Rough SABR Forward Market Model

Reo Adachi[*], Masaaki Fukasawa[†a], Naoki Iida[*], Mitsumasa Ikeda[*], Yo Nakatsu[*], Ryota Tsurumi[*b], and Tomohisa Yamakami[*‡c]

Abstract This paper advances interest rate modeling in the post-LIBOR era by introducing rough stochastic volatility into the Forward Market Model (FMM). We establish a rigorous asymptotic expansion of swaption implied volatility, connecting the FMM to a rough Bergomi-type framework for forward swap rates. This contribution bridges the gap between Heath-Jarrow-Morton (HJM)-consistent forward term rate models and forward swap rate models with stochastic volatility, offering a parsimonious yet precise framework for modeling swaption volatility surfaces. Furthermore, we justify and generalize the widely used "freezing" approximation within a rigorous mathematical framework. The proposed approach enhances the representation of persistent skew and term structure, addressing key challenges in modern fixed income markets.

Key words: Rough Volatility, Forward Market Model (FMM), Volatility Term Structure, Swaption IV, Asymptotic Expansion, Calibration, vol-of-vol

1 Introduction

Interest rate modeling serves as the cornerstone for pricing, hedging, and managing risk in fixed income markets. A diverse array of models has been developed to track

---

[*] Mizuho–DL Financial Technology Co., Ltd., Kojimachi-odori Building, 2-4-1 Kojimachi, Chiyoda-ku, Tokyo 102-0083, Japan.

[†] Graduate School of Engineering Science, The University of Osaka, 1-3 Machikaneyama, Toyonaka, Osaka 560–8531, Japan.

[‡] Graduate School of Economics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

[a] Email: fukasawa@sigmath.es.osaka-u.ac.jp

[b] The corresponding author. Email: ryota-tsurumi@fintec.co.jp

[c] Email: tomohisa-yamakami@fintec.co.jp

the evolution of interest rates across maturities, striving to achieve a balance between theoretical rigor, empirical realism, and computational feasibility.

The Heath-Jarrow-Morton (HJM) framework provides a general arbitrage-free methodology for modeling the term structure of interest rates. By specifying the risk-neutral dynamics of forward rates, it accommodates a wide range of market behaviors. However, the HJM framework relies on infinitesimal forward rates, which are abstract quantities not directly observable or traded. To enable practical implementation and calibration, tractable submodels are required, focusing on the dynamics of market-observable quantities such as forward swap rates or rates linked to overnight indexed swaps (OIS).

Historically, the LIBOR Market Model (LMM), also referred to as the Brace-Gatarek-Musiela (BGM) model, dominated interest rate derivative modeling. The LMM's assumption of lognormal dynamics for forward LIBOR rates under their respective forward measures aligns with market practice, particularly in pricing caps, floors, and swaptions using Black-type formulas. Nonetheless, the model's reliance on deterministic volatility structures limits its ability to capture complex strike-dependent features of implied volatility surfaces, such as the observed volatility smile and skew in swaption markets. These limitations spurred the development of stochastic volatility extensions.

Rebonato, McKay, and White (2009) enhanced the LMM by introducing stochastic volatility and providing connections to the SABR model for forward swap rates. The SABR model, characterized by its flexible parametric form, captures the intricacies of implied volatility skew and smile, making it a standard tool for swaption volatility modeling. Rebonato's work offers an approximate mapping between SABR dynamics and modified LMM volatility structures, bridging the gap between practical volatility surface modeling and arbitrage-free interest rate models.

In this work, we build upon this trajectory by integrating rough stochastic volatility into the Forward Market Model (FMM), an arbitrage-free framework tailored to the post-LIBOR era and developed by Lyashenko and Mercurio (2019). We rigorously derive an asymptotic expansion for swaption implied volatility, linking the model to a rough Bergomi-type framework for forward swap rate dynamics. This formulation establishes a direct connection between an HJM-consistent interest rate model and a rough volatility framework for swaption pricing, offering a parsimonious yet precise representation of swaption volatility surfaces, including their persistent skew and term structure.

Rough volatility models, in which volatility has lower Hölder regularity than Brownian motions, are known to explain a power-law type term structure of the volatility skew (see, e.g., Fukasawa (2021)). The roughness has a market microstructural foundation El Euch, Fukasawa, and Rosenbaum (2018), which potentially applies to bond markets and hence, forward term rates. Indeed, by a calibration experiment, we demonstrate that a rough volatility model exhibits a better performance than the lognormal SABR model, its non-rough counterpart.

As a byproduct of this work, we provide a theoretically sound and verifiable formulation of the commonly used "freezing" technique in swaption modeling. While the freezing technique is traditionally applied as a heuristic simplification -

assuming certain variables such as forward rates or volatilities remain constant - it lacks formal justification in many practical implementations. By deriving a rigorous mathematical framework, this paper not only validates the freezing approximation in specific cases but also provides alternative version that works in general. This advancement offers a more robust foundation for practical swaption pricing, bridging empirical convenience with theoretical consistency.

In Section 2, we introduce the basic notation and assumptions, followed by a key lemma that forms the foundation of the freezing technique. Section 3 reviews the HJM framework and presents a rough SABR forward market model. In Section 4, we derive a short-term asymptotic expansion of the swaption implied volatility under this model and construct a forward swap rate model that produces the same expansion. All proofs are provided in Section 5, while Section 6 reports a calibration experiment assessing the effectiveness of the proposed models.

2 Forward term rates and a forward swap rate

2.1 Notation

Here we introduce notation and basic assumptions. Let $M = \{M_t\}$ denote a money market account process

$$M_t = \exp\left\{\int_0^t r_s \mathrm{d}s\right\}, \tag{1}$$

where $r = \{r_t\}$ is a piecewise continuous adapted process, called a spot interest rate, defined on a filtered probability space $(\Omega, \mathscr{F}, \mathsf{P}, \{\mathscr{F}_t\}_{t \in [0,\infty)})$. We assume $\mathscr{F}_0$ to consist of the events with probability 0 or 1. Let $P_t(T)$ denote the price of the zero coupon bond with face value 1 and maturity T at time $t \le T$. We extend $P_t(T)$ for $t \ge T$ by

$$P_t(T) = \exp\left\{\int_T^t r_s \,\mathrm{d}s\right\} = \frac{M_t}{M_T}. \tag{2}$$

In light of the fundamental theorem of asset pricing, we assume that there exists an equivalent probability measure Q, called a risk neutral measure, such that

$$\mathsf{E}_\mathsf{Q}\left[\frac{P_t(u)}{M_t}\bigg|\mathscr{F}_s\right] = \frac{P_{s \wedge t}(u)}{M_{s \wedge t}} \tag{3}$$

for all $s, t, u \in [0, \infty)$, or equivalently, for all $s, t \in [0, \infty)$,

$$P_s(t) = \mathsf{E}_\mathsf{Q}\left[\frac{M_s}{M_t}\bigg|\mathscr{F}_s\right] = \mathsf{E}_\mathsf{Q}\left[\exp\left(-\int_s^t r_u \,\mathrm{d}u\right)\bigg|\mathscr{F}_s\right]. \tag{4}$$

We also assume that $P(T) = \{P_t(T)\}$ is a continuous process for each $T \ge 0$.

For a tenor structure $0 = T_0 < T_1 < T_2 < \cdots < T_N$, we define the forward term rate R^j for the term $(T_{j-1}, T_j]$ by

$$R_t^j = \frac{1}{\theta_j} \left(\frac{P_t(T_{j-1})}{P_t(T_j)} - 1 \right), \quad j = 1, 2, \ldots, N,$$

where $\theta_j = T_j - T_{j-1}$. Remark that $R_t^j = R_{T_j}^j$ for $t \geq T_j$ for each j. For a given pair (I, J) with $1 \leq I < J \leq N$, the forward swap rate $S^{IJ} = \{S_t^{IJ}\}$ for the period $(T_I, T_J]$ is defined by

$$S_t^{IJ} = \frac{P_t(T_I) - P_t(T_J)}{A_t^{IJ}}, \quad A_t^{IJ} = \sum_{j=I+1}^{J} \theta_j P_t(T_j). \tag{5}$$

Since $P(T)$ is continuous and adapted, so are R^j, A^{IJ} and S^{IJ}. Except in Section 6, the pair (I, J) is fixed. Therefore, we ease notation by setting $S = S^{IJ}$ and $A = A^{IJ}$ in the sequel.

