitrage free, but lucks mathematic prove. The calculation can also be slow when $\beta \neq 1$, it involves non-central chi-squared calculation for CEV process.

With similar ideas, in this article we will use shifted-Black model to fix the low strike of the original Hagan's formula. Shifted-Black model has been used to extrapolate the SABR formula for high strike [15]. Instead of using the standard shifted-Black, we will use Dirichlet boundary condition, which will guarantee the low strike always having Black implied volatility as the original SABR model. This will be described in section 3.

Antonov *et al* [5] made breakthrough for 0-correlation SABR model ($\rho = 0$). They simplify the call option pricing into a two-dimension integration (or one-dimension if approximation is used). They also proposed a mapping technique to map the general SABR model into solvable 0-correlation SABR. But they did not find a consistent method for different model parameters. In this article we will use the standard Hagan formula to improve the mapping technique, and we find it works for all model parameters. The problem with the mapping technique is that, not all SABR model can be map to 0-correlation SABR, and the mapping does not guarantee arbitrage free. This will be described in section 4.

Besides the exact solution of SABR model, either analytic solution or numerical method such as PDE, directly working on the probability distribution is another approach to guarantee arbitrage free. Doust[8] and Hagan et al [3] recently publish some work with this approach. Doust used the perturbation method introduced in the original SABR model to derive out the approximated probability distribution function, and then modified it with arbitrage free condition, such as conservation of total probability and forward rate, then the option price is just a one-dimension integration on the payoff function. His method involves a MC calculation on the total probability at the boundary of $f = 0$. Hagan and Kumar[3] first approximated the SABR model with a local volatility model, and then use a one-dimension PDE solve the local volatility model to get the probability distribution. The arbitrage conditions can be easily satisfied in the PDE algorithm. The only problem with this approach is that one-dimension PDE is still slow if we use it on a swaption cube.

Following similar approach as Doust and Hagan, in this article we will directly work on the probability distribution function. We use the improved SABR approximation method based on differential geometry [7][4] to get the probability distribution function and then adjust it to satisfy the conservation of forward rate. We do not directly work on the conservation of the total probability since there is non-zero mass at boundary. With this idea an option pricing is just a one-dimension integration, which will be much more efficient that the above two approaches. We compared the result with MC and find this method is very good. This will be described in section 5.

II. SABR Calibration Speedup.

The analytic approximation result proposed by Hagan et al [9] is the market standard for SABR model implementation. The implied Black volatility is approximated as:

$$\sigma_B(K,F_0)=\frac{\nu\log\left(\frac{F_0}{K}\right)}{\chi(\zeta)}\left\{1+\left[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(F_0K)^{1-\beta}}+\frac{1}{4}\frac{\rho\alpha\beta\nu}{(F_0K)^{(1-\beta)/2}}+\frac{2-3\rho^2}{24}\nu^2\right]\tau_{ex}\right\}$$

With

$$\zeta=\frac{\nu}{\alpha}\frac{{F_0}^{1-\beta}-K^{1-\beta}}{1-\beta},\ \chi(\zeta)=\log\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2}-\rho+\zeta}{1-\rho}\right),\alpha=\sigma(0)$$

A standard swaption volatility cube usually has very large number of grids (with 1 moth to 30 year term dimension, 3 month to 30 year tenor dimension, 1% or lower to 10% or higher strike dimension). An efficient SABR calibration is still very important even using the analytic approximation formula.

In derivative market, usually at the money (ATM) options have most liquidity. There are some market quotes for out of the money (OTM) strikes, but not as liquid as ATM options, this is especially true for swaption market. To calibrate SABR model to swaption market, we should try to match ATM quote exactly.

For ATM volatility the Hagan formula will be reduced to:

$$\sigma_{ATM}=I_0\left\{1+\left[\frac{(1-\beta)^2}{24}I_0^2+\frac{1}{4}\rho\beta\nu I_0+\frac{2-3\rho^2}{24}\nu^2\right]\tau_{ex}\right\}$$

with

$$I_0=\frac{\alpha}{{F_0}^{1-\beta}}$$

We can solve α exactly for given ATM volatility σ_{ATM} and ρ,β,ν.

$$\alpha=I_0{F_0}^{1-\beta}$$

with

$$\begin{aligned}
I_0&=-\frac{B}{3A}+\frac{b_0}{3A}\sqrt[3]{\frac{2}{b_2}}-\frac{1}{3A}\sqrt[3]{\frac{b_2}{2}}\\
b_0&=3A-B^2+3AC\\
b_1&=-9AB+2B^3-9ABC-27A^2\sigma_{ATM}\\
b_2&=b_1+\sqrt{4(b_0)^3+(b_1)^2}\\
A&=\frac{(1-\beta)^2}{24}\tau_{ex}\\
B&=\frac{1}{4}\rho\beta\nu\tau_{ex}\\
C&=\frac{2-3\rho^2}{24}\nu^2\tau_{ex}
\end{aligned}$$

Then we successfully removed α from parameter space and the 4 parameter SABR will reduce to 3 parameters.

As discovered in the original SABR paper, both β and ρ have very simula impact the volatility skew, so people usually will fix β or ρ and SABR model will reduce to two parameters: $\beta(\rho)$ and ν. It will be very easy and fast for any calibration optimization algorithm. Based on our experiments a two parameter model will be about 8 times faster than three parameter SABR model if we try to calibrate the full swaption cube.