For $T \geq 0$, the T-forward measure Q^T is defined by

$$\mathsf{Q}^T(F) = \frac{1}{P_0(T)} \mathsf{E}_{\mathsf{Q}} \left[\frac{1}{M_T}; F \right], \quad F \in \mathscr{F}.$$

For a process X, X/M is martingale under Q if and only if $X/P(T)$ is a martingale under Q^T. In particular, R^j is a martingale under Q^{T_j} for each j. By the fundamental theorem of asset pricing, a no arbitrage price of an European payoff $f(R_T^j)$ with maturity T at time t is given by

$$M_t \mathsf{E}_{\mathsf{Q}} \left[\frac{f(R_T^j)}{M_T} \middle| \mathscr{F}_t \right] = P_t(T) \mathsf{E}_{\mathsf{Q}^T} [f(R_T^j) | \mathscr{F}_t].$$

In particular, the unique no arbitrage price at time t for the payoff $R_{T_j}^j - K$ at T_j is

$$M_t \mathsf{E}_{\mathsf{Q}} \left[\frac{R_{T_j}^j - K}{M_T} \middle| \mathscr{F}_t \right] = P_t(T_j)(R_t^j - K)$$

for any constant K and $t \geq 0$. The unique no arbitrage price at time t of a backward-looking fixed-floating payer's swap with a fixed rate K, where the holder receives $\theta_j R_{T_j}^j$ and pay $\theta_j K$ at T_j, $j = I + 1, \ldots, J$, is therefore

$$\sum_{j=I+1}^{J} \theta_j P_t(T_j)(R_t^j - K) = P_t(T_I) - P_t(T_J) - K \sum_{j=I+1}^{J} \theta_j P_t(T_j) = (S_t - K)A_t$$

which is zero if and only if $K = S_t$ defined by (5). This means that the forward swap rate S is the unique fair rate for this swap contract.

The forward swap measure Q^* is defined by

$$Q^*(F) = \frac{1}{A_0} E_Q \left[\frac{A_{T_N}}{M_{T_N}}; F \right], \quad F \in \mathcal{F},$$

where A is defined by (5). Note that A/M is a martingale under Q and that $A_t/M_t = A_{T_N}/M_{T_N}$ for $t \geq T_N$. For a process X, X/M is martingale under Q if and only if X/A is a martingale under Q^*. In particular, S is a martingale under Q^*. A no arbitrage price at time t of the swaption with strike K and maturity T is given by

$$M_t E_Q \left[\frac{(S_T - K)_+ A_T}{M_T} \middle| \mathcal{F}_t \right] = A_t E^* [(S_T - K)_+ | \mathcal{F}_t],$$

where E^* denotes the expectation under Q^*.

2.2 A key lemma

Now, we are ready to state a key lemma for this paper which is seemingly new and important in itself. The proof is given in Section 5.1.

Lemma 1 *Let R^{j*} denote the local martingale part of the forward term rate R^j under Q^* for each $j = I + 1, \ldots, J$. Then,*

$$dS = \sum_{j=I+1}^{J} \Pi^j dR^{j*}, \tag{6}$$

where

$$\Pi^j = \frac{\theta_j P(T_j)}{A P(T_{j-1})} \left(P(T_J) + S \sum_{k=j}^{J} \theta_k P(T_k) \right).$$

Moreover, if $R^j \geq 0$ for all $j = I + 1, \ldots, J$, then $\Pi^j R^j \leq S$ for all $j = I + 1, \ldots, J$. Further, if $R^j > 0$ for all $j = I + 1, \ldots, J$, then

$$d \frac{\Pi^j R^j}{S} = \sum_{i=I+1}^{J} X^{ji} \frac{dR^i}{R^i} + \sum_{i,k=I+1}^{J} Y^{jik} d \langle \log R^i, \log R^k \rangle \tag{7}$$

for some bounded processes X^{ji} and Y^{jik}.

Remark 1 Writing

$$\Pi^j = \frac{\theta_j P(T_j)}{A} C^j,$$

we have a recursion formula for C^j:

$$C^{I+1} = 1, \quad C^{j+1} = (1 + \theta_j R^j) C^j - \theta_j S.$$

From this we observe that $C^j = 1$ for all j if and only if $R^j = S$ for all j.

An important implication from Lemma 1 is that

$$\langle S \rangle_t = \sum_{i,j=I+1}^{J} \int_0^t \Pi_s^i \Pi_s^j \, d\langle R^i, R^j \rangle_s \approx \sum_{i,j=I+1}^{J} \Pi_0^i \Pi_0^j \langle R^i, R^j \rangle_t$$

for small $t \geq 0$, which in turn verifies an approximation of the swap rate by a basket of forward term rates:

$$S_t \approx S_0 + \sum_{j=I+1}^{J} \Pi_0^j (R_t^j - R_0^j).$$

Notice that this coincides with a well-known freezing approximation

$$S_t = \sum_{j=I+1}^{J} \frac{\theta_j P_t(T_j)}{A_t} R_t^j \approx \sum_{j=I+1}^{J} \frac{\theta_j P_0(T_j)}{A_0} R_t^j$$

only when $R_0^j = S_0$ for all j by Remark 1. The freezing approximation is a common market and academic practice of which the validity has been unclear; see Gaspar and Pimentel (2017) and the references therein for details. On one hand, Lemma 1 justifies this approximation for the very special case of $R_0^j = S_0$, that is, when the interest rate term structure at time 0 is flat. On the other hand, Lemma 1 suggests a better approximation which works in general case.

3 A Forward Market Model with rough volatility

3.1 The HJM model

We begin with recalling the Heath-Jarrow-Morton (HJM) framework. The HJM model assumes

$$P_t(T) = \exp\left\{ -\int_t^T f_t(s) \, ds \right\} \tag{8}$$

for all t and T with a family of continuous adapted processes $\{f_t(s)\}$ of the form

$$f_t(s) = f_0(s) + \int_0^t \mu_u(s) \, du + \int_0^t \sigma_u(s)^\top dW_u$$

called the forward rate, where W is an n-dimensional Brownian motion under Q and $\{\mu_u(s)\}$ and $\{\sigma_u(s)\}$ are integrable processes in suitable senses. By (2), we have $f_t(s) = r_s$ for $t \geq s$. This requires $\mu_u(s) = 0$ and $\sigma_u(s) = 0$ for $u > s$.

Further, to be consistent with the martingale property of $P(T)/M$ under Q, we have

$$\frac{P_t(T)}{M_t} = P_0(T) \exp\left\{-\frac{1}{2}\int_0^t |\Sigma_u(T)|^2 \,\mathrm{d}u - \int_0^t \Sigma_u(T)^\top \mathrm{d}W_u\right\} \tag{9}$$

for all $T \geq 0$, where

$$\Sigma_u(T) = \int_0^T \sigma_u(s) \,\mathrm{d}s = \int_u^T \sigma_u(s) \,\mathrm{d}s,$$

or equivalently,

$$\mu_u(s) = \sigma_u(s)^\top \int_0^s \sigma_u(v) \mathrm{d}v. \tag{10}$$

From (9) and (33) in Appendix A, we derive the dynamics of forward term rates

$$\mathrm{d}R_t^j = \frac{1 + \theta_j R_t^j}{\theta_j} \left(\Sigma_t(T_j) - \Sigma_t(T_{j-1})\right)^\top \mathrm{d}W_t + \sum_{i=1}^j \frac{\theta_i}{1 + \theta_i R_t^i} \mathrm{d}\langle R^i, R^j\rangle_t \tag{11}$$

for $j = 1, \ldots, N$ under the HJM model.

Under the particular structure

$$\sigma_u(s) = Z_u g(s) 1_{[0,s]}(u) \tag{12}$$

with an $n \times m$ matrix-valued process Z and an $\mathbb{R}^m$-valued deterministic continuous function g for the HJM model (10), we have

$$\Sigma_u(T) = Z_u(G(u \vee T) - G(u)), \quad G(u) = \int_0^u g(v) \,\mathrm{d}v$$

and

$$\begin{aligned}
f_t(s) &= f_0(s) + \int_0^t \sigma_u(s)^\top \int_0^s \sigma_u(v) \,\mathrm{d}v \,\mathrm{d}u + \int_0^t \sigma_u(s)^\top \mathrm{d}W_u \\
&= f_0(s) + g(s)^\top \int_0^{t \wedge s} Z_u^\top Z_u \int_u^s g(v) \,\mathrm{d}v \,\mathrm{d}u + g(s)^\top \int_0^{t \wedge s} Z_u^\top \mathrm{d}W_u \\
&= f_0(s) + g(s)^\top \int_0^{t \wedge s} \mathrm{d}H_u(G(s) - G(u)) + g(s)^\top \int_0^{t \wedge s} Z_u^\top \mathrm{d}W_u \\
&= f_0(s) + g(s)^\top H_{t \wedge s}(G(s) - G(t \wedge s)) + g(s)^\top F_{t \wedge s},
\end{aligned}$$

where

$$H_t = \int_0^t Z_u^\top Z_u \mathrm{d}u, \quad F_t = \int_0^t H_u g(u) \mathrm{d}u + \int_0^t Z_u^\top \mathrm{d}W_u.$$

In particular,

$$r_t = f_t(t) = f_0(t) + g(t)^\top F_t$$

and for $t \leq T$,

$$P_t(T) = \frac{P_0(T)}{P_0(t)} \exp\left\{ -\int_t^T g(s)^\top H_t(G(s) - G(t)) \, ds - (G(T) - G(t))^\top F_t \right\}$$

by (9).