III. Arbitrage free extrapolation of low strike.

Hagan's approximated SABR formula is based on short term with strikes around ATM, it also works well for high strike ranges. But it introduces negative probability for low strike when the option has long time to expire. In swaption market, a 20-30 year time to expiring is very common. After 2008 financial crisis, interest rate becomes very low globally. So the negative probability domain of Hagan's formula becomes a practical issue.

The shifted-Black model is a well-known simply model with skew feature. We will use shifted-Black model for low strike range, and then glue with Hagan's formula for high strike range to get the full smile curve.

The dynamic of shifted-Black is

$$dF = \sigma_d(F + d)dW$$

d is the shift, and σ_d is the volatility of the shifted process.

SABR has boundary at $F = 0$, there is non-zero mass at that point in general. If we also define a Dirichlet boundary condition at $F = 0$ for shifted-Black process. Let $P(T, F_0, \sigma_d, d; F)$ is the PDF(probability distribution function) of the shifted-Black process with free boundary condition, then the PDF with Dirichlet boundary condition will be:

$$P^D(T, F_0, \sigma_d, d; F) = P(T, F_0, \sigma_d, d; F) - P(T, F_0, \sigma_d, d; -F)$$

And we have

$$P^D(T, F_0, \sigma_d, d; 0) = 0$$

The total mass of $P^D(T, F_0, \sigma_d, d; F)$ will be less than 1, the remaining probability will be a Dirac's delta function at $F = 0$.

Call or put option with strike K will be priced easily with shifted-Black process:

$$\begin{aligned} PV_{call}(T, F_0, \sigma_d, d; K) &= \int_0^\infty [F - K]^+ P(T, F_0, \sigma_d, d; F)dF \\ &= (F_0 + d)\Phi_{Norm}(d_1) - (K + d)\Phi_{Norm}(d_2) \end{aligned}$$

$$PV_{put}(T, F_0, \sigma_d, d; K) = \int_0^\infty [K - F]^+ P(T, F_0, \sigma_d, d; F) dF$$
$$= (K + d)\Phi_{Norm}(-d_2) - (F_0 + d)\Phi_{Norm}(-d_1)$$

with $d_1 = \frac{\log\left(\frac{F_0+d}{K+d}\right)}{\sigma_d\sqrt{T}} + \frac{1}{2}\sigma_d\sqrt{T}$, $d_2 = \frac{\log\left(\frac{F_0+d}{K+d}\right)}{\sigma_d\sqrt{T}} - \frac{1}{2}\sigma_d\sqrt{T}$. Φ_{Norm} is the cumulated normal distribution.

A call option with Dirichlet boundary condition is:

$$PV_{call}^D(T, F_0, \sigma_d, d; K) = \int_0^\infty [F - K]^+ P^D(T, F_0, \sigma_d, d; F) dF$$
$$= PV_{call}(T, F_0, \sigma_d, d; K) - PV_{put}(T, F_0, \sigma_d, d; -K)$$

$$PV_{put}(T, F_0, \sigma_d, d; -K) = 0, \text{ when } K > d$$

Then the cumulated distribution of the shifted-Black process with Dirichlet boundary condition Φ^Dcan be easily calculated. It is just a differential on the call price.

Next step is to glue the shifted-Black model with Hagan's SABR formula. To achieve this, we will find the glue strike K_g such that:

$$PV_{call}^D(T, F_0, \sigma_d, d; K_g) = PV_{call}^{SABR}(T, F_0, \alpha, \beta, \upsilon, \rho, K_g)$$
$$\Phi^D(T, F_0, \sigma_d, d; K_g) = \Phi^{SABR}(T, F_0, \alpha, \beta, \upsilon, \rho, K_g)$$

So the option price and cumulated probability distribution both match with these two models, and we get an arbitrage free modified cumulated distribution:

$$\Phi^M(T, F_0, \sigma_d, d, \alpha, \beta, \upsilon, \rho; F) = \begin{cases} \Phi^D(T, F_0, \sigma_d, d, F), & F \le K_g \\ \Phi^{SABR}(T, F_0, \alpha, \beta, \upsilon, \rho, F), & F > K_g \end{cases}$$

In practice, this arbitrage free extrapolation technique is very efficient, and can be used when fast SABR calibration is a needed.

In Figure 1, we show an example how the shifted-Black glues with Hagan's SABR formula. The SABR model parameters are: $= 0.06$, $\beta = 0.6$, $\nu = 0.33$, $\rho = -0.2$, $T = 20$, and $F_0 = 4.0\%$. The cumulated probability and PDF are calculated numerically. We can see the PDF is guaranteed positive with this method.