3.2 The Forward Market Model

Here we describe a forward market model following Lyashenko and Mercurio (2019). A forward market model assumes (12) with $m = n = N$ and

$$g(s) = (g_1(s), \ldots, g_N(s))^\top, \quad g_j(s) = g_j(s) 1_{(T_{j-1}, T_j]}(s), \quad j = 1, \ldots, N.$$

In this case, only the jth element of $G(T_j \vee t) - G(T_{j-1} \vee t)$ is nonzero. Therefore, putting

$$\gamma_j(t) = \int_{T_{j-1} \vee t}^{T_j \vee t} g_j(s) \, ds, \quad \sigma_t^j = \frac{1 + \theta_j R_t^j}{\theta_j} Z_t^j, \quad j = 1, \ldots, N,$$

where Z^j is the jth column vector of Z, by (11), we have

$$dR_t^j = \gamma_j(t)(\sigma_t^j)^\top \left[dW_t + \sum_{i=1}^j \frac{\theta_i}{1 + \theta_i R_t^i} \gamma_i(t) \sigma_t^i dt \right], \quad j = 1, \ldots, N.$$

Note that $\gamma_j(t)$ is constant until $t \leq T_{j-1}$ and $\gamma_j(t) = 0$ for $t \geq T_j$. In terms of

$$\tilde{W}_t^j := \int_0^t \frac{1}{|\sigma_s^j|} (\sigma_s^j)^\top dW_s = \int_0^t \frac{1}{|Z_s^j|} (Z_s^j)^\top dW_s, \quad j = 1, \ldots, N,$$

we have

$$\langle \tilde{W}^j \rangle_t = t, \quad \langle \tilde{W}^i, \tilde{W}^j \rangle_t = \int_0^t \rho_s^{ij} \, ds, \quad \rho_t^{ij} := \frac{1}{|\sigma_t^i||\sigma_t^j|} (\sigma_t^i)^\top \sigma_t^j$$

and

$$dR_t^j = \gamma_j(t)|\sigma_t^j| \left[d\tilde{W}_t^j + \sum_{i=1}^j \frac{\theta_i}{1 + \theta_i R_t^i} \gamma_i(t)|\sigma_t^i|\rho_t^{ij} dt \right], \quad j = 1, \ldots, N. \quad (13)$$

The simplest among the forward market models is achieved by

$$g_j(s) = 1_{(T_{j-1}, T_j]}(s), \quad Z_t^j = \frac{\theta_j R_t^j}{1 + \theta_j R_t^j} \varphi_j, \quad j = 1, \ldots, N$$

with constant vectors $\varphi_j \in \mathbb{R}^N$. Indeed, R^j follows a time-inhomogeneous Black-Scholes model under Q^{T_j} with volatility $\gamma_j(t)|\varphi_j|$, where

$$\gamma_j(t) = \begin{cases} 1 & t \leq T_{j-1} \\ (T_j - t)/(T_j - T_{j-1}), & t \in (T_{j-1}, T_j], \\ 0 & t > T_j. \end{cases} \tag{14}$$

3.3 A rough SABR forward market model

To accommodate a volatility skew for each caplet market, we propose a local-stochastic volatility model of the form

$$\mathrm{d}R_t^j = \gamma_j(t)\eta_j(R_t^j)\sqrt{V_t^j}\mathrm{d}W_t^{j\dagger},$$
$$\log V_t^j = \log \xi_j(t) - \frac{1}{2}\int_0^t \zeta_j(t-s)^2 \mathrm{d}s + \int_0^t \zeta_j(t-s)\mathrm{d}\bar{W}_s, \tag{15}$$

where

$$W^{j\dagger} = W^j + \sum_{i=1}^{j}\int_0^{\cdot} \frac{\theta_i}{1+\theta_i R_t^i}\gamma_i(t)\eta_i(R_t^i)\sqrt{V_t^i}\rho_{ij}\mathrm{d}t,$$

$$\bar{W} = W^0 + \sum_{i=1}^{N}\int_0^{\cdot} \frac{\theta_i}{1+\theta_i R_t^i}\gamma_i(t)\eta_i(R_t^i)\sqrt{V_t^i}\rho_{i0}\mathrm{d}t,$$

$(W^0, W^1, \ldots, W^N)$ is a correlated Brownian motion under Q with

$$\langle W^j \rangle_t = t, \quad \langle W^i, W^j \rangle_t = \rho_{ij}t \tag{16}$$

for a constant correlation matrix $[\rho_{ij}]$, γ_j is defined by (14), and ξ_j, η_j and ζ_j are deterministic positive continuous functions on their respective domains. For ζ_j, the domain is $(0, \infty)$ and the integrability

$$\int_0^t \zeta_j(s)^2 \, \mathrm{d}s < \infty$$

is assumed for all $t > 0$. The well-posedness of the stochastic differential equation (15) under a suitable condition is proved in Appendix C.

This model is consistent to the HJM framework. Indeed, it is derived from (12) by taking $m = n = N$ and

$$g_j(s) = 1_{(T_{j-1}, T_j]}(s), \quad Z_t^j = \frac{\theta_j \eta_j(R_t^j)}{1+\theta_j R_t^j}\sqrt{V_t^j}\varphi_j, \quad j = 1, \ldots, N,$$

where $\varphi_j \in \mathbb{R}^N$ are such constant vectors that

$$(\varphi_i)^\top \varphi_j = \rho_{ij}, \quad i, j = 1, \dots, N.$$

For any $[\rho_{ij}]$, such $[\varphi_1, \dots, \varphi_N]$ is obtained by the Cholesky decomposition. Here, W^j corresponds to $\tilde{W}^j$ in (13) for each j. The volatility processes V^j are driven by a common Brownian motion W^0 which is correlated to W^j according to (16).

In order to clarify some advantages of the model, let us first consider a special case of model:

$$\log \xi_j(t) = 2\log \alpha_j + \frac{1}{4}\kappa_j^2 t, \quad \eta_j(r) = |r + \delta_j|^{\beta_j}, \quad \zeta_j(t) = \kappa_j$$

for some positive constants $\alpha_j, \beta_j, \delta_j$, and κ_j. Then, the model reduces to

$$dR_t^j = \gamma_j(t)|R_t^j + \delta_j|^{\beta_j}\sqrt{V_t^j}dW_t^{j\dagger},$$

$$d\sqrt{V_t^j} = \kappa_j\sqrt{V_t^j}d\bar{W}_t, \quad \sqrt{V_0^j} = \alpha_j,$$

which is a time-inhomogeneous shifted SABR model, widely used in practice.

While these Markovian SABR models successfully recover the volatility smile of each caplet and its dynamics, the common volatility factor W^0 for all j is not considered to work sufficiently for the joint calibration for swaptions. The existing approach is to introduce respective Brownian motions as the driving factors of the volatility processes, which amounts to introducing a large number of parameters describing the correlation among those Brownian motions. Practical challenges arise from not only the large number of the parameters to be calibrated but also the constraint on the parameters for describing a correlation matrix.

Our approach is to add flexibility by going beyond the finite-dimensional Markovian framework with parsimonious parameters. The deterministic functions ξ_j are introduced to endow the model with an infinite dimensional Markov property which ensures time-consistent nature of the model; see Fukasawa, Horvath, and Tankov (2023) for details.

We motivate the drift structure of $\bar{W}$. By (33) in Appendix A,

$$\sum_{i=1}^j \int_0^\cdot \frac{\theta_i}{1 + \theta_i R_t^i}\gamma_i(t)\eta_i(R_t^i)\sqrt{V_t^i}\rho_{ij}dt = -d\left\langle W^j, \log\frac{P(T_j)}{M}\right\rangle,$$

$$\sum_{i=1}^j \int_0^\cdot \frac{\theta_i}{1 + \theta_i R_t^i}\gamma_i(t)\eta_i(R_t^i)\sqrt{V_t^i}\rho_{i0}dt = -d\left\langle W^0, \log\frac{P(T_j)}{M}\right\rangle$$

for each $j = 1, \dots, N$ and so, $(W^{j\dagger}, \bar{W}^{j\dagger})$ is a correlated two dimensional Brownian motion under $\mathbb{Q}^{T_j}$ by the Girsanov-Maruyama theorem, where

$$\bar{W}^{j\dagger} = W^0 + \sum_{i=1}^{j} \int_0^{\cdot} \frac{\theta_i}{1 + \theta_i R_t^i} \gamma_i(t) \eta_i(R_t^i) \sqrt{V_t^i} \rho_{i0} dt$$

$$= \bar{W} - \sum_{i=j+1}^{N} \int_0^{\cdot} \frac{\theta_i}{1 + \theta_i R_t^i} \gamma_i(t) \eta_i(R_t^i) \sqrt{V_t^i} \rho_{i0} dt.$$

==This implies that the drift of $\bar{W}$ under Q^{T_j} is nonpositive if all ρ_{i0} are nonpositive==, which enable us to show the following result, meaning that for each j, under Q^{T_j}, R^j follows an extension of a rough SABR model Fukasawa (2022); Fukasawa and Gatheral (2022); Fukasawa et al. (2023).