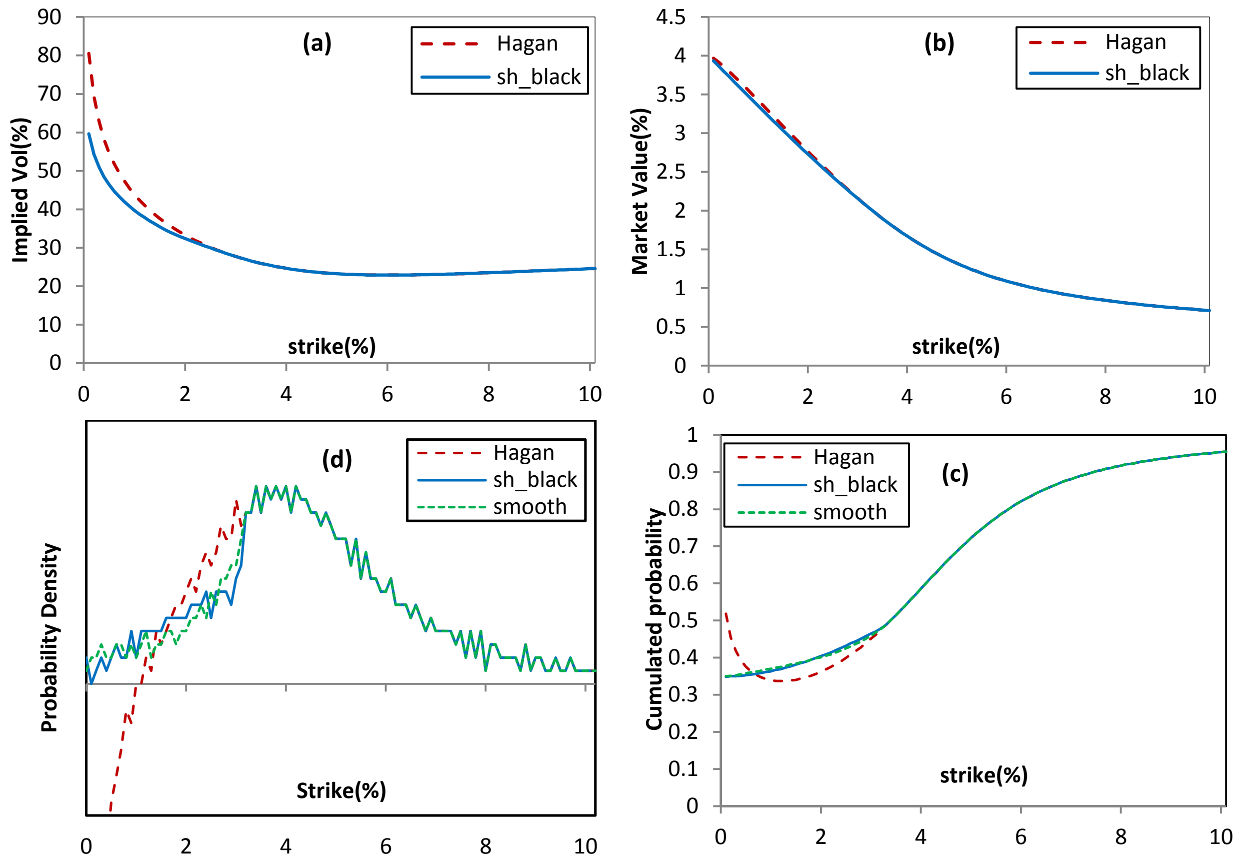

Figure 1. Extrapolate the Hagan volatility with shifted-black volatility. The SABR model parameters are: $= \mathbf{0.06}, \boldsymbol{\beta} = \mathbf{0.6}, \boldsymbol{\nu} = \mathbf{0.33}, \boldsymbol{\rho} = -\mathbf{0.2}, \boldsymbol{T} = \mathbf{20}$, and $\boldsymbol{F_0} = \mathbf{4.0}\%$.

People may want to use the shifted-Black process with free boundary instead of Dirichlet boundary proposed here. But with free boundary, you could end up a call option price which will be higher than the forward rate, which means there is no valid implied Black volatility. With Dirichlet boundary, there is always a valid implied Black volatility. And also the non-zero probability mass at $F = 0$ for this model is very similar to SABR model. So we would expect shifted-Black with Dirichlet boundary is a better model.

When we glue the two models, we requires the cumulated probability being continues but not guarantee smoothness, in theory the probability distribution function could have jump when switching between the two models. That's the drawback of this method, what we can do is we can smoothly transit the price from the shifted-Black model into SABR model. The result is show in the figure (c) (d), since the difference of price and volatility is so small we did not show the smooth result in (a) (b). We can see the PDF is smoother.

IV. Improving the mapping technique for zero-correlation SABR model.

Up to now we only discussed the standard Hagan's SABR formula. Hagan's formula is based on perturbation technique, the volatility will be off from true value when time to expiring is large. If we are interested in the true SABR model result, we have to look for alternative methods.

Recently Antonov et al [5]simplified the analytic result of SABR when $\rho = 0$ into a two-dimension integration, and proposed a mapping method to calculate the general SABR model:

$$
\begin{aligned}
&\tilde{\beta} = \beta \\
&\tilde{\nu}^2 = \nu^2 - \frac{3}{2}\left\{\nu^2\rho^2 + \rho\alpha\nu(1-\beta)F_0^{\beta-1}\right\} \\
&\tilde{\alpha} = \tilde{\alpha}_0 + T\tilde{\alpha}_1 \\
&\tilde{\alpha}_0 = \frac{2\phi\tilde{q}\tilde{\nu}}{\phi^2 - 1} \\
&\frac{\tilde{\alpha}_1}{\tilde{\alpha}_0} = f(\alpha, \beta, \nu, \rho, F_0, K)
\end{aligned}
$$

The detail of function $f(\alpha, \beta, \nu, \rho, F_0, K)$ is in the original paper. $\tilde{\alpha}, \tilde{\beta}, \tilde{\nu}$ and $\tilde{\rho} = 0$ are the parameters of the new mapped SABR model, which has analytic solution. So a general SABR model can be approximated with the analytic solvable 0 correlation SABR model. From their MC comparison, this mapping is good in general. To get a better result, they also proposed a hybrid mapping: for different parameter range used different mapping method.

The assumption behind the mapping idea is that, the approximated SABR has error from the true SABR value, if the approximated results of two different SABR models are same, then the true results of those two SABR models will be very close to each other. Antonov et al proposed mapping of $\tilde{\beta}, \tilde{\nu}$, and then solved $\tilde{\alpha}$ up to second order (0(T) order) based Paulot's [7] implied volatility.

Instead of using Paulots formula and analytically solving $\tilde{\alpha}$ up to second order, here we will use Hagan's original SABR formula, and numerically solve $\tilde{\alpha}$ such that the implied volatilities of the SABR models are the same. Since Hagan's formula is very simple, the numerical solution is very fast. With this simple change, we find the result can be improved a lot comparing with Antonov's original idea.

In following tables and Figure 2, we show some results. Here the Monte-Carlo (MC) and Antonov's mapping--orginal zero correlation mapping (Antonov) and Hybrid mapping (An-H Map) -- results are copied from paper [5]. Hagan's result is based formula in this paper. NewMap is the new mapping method proposed here.

We can see in Table 1 the hybrid mapping method has better result than original mapping method, in Table 2, the results are reversed. For both cases, the new mapping method gives good stable results, even though it does not always have the best answer.

Figure 2, we plot the relative errors of different methods comparing with MC,

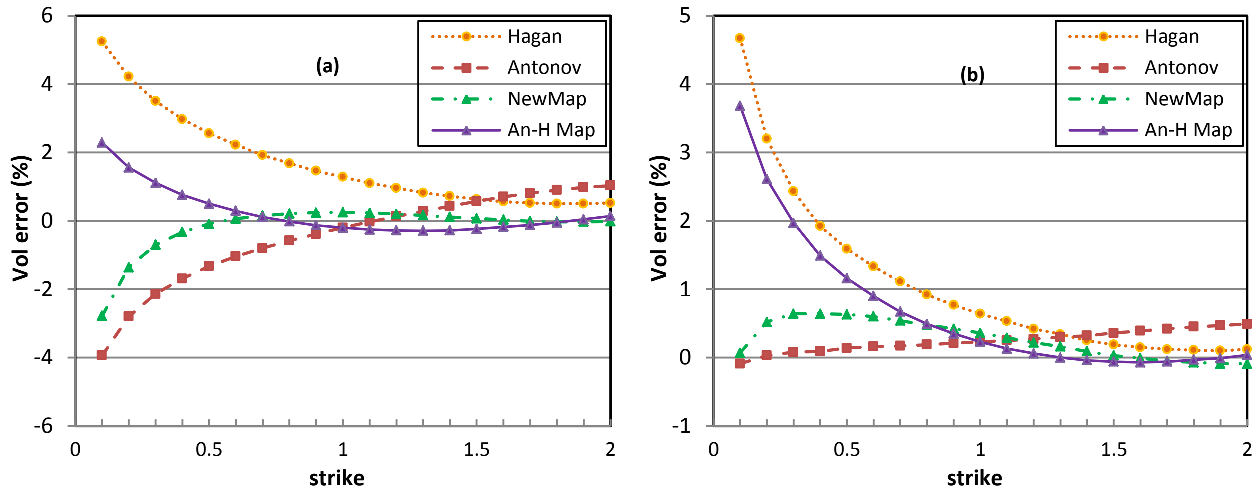

Figure 2. Implied volatility error of different methods comparing with MC.

K	Value (%)					Difference (bps)			
	MC	Hagan	Antonov	An-H Map	NewMap	Hagan	Antonov	An-H Map	NewMap
0.1	42.37	47.61	38.43	44.66	**39.59**	524	-394	229	***-278***
0.2	36.28	40.49	33.48	37.84	**34.92**	421	-280	156	***-136***
0.3	32.44	35.94	30.30	33.55	**31.74**	350	-214	111	***-70***
0.4	29.58	32.55	27.89	30.34	**29.25**	297	-169	76	***-33***
0.5	27.27	29.83	25.94	27.77	**27.18**	256	-133	50	***-9***
0.6	25.32	27.54	24.28	25.61	**25.38**	222	-104	29	***6***
0.7	23.64	25.56	22.84	23.75	**23.79**	192	-80	11	***15***
0.8	22.14	23.82	21.56	22.12	**22.35**	168	-58	-2	***21***
0.9	20.80	22.26	20.41	20.67	**21.04**	146	-39	-13	***24***
1.0	19.58	20.86	19.38	19.38	**19.83**	128	-20	-20	***25***
1.1	18.48	19.58	18.45	18.22	**18.71**	110	-3	-26	***23***
1.2	17.47	18.43	17.60	17.19	**17.67**	96	13	-28	***20***
1.3	16.56	17.38	16.84	16.27	**16.72**	82	28	-29	***16***
1.4	15.73	16.45	16.16	15.45	**15.84**	72	43	-28	***11***
1.5	14.99	15.62	15.56	14.75	**15.06**	63	57	-24	***7***
1.6	14.33	14.89	15.03	14.15	**14.36**	56	70	-18	***3***
1.7	13.77	14.29	14.58	13.65	**13.76**	52	81	-12	***-1***
1.8	13.29	13.79	14.19	13.24	**13.26**	50	90	-5	***-3***
1.9	12.89	13.39	13.87	12.93	**12.86**	50	98	4	***-3***
2.0	12.56	13.08	13.59	12.70	**12.54**	52	103	14	***-2***