Proposition 1 *If $(R^1, \ldots, R^N)$ solves (15) with $R_0^i > 0$,*

$$\sup_{r>0} \frac{\eta_i(r)}{r} < \infty \tag{17}$$

and

$$\rho_{0i} \le 0 \tag{18}$$

for all $i = 1, \ldots, N$, then

$$R_t^j = R_0^j \exp\left(\int_0^t \frac{\eta(R_s^j)}{R_s^j} \gamma_j(s) \sqrt{V_s^j} \, dW_s^{j\dagger} - \frac{1}{2} \int_0^t \left| \frac{\eta(R_s^j)}{R_s^j} \gamma_j(s) \right|^2 V_s^j \, ds \right),$$

$$V_t^j = \xi_j(t) \exp\left(\int_0^t \zeta_j(t-s) \, d\bar{W}_s - \frac{1}{2} \int_0^t \zeta_j(t-s)^2 \, ds \right) \tag{19}$$

and R^j is a martingale under Q^{T_j} for each $j = 1, \ldots, N$.

The proof is given in Section 5.2.

4 A rough Bergomi forward swap model

4.1 An asymptotic expansion of the implied volatility

The swaption implied volatility $\sigma(k, t)$ at time 0 is defined as the unique volatility parameter that equates the Black-Scholes put price with the put option price written on the forward swap rate S:

$$p_{\text{BS}}\left(k, \sigma(k, t) \sqrt{t} \right) = E^* \left[\left(e^k - \frac{S_t}{S_0} \right)_+ \bigg| \mathscr{F}_0 \right],$$

where

$$p_{\mathrm{BS}}(k,v) = e^k \Phi(-d_-(k,v)) - \Phi(-d_+(k,v)), \quad d_{\pm}(k,v) = -\frac{k}{v} \pm \frac{v}{2}.$$

To derive an asymptotic formula for $\sigma(k,t)$, we focus on a simplified version of (15). Namely, we assume for all i,

$$\zeta_i(t) = \kappa t^{H-1/2} \tag{20}$$

with $\kappa > 0$ and $H \in (0, 1/2)$ and that η_i are nonnegative C^2 functions with

$$\sup_{r>0} \frac{\eta_i(r)}{r} + \sup_{r>0} |\eta_i'(r)| + \sup_{r>0} r|\eta_i''(r)| < \infty. \tag{21}$$

Put

$$v(t) = \sum_{i,j=I+1}^{J} \rho_{ij} \pi_i \pi_j \sqrt{\xi_i(t)\xi_j(t)}, \quad \pi_j = \frac{\Pi_0^j \eta_j(R_0^j)}{S_0}. \tag{22}$$

Theorem 1 *If* $(R^1, \ldots, R^N)$ *solves* (15) *with* $R_0^i > 0$, (18), (20), *and* (21), *then*

$$\sigma(k,t) = \sqrt{\bar{v}(t)} \left(1 + \psi k t^{H-1/2}\right) + o(t^H)$$

uniformly in $k \in \{k \in \mathbb{R}; |k| \leq a\sqrt{t}\}$ *as* $t \to 0$ *for any* $a > 0$, *where*

$$\bar{v}(t) = \int_0^1 v(ts)\, ds,$$

$$\psi = \frac{\kappa}{(2H+1)(H+3/2)v(0)} \sum_{j=I+1}^{J} \rho_{0j} \pi_j \sqrt{\xi_j(0)}.$$

The proof is given in Section 5.3

4.2 A forward swap model with consistent asymptotics

Here we propose an approximate model for the forward swap rate S under (15) with (20), (21) and (18). Define S^* by

$$\frac{dS_t^*}{S_t^*} = \sqrt{V_t}\, dW_t^*,$$
$$V_t = v(t) \exp\left(\int_0^t \zeta(t-s)\, dW_t^{0*} - \frac{1}{2} \int_0^t \zeta(t-s)^2\, ds\right) \tag{23}$$

with $S_0^* = S_0$, where $\zeta(t) = \kappa t^{H-1/2}$ and (W^{0*}, W^*) is a correlated Brownian motion under Q^* such that

$$\langle W^{0*}, W^* \rangle_t = \rho t, \quad \rho = \frac{1}{\sqrt{v(0)}} \sum_{j=I+1}^{J} \rho_{0j} \pi_j \sqrt{\xi_j(0)}. \tag{24}$$

Such a Brownian motion exists; let W^{0*} be the local martingale part of W^0 under Q^* and

$$W^* = \frac{1}{\sqrt{v(0)}} \sum_{j=I+1}^{J} \pi_j \sqrt{\xi_j(0)} W^{j*},$$

where W^{j*} is the local martingale part of $W^{j\dagger}$ under Q^*. This means that $|\rho| \leq 1$.

This is an approximative model in the sense that the implied volatility under S^* has the same short-time asymptotics as $\sigma(k, t)$. To be precise, let $\sigma^*(k, t)$ denote the implied volatility, that is, the solution to

$$p_{\mathrm{BS}} \left(k, \sigma^*(k, t) \sqrt{t} \right) = E^* \left[\left(e^k - \frac{S_t^*}{S_0^*} \right)_+ \Big| \mathcal{F}_0 \right]$$

for $k \in \mathbb{R}$ and $t \geq 0$. Noting that S^* follows a rough Bergomi model under Q^*, we deduce the next result by combining Theorem 1 above, Theorem 2.1 and Remark 2.5 of Fukasawa (2021).

Theorem 2 *If* $(R^1, \ldots, R^N)$ *solves* (15) *with* $R_0^i > 0$, (18), (20), *and* (21), *then*

$$\sigma^*(k, t) = \sigma(k, t) + o(t^H)$$

uniformly in $k \in \{k \in \mathbb{R}; |k| \leq a\sqrt{t}\}$ *as* $t \to 0$ *for any* $a > 0$.

Remark 2 The forward swap rate model (23) depends on $[\rho_{ij}]_{1 \leq i, j \leq N}$ only via $v(t)$ defined by (22).

Remark 3 Rebonato et al. (2009) proposed a similar but different mapping between the correlation parameters of their SABR-LMM model and an approximate forward swap rate model. While the derivation was heuristic in Rebonato et al. (2009), our mapping (24) is justified by Theorem 2.

5 Proofs

5.1 Proof of Lemma 1

We write F to represent a process of finite variation, which may change line by line. Since S is a martingale under Q^*, it suffices to show

$$dS = \sum_{j=I+1}^{J} \Pi^j \, dR^j + dF. \tag{25}$$

We put $P^j = P(T_j)$ for brevity. Since

$$\frac{P^I}{A} = \frac{P^J}{A} + S,$$

using Itô's formula, we obtain

$$
\begin{aligned}
\mathrm{d}S &= \frac{\mathrm{d}P^I}{A} - \frac{\mathrm{d}P^J}{A} - S\frac{\mathrm{d}A}{A} + \mathrm{d}F \\
&= \frac{P^J}{A}\left(\frac{\mathrm{d}P^I}{P^I} - \frac{\mathrm{d}P^J}{P^J}\right) + S\left(\frac{\mathrm{d}P^I}{P^I} - \frac{\mathrm{d}A}{A}\right) + \mathrm{d}F \\
&= \frac{P^J}{A}\left(\frac{\mathrm{d}P^I}{P^I} - \frac{\mathrm{d}P^J}{P^J}\right) + \frac{S}{A}\sum_{k=I+1}^{J}\theta_k P^k\left(\frac{\mathrm{d}P^I}{P^I} - \frac{\mathrm{d}P^k}{P^k}\right) + \mathrm{d}F \\
&= \frac{P^J}{A}\sum_{j=I+1}^{J}\left(\frac{\mathrm{d}P^{j-1}}{P^{j-1}} - \frac{\mathrm{d}P^j}{P^j}\right) + \frac{S}{A}\sum_{k=I+1}^{J}\theta_k P^k\sum_{j=I+1}^{k}\left(\frac{\mathrm{d}P^{j-1}}{P^{j-1}} - \frac{\mathrm{d}P^j}{P^j}\right) + \mathrm{d}F \\
&= \sum_{j=I+1}^{J}\left(\frac{P^J}{A} + \frac{S}{A}\sum_{k=j}^{J}\theta_k P^k\right)\left(\frac{\mathrm{d}P^{j-1}}{P^{j-1}} - \frac{\mathrm{d}P^j}{P^j}\right) + \mathrm{d}F.
\end{aligned}
$$

Then, (25) follows from another application of Itô's formula:

$$\mathrm{d}R^j = \frac{P^{j-1}}{\theta_j P^j}\left(\frac{\mathrm{d}P^{j-1}}{P^{j-1}} - \frac{\mathrm{d}P^j}{Pj}\right) + \mathrm{d}F.$$

Since

$$
\begin{aligned}
\frac{R^j \Pi^j}{S} &= \frac{1}{SA}\frac{\theta_j R^j}{1 + \theta_j R^j}\left(P^J + S\sum_{k=j}^{J}\theta_k P^k\right) \\
&= \frac{\theta_j R^j}{1 + \theta_j R^j}\left(\frac{P^J}{P^I - P^J} + \frac{1}{A}\sum_{k=j}^{J}\theta_k P^k\right) \\
&= \frac{\theta_j R^j}{1 + \theta_j R^j}\left(\frac{1}{\prod_{i=I+1}^{J}(1 + \theta_i R^i) - 1} + \frac{1}{A}\sum_{k=j}^{J}\theta_k P^k\right),
\end{aligned}
\tag{26}
$$

if all R^j are nonnegative, then we have

$$\frac{R^j \Pi^j}{S} \le \frac{\theta_j R^j}{1 + \theta_j R^j}\left(\frac{1}{1 + \theta_j R^j - 1} + 1\right) = 1.$$

Now, suppose $R^j > 0$ for all j. Note that

$$\frac{A}{P^J} = \theta_J + \sum_{k=I+1}^{J-1} \theta_k \prod_{i=k+1}^{J} \frac{P^{i-1}}{P^i} = \theta_J + \sum_{j=I+1}^{J-1} \theta_k \prod_{i=k+1}^{J} (1 + \theta_i R^i)$$

so that

$$\frac{1}{A} \sum_{k=j}^{J} \theta_k P^k = \frac{P^J}{A} \sum_{k=j}^{J} \theta_k \frac{P^k}{P^J} = \frac{\theta_J + \sum_{k=j}^{J-1} \theta_k \prod_{i=k+1}^{J}(1 + \theta_i R^i)}{\theta_J + \sum_{k=I+1}^{J-1} \theta_k \prod_{i=k+1}^{J}(1 + \theta_i R^i)}.$$

From this, we observe that

$$\mathrm{d} \log\left(\frac{1}{A} \sum_{k=j}^{J} \theta_k P^k \right) = \sum_{i=I+1}^{J} \hat{X}^{ji} \frac{\mathrm{d} R^i}{R^i} + \sum_{i,k=I+1}^{J} \hat{Y}^{jik} \mathrm{d}\langle \log R^i, \log R^k \rangle \qquad (27)$$

for bounded processes $\hat{X}^{ji}$ and $\hat{Y}^{jik}$. Also,

$$\mathrm{d} \log\left(\prod_{i=I+1}^{J} (1 + \theta_i R^i) - 1 \right) = \sum_{i=I+1}^{J} \tilde{X}^{i} \frac{\mathrm{d} R^i}{R^i} + \sum_{i,k=I+1}^{J} \tilde{Y}^{ik} \mathrm{d}\langle \log R^i, \log R^k \rangle \qquad (28)$$

for bounded processes $\tilde{X}^{i}$ and $\tilde{Y}^{ik}$ since

$$\frac{1}{\prod_{i=I+1}^{J}(1 + \theta_i R^i) - 1} \le \frac{1}{\theta_j R^j \prod_{i=I+1, i \neq j}^{J}(1 + \theta_i R^i)} \qquad (29)$$

for any $j = I + 1, \ldots, J$. By (26),

$$\mathrm{d} \frac{R^j \Pi^j}{S} = \mathrm{d} G^j + \mathrm{d} H^j,$$

where

$$G^j = \frac{\theta_j R^j}{1 + \theta_j R^j} \frac{1}{\prod_{i=I+1}^{J}(1 + \theta_i R^i) - 1}, \quad H^j = \frac{\theta_j R^j}{1 + \theta_j R^j} \frac{1}{A} \sum_{k=j}^{J} \theta_k P^k.$$

By (29), we have

$$0 < G^j \le \frac{1}{\prod_{i=I+1}^{J}(1 + \theta_i R^i)} \le 1,$$

and by (28), we observe that

$$\mathrm{d} G^j = G^j \, \mathrm{d} \log G^j + \frac{1}{2} G^j \, \mathrm{d}\langle \log G^j \rangle$$

$$= \sum_{i=I+1}^{J} \check{X}^{ji} \frac{\mathrm{d} R^i}{R^i} + \sum_{i,k=I+1}^{J} \check{Y}^{jik} \mathrm{d}\langle \log R^i, \log R^k \rangle$$

for bounded processes $\check{X}^{ji}$ and $\check{Y}^{jik}$. Further by (27),

$$
\begin{aligned}
\mathrm{d}H^j &= G^j \, \mathrm{d}\log H^j + \frac{1}{2} H^j \, \mathrm{d}\langle \log H^j \rangle \\
&= \sum_{i=I+1}^{J} \bar{X}^{ji} \frac{\mathrm{d}R^i}{R^i} + \sum_{i,k=I+1}^{J} \bar{Y}^{jik} \mathrm{d}\langle \log R^i, \log R^k \rangle
\end{aligned}
$$

for bounded processes $\bar{X}^{ji}$ and $\bar{Y}^{jik}$.

5.2 Proof of Proposition 1

The representation (19) under (17) follows that the unique solution to the equation

$$
\mathrm{d}R = R\mathrm{d}X
$$

is given by

$$
R = R_0 \exp\left(X - \frac{1}{2}\langle X \rangle \right)
$$

for any continuous semimartingale X. Apply this to

$$
X_t = \int_0^{t \wedge \tau_n} \frac{\eta(R_s^j)}{R_s^j} \gamma_j(s) \sqrt{V_s^j} \, \mathrm{d}W_s^{j\dagger},
$$

where

$$
\tau_n = \inf\left\{ t \geq 0; R_t^j = \frac{1}{n} \right\}
$$

for each $n = 1, 2 \ldots$, and then argue that $\lim \tau_n = \infty$. To prove that R^j is a martingale, we extend the argument of Gassiat (2019). We already know that R^j is a nonnegative local martingale by (19), it suffices to show $\mathsf{E}_{\mathsf{Q}^{T_j}}[R_T^j] = R_0^j$ for any $T \geq 0$. Let

$$
\hat{\tau}_n = \inf\{ t \geq 0; Y_t = n \}, \quad Y_t = \int_0^t \zeta_j(t-s) \, \mathrm{d}\bar{W}_s
$$

and notice that $\{R_{t \wedge \hat{\tau}_n}^j\}$ is a positive martingale under Q^{T_j} for each n by (19). Define an equivalent probability measure $\hat{\mathsf{Q}}_n$ by

$$
\frac{\mathrm{d}\hat{\mathsf{Q}}_n}{\mathrm{d}\mathsf{Q}^{T_j}} = \frac{R_{\hat{\tau}_n}^j}{R_0^j}.
$$

Then,

$$R_0^j = \mathsf{E}_{\mathsf{Q}^{T_j}}[R_{T \wedge \hat{\tau}_n}^j] = \mathsf{E}_{\mathsf{Q}^{T_j}}[R_{\hat{\tau}_n}^j 1_{\{\hat{\tau}_n \le T\}}] + \mathsf{E}_{\mathsf{Q}^{T_j}}[R_T^j 1_{\{\hat{\tau}_n > T\}}]$$

$$= R_0^j \hat{\mathsf{Q}}_n(\hat{\tau}_n \le T) + \mathsf{E}_{\mathsf{Q}^{T_j}}[R_T^j 1_{\{\hat{\tau}_n > T\}}].$$

By the monotone convergence theorem, we have

$$\mathsf{E}_{\mathsf{Q}^{T_j}}[R_T^j 1_{\{\hat{\tau}_n > T\}}] \to \mathsf{E}_{\mathsf{Q}^{T_j}}[R_T^j]$$

as $n \to \infty$. Under $\hat{\mathsf{Q}}_n$, by the Girsanov-Maruyama theorem,

$$\hat{W}^n := \bar{W}^{j\dagger} - \langle \bar{W}^{j\dagger}, \log R^j \rangle_{\cdot \wedge \hat{\tau}_n} = \bar{W} - \sum_{i=j+1}^{N} \int_0^{\cdot} \frac{\theta_i}{1 + \theta_i R_t^i} \gamma_i(t) \eta_i(R_t^i) \sqrt{V_t^i} \rho_{i0} dt$$

$$- \int_0^{\cdot \wedge \hat{\tau}_n} \frac{\eta(R_s^j)}{R_s^j} \gamma_j(s) \sqrt{V_s^j} \rho_{j0} \, ds$$

is a Brownian motion. Under (21), we have

$$Y_t = \int_0^t \zeta_j(t-s) \, d\bar{W}_s \le \int_0^t \zeta_j(t-s) \, d\hat{W}_s^n =: \hat{Y}_t^n.$$

The law of $\hat{Y}^n$ under $\hat{\mathsf{Q}}_n$ does not depend on n and

$$\hat{\mathsf{Q}}_n(\hat{\tau}_n \le T) \le \hat{\mathsf{Q}}_n \left(\sup_{t \le T} \hat{Y}_t^n \ge n \right) \to 0$$

as $n \to \infty$, which concludes the proof.