Table 1: Implied vol and its error for different methods, 20 year maturity, $\alpha = 0.25, \beta = 0.6, \nu = 0.3, \rho = -0.8$. The Monte-Carlo (MC) and Antonov's two mapping-orginal mapping (Antonov) and Hybrid mapping (An-H Map) -results are from paper [5]. Hagan's result is based formula in this paper. NewMap is the new mapping method proposed here.

K	Value (%)					Difference (bps)			
	MC	Hagan	Antonov	An-H Map	NewMap	Hagan	Antonov	An-H Map	NewMap
0.1	51.23	55.90	51.14	54.91	**51.30**	467	-9	368	**7**
0.2	43.22	46.42	43.25	45.83	**43.74**	320	3	261	**52**
0.3	38.27	40.70	38.35	40.24	**38.91**	243	8	197	**64**
0.4	34.65	36.57	34.74	36.14	**35.29**	192	9	149	**64**
0.5	31.73	33.32	31.87	32.89	**32.36**	159	14	116	**63**
0.6	29.30	30.63	29.46	30.20	**29.90**	133	16	90	**60**
0.7	27.22	28.33	27.39	27.89	**27.76**	111	17	67	**54**
0.8	25.39	26.31	25.58	25.88	**25.87**	92	19	49	**48**
0.9	23.76	24.53	23.97	24.11	**24.18**	77	21	35	**42**
1.0	22.29	22.93	22.52	22.52	**22.65**	64	23	23	**36**
1.1	20.96	21.49	21.21	21.09	**21.25**	53	25	13	**29**
1.2	19.76	20.18	20.03	19.82	**19.98**	42	27	6	**22**
1.3	18.67	19.01	18.97	18.67	**18.83**	34	30	0	**16**
1.4	17.70	17.95	18.02	17.66	**17.79**	25	32	-4	**9**
1.5	16.83	17.02	17.19	16.77	**16.86**	19	36	-6	**3**
1.6	16.07	16.22	16.46	16.00	**16.06**	15	39	-7	**-1**
1.7	15.42	15.54	15.84	15.36	**15.37**	12	42	-6	**-5**
1.8	14.87	14.98	15.32	14.84	**14.80**	11	45	-3	**-7**
1.9	14.43	14.53	14.90	14.42	**14.34**	10	47	-1	**-9**
2.0	14.07	14.19	14.56	14.11	**13.98**	12	49	4	**-9**

Table 2: Implied vol and its error for different methods, 10 year maturity, $\alpha = 0.25, \beta = 0.6, \nu = 0.3, \rho = -0.8$.

As shown above zero-correlated SABR combining with mapping technique gives very good result for general SABR model. But there are a few issues with this approach.

1) The mapping technique could fail, for example $\tilde{\nu}$ does not have a real solution;

2) Even if there is valid mapping parameters, but just like other SABR approximations, it is mapping different strike options into different models (in this case different $\tilde{\alpha}$), then it will not guarantee arbitrage free. You can easily find some example that the mapped SABR has negative probability. If you map only the ATM strike option to get one $\tilde{\alpha}$ and apply it for the full smiles, it is arbitrage free since analytic zero-correlated SABR is arbitrage free, but as show in Antonov's original results, sometimes the error is too big.

3) Two-dimension integration is slow, even there is one-dimension approximation [6], to make the one-dimension integration stable for all cases is needed, this could be a technique challenge.

In the following we will introduce a different approach, which is efficient and arbitrage free, in the meantime the result is very close to the true SABR MC.

V. Efficient Arbitrage Free SABR implementation.

Simila to the works has been done by Doust[8] and Hagan et al [3], we directly work on the probability density function(PDF). The call option price can then be numerical calculated via:

$$PV_{call}(T,F_0,\sigma;K)=\int_0^{\infty}[F-K]^{+}P(T,F_0,\sigma;F)dF$$

We follow the approximation method to get the PDF as Doust[8]:

$$P(\tau,F_0,\sigma;F)=\frac{1}{\sigma F^{\beta}}\frac{J(z)^{-\frac{3}{2}}}{\sqrt{2\pi\tau}}\sqrt{\frac{F_0^{\beta}}{F^{\beta}}}\exp\left(-\frac{x(z)^2}{2\tau}+\left(\tilde{h}(z)+k(z_0)\tau\right)\right)$$

$$z=\frac{1}{\sigma}\int_F^{F_0}\frac{df}{C(f)}=z_{F_0}-z_F$$

$$z_F=\frac{1}{\sigma}\int_0^F\frac{df}{C(f)}=\frac{F^{1-\beta}}{\sigma(1-\beta)},z_{F_0}=\frac{1}{\sigma}\int_0^{F_0}\frac{df}{C(f)}=\frac{{F_0}^{1-\beta}}{\sigma(1-\beta)}$$