5.3 Proof of Theorem 1

Put

$$v^*(t) = \mathsf{E}^*[V_t^* | \mathcal{F}_0], \quad V_t^* = \sum_{i,j=I+1}^{J} \hat{\Pi}_t^i \hat{\Pi}_t^j \sqrt{V_t^i} \sqrt{V_u^j} \rho_{ij}, \quad \hat{\Pi}_t^j = \frac{\Pi_t^j \eta_j(R_t^j)}{S_t}.$$

By Lemma 1, we have

$$d\langle \log S \rangle_t = V_t^* dt.$$

Therefore by Theorem 2.1 of Fukasawa (2021), it suffices to show (i):

$$v(t) = v^*(t) + o(t^H),$$

(ii):

$$\frac{1}{t^H} \left(\frac{V_t^*}{v^*(t)} - 1 \right)$$

is uniformly (in t) integrable, and (iii):

$$\left(\frac{1}{\sqrt{t}}\left(\frac{S_t}{S_0} - 1\right), \frac{1}{t^H}\left(\frac{V_t^*}{v^*(t)} - 1\right)\right)$$

converges in law to the normal distribution with mean 0 and covariance $[\Sigma_{ij}]$,

$$\Sigma_{12} = \frac{2}{2H+1} \sum_{j=I+1}^{J} \rho_{0j} \pi_j \sqrt{\xi_j(0)}$$

as $t \to 0$. As a preliminary, first note that by (19) all R^j are positive and so, by Lemma 1, we have that $\Pi^j R^j \leq S$ for all j and (7). Moreover, this also implies that for any $p \geq 1$ and $T > 0$, there exists a constant $C_p > 0$ such that

$$\mathsf{E}^* \left[\langle \log R^j \rangle_t^p\right] \leq C_p t^p.$$

for all j and $t \in [0, T]$, which in turn implies that for any $p \geq 1$ and $T > 0$, there exists a constant $C_p > 0$ such that

$$\mathsf{E}^* \left[\left|\hat{\Pi}_t^j - \pi_j\right|^{2p}\right] \leq C_p t^p \tag{30}$$

for all j and $t \in [0, T]$ by (7) and (21). Second, by (34), we have

$$\bar{W}^* := W^0 + \int_0^\cdot \sum_{i=1}^{J} \left(\sum_{k=(I+1)\vee i}^{J} \frac{\theta_k P(T_k)}{A}\right) \frac{\theta_i}{1 + \theta_i R^i} \mathrm{d}\langle W^0, R^i \rangle$$

is a Brownian motion under Q^*. Note that $\bar{W} = \bar{W}^* + D$ and

$$V_t^j = \xi_j(t) \exp\left(\kappa \int_0^t (t-s)^{H-1/2} \mathrm{d}D_s\right) V_t^0,$$

where

$$D = \sum_{i=J+1}^{N} \int_0^\cdot \frac{\theta_i}{1 + \theta_i R^i} \mathrm{d}\langle W^0, R^i \rangle + \sum_{i=I+2}^{J} \left(\sum_{k=I+1}^{i-1} \frac{\theta_k P(T_k)}{A}\right) \int_0^\cdot \frac{\theta_i}{1 + \theta_i R^i} \mathrm{d}\langle W^0, R^i \rangle$$

and

$$V_t^0 = \exp\left(\kappa \int_0^t (t-s)^{H-1/2} \mathrm{d}\bar{W}_s^* - \frac{\kappa^2 t^{2H}}{4H}\right).$$

Let us show (i). Using $\mathsf{E}^*[V_t^0] = 1$, we have

$$v^*(t) - v(t) = \sum_{i,j=I+1}^{J} \mathsf{E}^* \left[(\hat{\Pi}_t^i \hat{\Pi}_t^j - \pi_i \pi_j) \sqrt{V_u^i} \sqrt{V_u^j} \right] \rho_{ij}$$

$$+ v(t) \mathsf{E}^* \left[\left(\exp \left(\kappa \int_0^t (t-s)^{H-1/2} \mathrm{d}D_s \right) - 1 \right) V_t^0 \right].$$

Under (18),

$$\kappa \int_0^t (t-s)^{H-1/2} \mathrm{d}D_s \leq 0 \tag{31}$$

and so

$$0 \leq \mathsf{E}^* \left[V_t^0 \left(1 - \exp \left(\kappa \int_0^t (t-s)^{H-1/2} \mathrm{d}D_s \right) \right) \right]$$

$$\leq -\mathsf{E}^* \left[V_t^0 \kappa \int_0^t (t-s)^{H-1/2} \mathrm{d}D_s \right] = O(t^{H+1/2}).$$

By (30) and (31), we also have

$$\mathsf{E}^* \left[(\hat{\Pi}_t^i \hat{\Pi}_t^j - \pi_i \pi_j) \sqrt{V_u^i} \sqrt{V_u^j} \right] = O(t^{1/2}),$$

which concludes (i).

Next we show (ii). By the above computation for (i), we observe that

$$\mathsf{E}^* [|V_t^* - v(t) V_t^0|^2] = O(t).$$

Therefore, it suffices to show that $t^{-H}(V_t^0 - 1)$ is uniformly integrable. By a straightforward computation,

$$\mathsf{E}^* [|t^{-H}(V_t^0 - 1)|^2] = t^{-2H} (\mathsf{E}^* [|V_t^0|^2] - 1) = \frac{1}{t^{2H}} \left(\exp \left(\frac{\kappa^2 t^{2H}}{4H} \right) - 1 \right) = O(1),$$

which implies the uniform integrability.

Finally, we show (iii). By Lemma 1 and (30), we have

$$\frac{1}{\sqrt{t}} \left(\frac{S_t}{S_0} - 1 \right) = \sum_{j=I+1}^{J} \pi_j \sqrt{\xi_j(t)} \frac{W_t^{j*}}{\sqrt{t}} + o_P(1),$$

where W^{j*} is the local martingale part of $W^{j\dagger}$ under Q^*. Also,

$$\frac{1}{t^H} \left(\frac{V_t^*}{v^*(t)} - 1 \right) = \frac{1}{t^H} \left(V_t^0 - 1 \right) + o_P(1) = \frac{\kappa}{t^H} \int_0^t (t-s)^{H-1/2} \mathrm{d}\bar{W}_s^* + o_P(1)$$

as seen above. These leading terms are Gaussian with covariance

$$\mathsf{E}^* \left[\left(\sum_{j=I+1}^{J} \pi_j \sqrt{\xi_j(t)} \frac{W_t^{j*}}{\sqrt{t}} \right) \frac{\kappa}{t^H} \int_0^t (t-s)^{H-1/2} \mathrm{d}\bar{W}_s^* \right] = \frac{2}{2H+1} \sum_{j=I+1}^{J} \pi_j \sqrt{\xi_j(t)} \rho_{j0},$$

which concludes (iii).

6 Numerical Analysis

6.1 Market Data and Separate Tenor Calibration

Here we examine the effectiveness of the proposed model by a calibration experiment to market data. We use USD secured overnight financing rate (SOFR) swap rate and SOFR swaption smile data on 9 December 2024 extracted from Bloomberg[1].

The smile data consist of swaption implied volatilities (IVs) for strikes in the set

$$\mathcal{K} = K_{\mathrm{ATM}} + \{-2\%, -1\%, -0.5\%, -0.25\%, 0\%, 0.25\%, 0.5\%, 1\%, 2\%\},$$

where K_{ATM} is the ATM strike, and for some maturities and tenors designated by S in Table 1. For example, we have a 3Y × 5Y smile, that is for a swaption with maturity $t = T^I = 3$ years and $T^J = T^I + 5$ years. For other maturities and tenors,

Table 1 Swaption Dataset (row:maturity, column:underlying tenor)

	1Y	2Y	3Y	4Y	5Y	6Y	7Y	8Y	9Y	10Y
1Y	S	S			S					S
2Y										
3Y	S	S			S					
4Y										
5Y	S	S			S					
6Y										
7Y										
8Y										
9Y										
10Y	S									

say, 1Y × 3Y, we have ATM volatility only. The convention of the obtained swaption volatility is that the underlying swap has annual payment intervals in both legs. So, the year-fraction $\theta_j = T_j - T_{j-1}$ is 1 year for all j.