$$J(z)=\sqrt{1-2\rho\nu z+\nu^2z^2}$$

$$x(z)=\int_0^z\frac{d\theta}{J(\theta)}=\frac{1}{\nu}\ln\left(\frac{\sqrt{1-2\rho\nu z+\nu^2z^2}-\rho+\nu z}{1-\rho}\right),x'(z)=\frac{1}{J(z)}$$

$$k(z)=\frac{1}{8}\left(2\ -3\frac{(\rho-\nu z)^2}{J(z)^2}\ \right)\nu^2-\frac{1}{4}\frac{\rho\nu\beta}{(1-\beta)(z_F+z)}-\frac{\beta(2-\beta)}{8(1-\beta)^2(z_F+z)^2}$$

$$\tilde{h}(z)=\begin{cases}\frac{1}{2}\rho\nu\sigma\int_0^z\frac{\xi}{J(\xi)^2}d\xi\,, & \beta=1\\ \frac{1}{2}\rho\nu\sigma\int_0^z\frac{\beta}{\sigma\xi(1-\beta)+F^{1-\beta}}\frac{\xi}{J(\xi)^2}d\xi\,, & 0\le\beta<1\end{cases}$$

$$=\begin{cases}\frac{\rho\sigma}{2\nu}\left[\ln\big(J(z)\big)+\frac{\rho}{\sqrt{1-\rho^2}}\left(\tan^{-1}\frac{\nu z-\rho}{\sqrt{1-\rho^2}}+\tan^{-1}\frac{\rho}{\sqrt{1-\rho^2}}\right)\right], & \beta=1\\ \frac{1}{2}\frac{\beta\rho}{(1-\beta)}\frac{1}{J(-z_F)^2}\left[\nu z_F\ln\left(\frac{z_FJ(z)}{z_{F_0}}\right)+\frac{1+\rho\nu z_F}{\sqrt{1-\rho^2}}\left(\tan^{-1}\frac{\nu z-\rho}{\sqrt{1-\rho^2}}+\tan^{-1}\frac{\rho}{\sqrt{1-\rho^2}}\right)\right], & 0\le\beta<1\end{cases}$$

except $k(z)$ is defined differently, most formulas can be found in Doust[8] paper. We found this new $k(z_0)$ give a better result with $z_0=z/2$.

There are some basic conditions for arbitrage free pricing: 1) the total probability sums up to 1; 2) the average rate based on PDF matches with the forward rate. Since we are using approximation method to get PDF, these conditions cannot be automatically satisfied.

To satisfy these arbitrage free conditions, Doust[8] used a MC to get the total probability mass at $f=0$, and then adjusted the approximated PDF by shifting $F_0$and scaling.

However, we find that if we ignore the conservation of the total probability, and just scale the approximated PDF with scaling factor L to satisfy the forward rate:

$$F_0 = \int_0^{\infty} F\tilde{P}(T, F_0, \sigma; F)dF$$
$$\tilde{P}(T, F_0, \sigma; F) = LP(T, F_0, \sigma; F)$$

It will greatly simplify the calculation.

Since we do not calculate the total probability mass at $f = 0$, we can assume that after the scaling of PFD, if the sum is different from 1.0, then the missing mass is stored at $f = 0$. In practice we find this is a very approximation, and the result is even better than Doust's result as shown in Figure2.

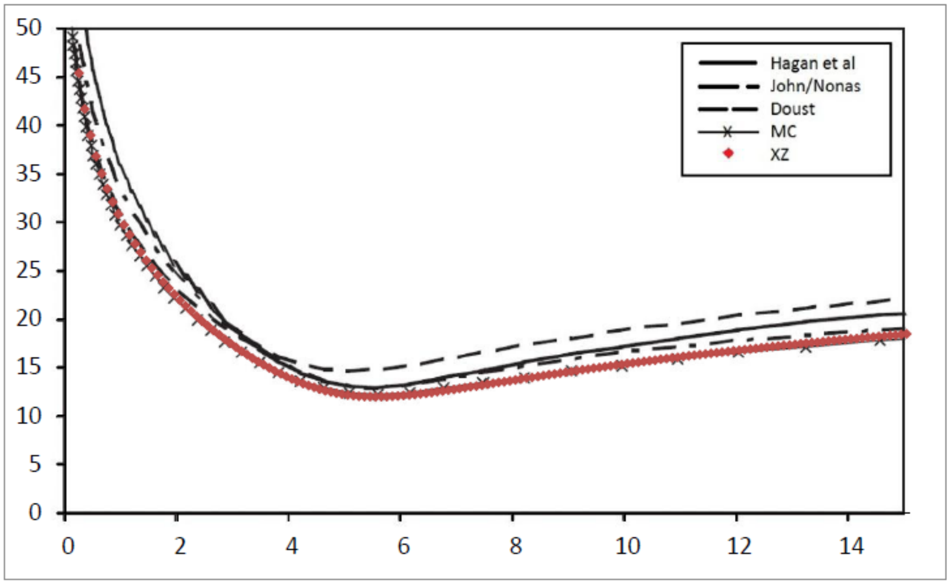

Figure 2. The implied volatility for different strikes based on different methods: Hagan et al formula, Johnson/Nonas formula, Doust formula, and Monte Carlo results are copied from Doust's original paper. XZ is the result based on this paper. The SABR model parameters are: $\alpha = 0.026, \beta = 0.5, \nu = 0.4, \rho = -0.1$, with time to expiring $T = 10$, and $F_0 = 4.88\%$.

In Figure 2, we showed the implied volatility for different strikes with different methods, the results of Hagan, Johnson/Nonas, Doust and Monte Carlo are copied directly from Doust paper. We can see the approximation method introduced in this paper is very accurate, it is very close to MC with the full range of strikes. We also found result of this method can be comparable with Antonov's mapping with the improvement based on section 5. Since for this set of model parameters, the difference is too small, we did not put the enhanced Antonov's result in the figure.

The above PDF is based on Hagan et al original perturbation method, which assumes short term and $f \to F_0$, Our modification works well for most long term cross all strikes. But in some very long term deal, for example above 30 year, there is errors for the implied volatility at very low strike range, even though it is arbitrage free. To further improve the result, we will use the WKB method based on differential geometry first introduced by Henry-Labordere [4], Paulot [7] gave more detail calculate for SABR model. Since this technique works for all strike, we expect the PDF derived out based on this will give a better answer.

Using this technique, the $0(T)$order correction for the PDF will be:

$$h_{WKB}(z) = -\frac{1}{2}\frac{\beta}{1-\beta}\frac{\rho}{\sqrt{1-\rho^2}}\left(\pi - \cos^{-1}\frac{\nu z - \rho}{J(z)} - \cos^{-1}\rho - I\right)$$

$$I = \begin{cases} \frac{2}{\sqrt{1-L^2}}\left(\tan^{-1}\frac{u_0 + L}{\sqrt{1-L^2}} - \tan^{-1}\frac{L}{\sqrt{1-L^2}}\right), & L < 1 \\ \frac{1}{\sqrt{L^2-1}}\ln\left(\frac{u_0\left(L+\sqrt{L^2-1}\right)+1}{u_0\left(L-\sqrt{L^2-1}\right)+1}\right), & L > 1 \end{cases}$$

$$u_0 = \frac{\rho}{\sqrt{1-\rho^2}} + \frac{J(z)-1}{\nu z\sqrt{1-\rho^2}}$$

$$L = \frac{J(z)}{\nu z_k\sqrt{1-\rho^2}}$$

We can also get the $0(T^2)$ order correction. But we find the above correction is already good enough for most test cases. In the following figures, we will show some result:

In Figure 3, we show the implied volatility for different strikes based on different method. MC and Antonov mapping results are same as Antonov's original paper. NewMap is the new mapping method introduced in previous section IV, ZWKB is the WKB based arbitrage free PDF method with the 0(T) order correcton. We can see PDF method is very accurate, and can be comparable with new mapping method, and be better than the original Antonov's mapping result. Comparing to two-dimension integration pricing for the mapping method, here we have one-dimension integration, this method will be more efficient. And most importantly, it always has positive PDF by design.

The technique described above generally works fine. In theory there could be some cases that when we scale the approximated PDF to satisfy the forward rate, the scale factor could be too big to push a total probability big than one, if this happens, then we have arbitrage. To solve this problem, we can just adopt Doust original idea by adjust the forward rate and scaling together to make sure the forward rate and total probability is correct. To simplify the calculation, here we can just assuming the mass at 0 is just 0. In practice, we do not find this kind of case in real market.

An application of this technique is to use it in LMM+SABR model. In LMM+SABR model, usually because we do not know the correct SABR distribution at large steps, then the Libor Monte Carlo usually is done with very small time steps, that will bring a lot of calculation burden. With our approximation of SABR PDF, we could directly simulate the Libor rate with big time steps, just like in the simply LMM model. And we believe it will greatly improve the performance.