First, we focus on smiles with 1 year tenor, that is 1Y × 1Y, 3Y × 1Y, 5Y × 1Y, and 10Y × 1Y. The underlying forward swap rates are related each other, but in order to examine whether our rough volatility model provides a reasonable term structure, we pretend as if they are independent and calibrate the rough Bergomi forward swap model (23) to obtain volatility of volatility parameter κ, while H fixed,

separately for each smile. Assuming $E^*[\sqrt{V_t^*}] = \sqrt{V_0^*}$, we set initial variance curve as $v(t) = v(0) \exp(\frac{\kappa^2 t^{2H}}{8H})$. Under this condition, when $H = 0.5$, the model coincides with the log normal SABR model.

The result is given in Table 2. The model IVs under different values of H are calculated by the Monte-Carlo method with 1,000,000 paths, using the hybrid scheme (Bennedsen, Lunde, and Pakkanen (2017)) with the Black-Scholes model paths as control variates. Since the model with $H = 0.5$ is a SABR model, we also present results calibrated using the Hagan formula at the first row.

Discrepancies in the parameter κ are observed when comparing the results between the Hagan formula and Monte Carlo simulation ($H = 0.5$). These discrepancies become more pronounced as the maturity increases. The Hagan formula performs well when $\kappa^2 T$ is small enough. In this swaption data, the estimated values of κ are mostly greater than 1 and deviate significantly from the Monte Carlo results, suggesting that the assumptions underlying the formula may not be satisfied.

We observe from Table 2 that the calibrated values of κ fall into smaller ranges when $H < 0.5$ than when $H = 0.5$. This result suggests the possibility of a parsimonious modeling of the forward swap rate processes with a single value of κ by choosing $H < 0.5$.

Table 2 Model Calibration Results for Different Hurst Exponents

	1Y	3Y	5Y	10Y
Hagan	1.743	1.095	1.001	0.842
H=0.5	1.952	1.328	1.415	3.525
H=0.45	1.821	1.299	1.394	1.938
H=0.4	1.693	1.266	1.372	1.739
H=0.35	1.566	1.230	1.347	1.635
H=0.3	1.442	1.188	1.318	1.568
H=0.25	1.320	1.142	1.282	1.516
H=0.2	1.201	1.091	1.241	1.473
H=0.15	1.083	1.033	1.190	1.417
H=0.1	0.960	0.961	1.116	1.048
H=0.05	0.782	0.835	0.944	0.762

6.2 Calibration of the Proposed Model

We calibrate the proposed model (15) with two steps. As the first-step calibration, we estimate the parameters $\kappa, \alpha_i, \rho_{0i}, i = 1, \cdots, N$ using swaption IVs with underlying of 1Y tenor. Since $\theta_j = 1$, we have $J = I + 1$ for each of such smiles. Therefore, $S^{IJ} = R^J$ and the swaption has the same value as the caplet (1Y tenor swaption) with maturity T_I, Based on this observation, we have calibrated the proposed model (15) with (20) and $\eta_j(r) = r$. We have parameterized the function $\xi_j(t)$ by $\alpha_j > 0$ as

$$\xi_j(t) = \alpha_j^2 \exp\left\{\frac{\kappa^2 t^{2H}}{8H} - \sum_{i=j+1}^{N} \frac{\theta_i \eta_i (R_0^i)}{1 + \theta_i R_0^i} \alpha_i \rho_{0i} \kappa \int_0^t (t-s)^{H-1/2} \gamma_i(s) \mathrm{d}s\right\},$$

where $T_N = 11$ years, so that $\mathsf{E}^{T_j}[\sqrt{V_t^j}] \approx \alpha_j$. Note that

$$\int_0^t (t-s)^{H-1/2} \gamma_i(s) \mathrm{d}s = \frac{t^{H+1/2}}{H+1/2} - \frac{(t-T_{i-1})_+^{H+3/2} - (t-T_i)_+^{H+3/2}}{(H+1/2)(H+3/2)(T_i - T_{i-1})}.$$

Furthermore, we define the model smile discrepancy in the first-step calibration as the sum of squared differences between the market IVs in smile tenors and those generated by the mapped SMM model (23). The parameters κ, ρ_{0i} (for $i = 2, 4, 6, 11$) are estimated by minimizing the model smile discrepancy. Given the parameters κ, ρ_{0i} for $i = 2, 4, 6, 11$, the remaining ρ_{0i} are determined by linear interpolation and flat extrapolation. We determine α_i such that the ATM swaption price computed under the mapped rough Bergomi forward swap model (23) matches the market IV. Swaption prices under the mapped model are computed using the Monte Carlo simulation. Instead of the Monte Carlo, we have also examined the use of the rough SABR formula proposed in Fukasawa and Gatheral (2022); the results are reported in Appendix B.

The first-step calibration is performed with variations in the Hurst exponent H. Table 3 reports the estimated parameters, under each condition, along with the corresponding model smile discrepancy expressed as the root mean squared error (RMSE). The RMSE-minimizing Hurst exponent is $H = 0.2$. These values deviate substantially from $H = 0.5$. This suggests that the assumption of a low Hurst exponent effectively captures the term structure of κ.

Table 3 First-Step Calibration Results

Hurst Exponent	κ	α_2	α_4	α_6	α_{11}	$\rho_{0,2}$	$\rho_{0,4}$	ρ_{06}	$\rho_{0,11}$	RMSE
H = 0.5	1.72065	0.315	0.315	0.343	0.401	-63.9%	-39.8%	-46.7%	-47.0%	0.728%
H = 0.45	1.62013	0.314	0.311	0.330	0.368	-63.3%	-40.1%	-46.9%	-47.0%	0.678%
H = 0.4	1.52652	0.312	0.305	0.318	0.339	-62.6%	-40.4%	-47.3%	-47.4%	0.625%
H = 0.35	1.43774	0.309	0.300	0.307	0.314	-61.9%	-40.8%	-48.1%	-48.4%	0.574%
H = 0.3	1.35083	0.304	0.293	0.296	0.292	-61.4%	-41.5%	-49.2%	-49.9%	0.532%
H = 0.25	1.26378	0.298	0.284	0.283	0.271	-61.2%	-42.4%	-50.8%	-52.2%	0.507%
H = 0.2	1.17658	0.288	0.271	0.268	0.251	-61.5%	-43.8%	-53.2%	-55.3%	0.505%
H = 0.15	1.08537	0.271	0.253	0.248	0.229	-63.3%	-46.5%	-57.1%	-60.6%	0.530%
H = 0.1	0.98173	0.241	0.223	0.218	0.199	-69.7%	-52.6%	-65.6%	-71.3%	0.585%
H = 0.05	0.81767	0.186	0.173	0.169	0.154	-97.1%	-74.2%	-94.2%	-99.8%	0.707%

As the second-step of the calibration, we estimate the correlation matrix among underlying assets R^j. In this paper, we estimate the correlation matrix of $(W_0, W_1, \cdots, W_N)$, namely $(N+1) \times (N+1)$-matrix with indices $i, j = 0, \cdots, N$. We construct the correlation matrix by a parametrized lower triangular matrix B, that is,

$$b_{i,j} = \begin{cases} \cos \omega_{i,j} \prod_{k=0}^{j-1} \sin \omega_{i,k}, & 0 < j < i \\ \prod_{k=0}^{j-1} \sin \omega_{i,k}, & j = i. \end{cases},$$

with $\prod_{k=0}^{-1} \cdot = 1$ for consistency. Then, $\Sigma = BB^{\mathsf{T}}$ satisfies the properties of a correlation matrix (Rapisarda, Brigo, and Mercurio (2007)). To embed the correlation estimated in the first-step calibration, we set $\rho_{i,0} = \cos \omega_{i,0}$ for $i = 1, \cdots, N$. Since the correlation related to W_1 cannot be estimated from the available swaptions, we impose the condition $\mathrm{d}\langle W_1, W_2 \rangle = \mathrm{d}t$ This is achieved by setting $\cos \omega_{2,1} = 1$ and $\cos \omega_{i,2} = 0$ for $i \geq 3$.