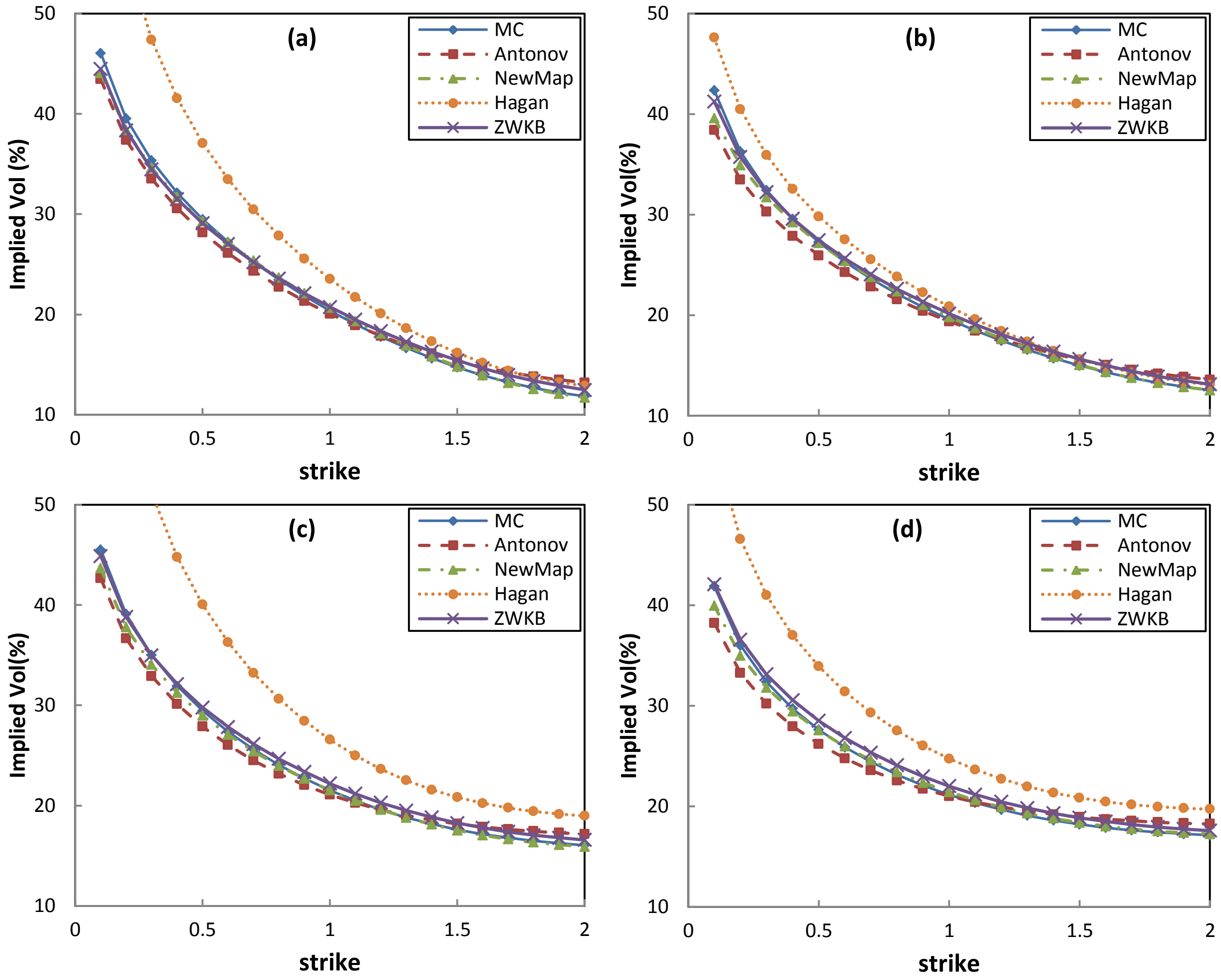

Figure 3. The implied volatility for different strikes based on different methods: Monte Carlo, Antonov et al mapping to 0 correlation (Antonov) results are taken from Antonov et al paper, the new mapping introduced in previous section (NewMap), and the WKB based PDF result (ZWKB) are also shown on top of each other. Time to expire is $T = 20$, and $F_0 = 1$. The SABR model parameters are: (a)$\alpha = 0.25, \beta = 0.3, \nu = 0.3, \rho = -0.8$; (b) $\alpha = 0.25, \beta = 0.6, \nu = 0.3, \rho = -0.8$; (c)$\alpha = 0.25, \beta = 0.3, \nu = 0.3, \rho = -0.5$; (d) $\alpha = 0.25, \beta = 0.6, \nu = 0.3, \rho = -0.5$.

VI. Reference

[1] Hagan P., Kumar D., Lesniewski A.S. and Woodward D. E.(2002), Managing Smile Risk, *Wilmott Magazine*, September 2002, 84-108.
[1] Hagan P., Kumar D., Lesniewski A.S. and Woodward D. E.(2002), Managing Smile Risk, *Wilmott Magazine*, September 2002, 84-108.
[2] Hagan P., Lesniewski A., Woodward D.(2005), Probability distribution in the SABR model of stochastic volatility, www.lesniewski.us/papers/working/ProbDistrForSABR.pdf

[3] Hagan P. and Kumar D. (2012), Arbitrage Free SABR, working paper.

[4] Henry-Labordere, P.(2005), A General Asymptotic Implied Volatility for Stochastic Volatility Models,
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=698601

[5] Antonov A. and Spector M.(2012), Advanced Analytics for the SABR model, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2026350

[6] Antonov A., Konikov M. and Spector M.(2013), SABR apreads its wings, *Risk*, August 2013, 58-63.

[7] Paulot L (2009), Asymptotic Implied Volatility at the Second order with Application to the SABR Model,
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1413649

[8] Doust P. (2012) No-arbitrage SABR, *The Journal of Computational Finance*, volume 15, 3-31.

[9] Obloj J. (2008), Fine-tune your smile correction to Hagan et al, *Wilmott Magazine*, June 2008, 102.

[10] Gauthier P. and Rivaille P.-Y. H.(2009), Fitting the smile: Smart parameters for SAVR and Heston, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1496982&rec=1&srcabs=1413649&alg=1&pos=1.

[11] Islah O.(2009), Solving SABR in exact form and unifying it with LIBOR market model, SSRN

[12] Benaim S, Dodgson M. and Kainth D. (2009), An arbitrage-free method for smile extrapolation. http://www.quarchome.org/RiskTailsPaper_v5.pdf

[13] Johnson S. and Nonas B. (2009), Arbitrage-free construction of swaption cube. SSRN

[14] Lee R. W (2004), The moment formula for implied volatility at extreme strikes, *Mathematical Finance*, vol 14, 469-480.

[15] Kluge W. (2012), Risk Management in the presence of extreme smiles, WBS-the 8th Fixed Income Conference, Vienna.

[16] Zhang J.(2013),