To calibrate the correlation matrix, we use ATM swaptions with an underlying tenor of 2 years or longer. Similar to the first-step calibration, the FMM (15) is mapped to rough Bergomi forward swap model (23), and swaption prices are then computed using either Monte Carlo simulation or the rough SABR formula (in Appendix B). Since the mapping only involves correlations of interest rate related to R^i included in the underlying swap, incremental estimation is possible by using co-terminal swaptions. Specifically, we treat the i-th row of $b_{i,j}$, as the search domain and calibrate them by minimizing the sum of squared differences between the model IV and the market IV for the i-Y co-terminal swaptions – i.e., the set of swaptions for which the sum of the maturity and the underlying tenor equals i years. Once the estimation is complete, we increment the row index and repeat the process. We estimate the correlation matrices with H = 0.5 and H = 0.2.

The obtained $\rho_{ij}(i, j = 1, \cdots, N)$ are shown in Table 4, 5. There is a tendency for the correlation between underlying assets with distant maturities to be lower. Additionally, when the Hurst exponent decreases, the correlation matrix tends to be estimated higher.

Table 4 Correlation Matrix Among Interest Rate Factors. (H=0.5)

1.000	0.8981	0.7390	0.7547	0.6308	0.4311	0.3862	0.3669	0.1519	0.4204
0.8981	1.000	0.9060	0.7303	0.6194	0.4596	0.3431	0.3240	0.0400	0.3230
0.7390	0.9060	1.000	0.8733	0.7541	0.6294	0.4746	0.4584	0.06683	0.3049
0.7547	0.7303	0.8733	1.000	0.8592	0.7046	0.6047	0.5897	0.1996	0.3979
0.6308	0.6194	0.7541	0.8592	1.000	0.8671	0.7051	0.7034	0.3427	0.3830
0.4311	0.4596	0.6294	0.7046	0.8671	1.000	0.8729	0.8814	0.6186	0.5210
0.3862	0.3431	0.4746	0.6047	0.7051	0.8729	1.000	0.9993	0.8787	0.8546
0.3669	0.3240	0.4584	0.5897	0.7034	0.8814	0.9993	1.000	0.8833	0.8414
0.1519	0.0400	0.06683	0.1996	0.3427	0.6186	0.8787	0.8833	1.000	0.8904
0.4204	0.3230	0.3049	0.3979	0.3830	0.5210	0.8546	0.8414	0.8904	1.000

Table 5 Correlation Matrix Among Interest Rate Factors. (H=0.2)

1.000	0.9593	0.8681	0.8775	0.8604	0.7131	0.6896	0.7382	0.5948	0.6950
0.9593	1.000	0.9728	0.9453	0.9217	0.8318	0.7761	0.8298	0.6773	0.7761
0.8681	0.9728	1.000	0.9464	0.9196	0.8861	0.8084	0.8621	0.7144	0.8042
0.8775	0.9453	0.9464	1.000	0.9477	0.9024	0.8942	0.9261	0.8168	0.8925
0.8604	0.9217	0.9196	0.9477	1.000	0.9507	0.8967	0.9384	0.8390	0.8905
0.7131	0.8318	0.8861	0.9024	0.9507	1.000	0.9534	0.9779	0.9240	0.9464
0.6896	0.7761	0.8084	0.8942	0.8967	0.9534	1.000	0.9932	0.9879	0.9995
0.7382	0.8298	0.8621	0.9261	0.9384	0.9779	0.9932	1.000	0.9683	0.9909
0.5948	0.6773	0.7144	0.8168	0.8390	0.9240	0.9879	0.9683	1.000	0.9879
0.6950	0.7761	0.8042	0.8925	0.8905	0.9464	0.9995	0.9909	0.9879	1.000

6.3 Validation of Market Fitting for Rough SABR Forward Market Model

We have calibrated the parameters of the proposed model (15) in the preceding subsection. Now we examine whether our approximation to the forward swap rate by (23) works or not, and whether the calibrated model is useful to predict out-of-sample volatility smiles.

For both parameter sets obtained in the previous section, we calculate the swaption IVs for the maturities and underlying where market smile data is available, and compare them with market prices. We plot the following IVs:

- Swaption IVs computed using Monte Carlo simulation under the FMM model "FMM",
- Swaption IVs under our approximative swap rate model (23) using the same calibrated parameters "Mapped SMM",
- Market IVs "Market".

Error bars corresponding to the ±2 standard deviations are shown for both "FMM" and "Mapped SMM".
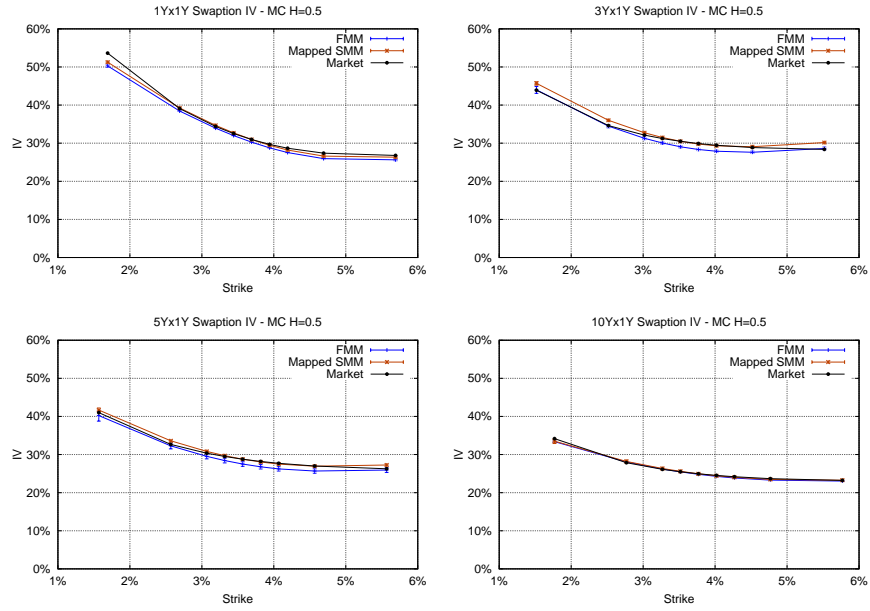
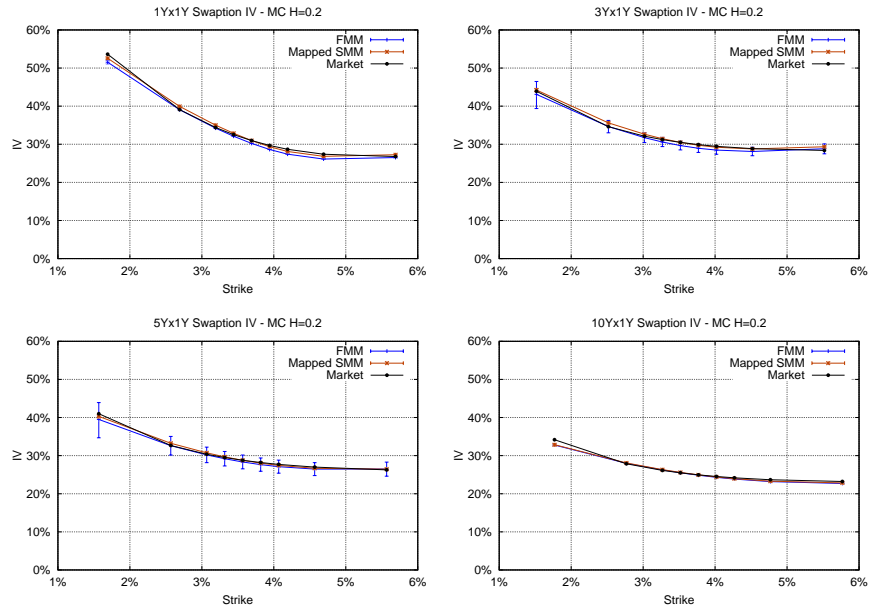**Fig. 1** Model and Market IVs of 1Y tenor at $H = 0.5$

Fig. 2 Model and Market IVs of 1Y tenor at $H = 0.2$

Figure 1, 2 show the quality of in-sample fit, as the calibration was done using swaptions with 1Y tenor. The fit of FMM to Market is almost perfect for $10Y \times 1Y$, while their volatility levels slightly deviate for other maturities. This is explained by our model on the drift; we have introduced negative drift terms of volatility in (15) for $j < N$ to ensure the martingale property of R^j under Q^{T_j}. For $T_N = 11Y$, there is no drift. The drift terms would induce some errors to our calibration of α_j for $j < N$. The approximative swap rate model (Mapped SMM), in spite of its much simpler structure, performs remarkably well to reproduce Market values. Not only the skew is accurate but also the level deviations in FMM are corrected. This would be explained by the fact that there is no drift of volatility in (23).

In Figure 3, 4, we present volatility smiles with tenor more than 2Y. These show the quality of out-of-sample fit, as the calibration was done without using these smile data other than the ATM volatility values.
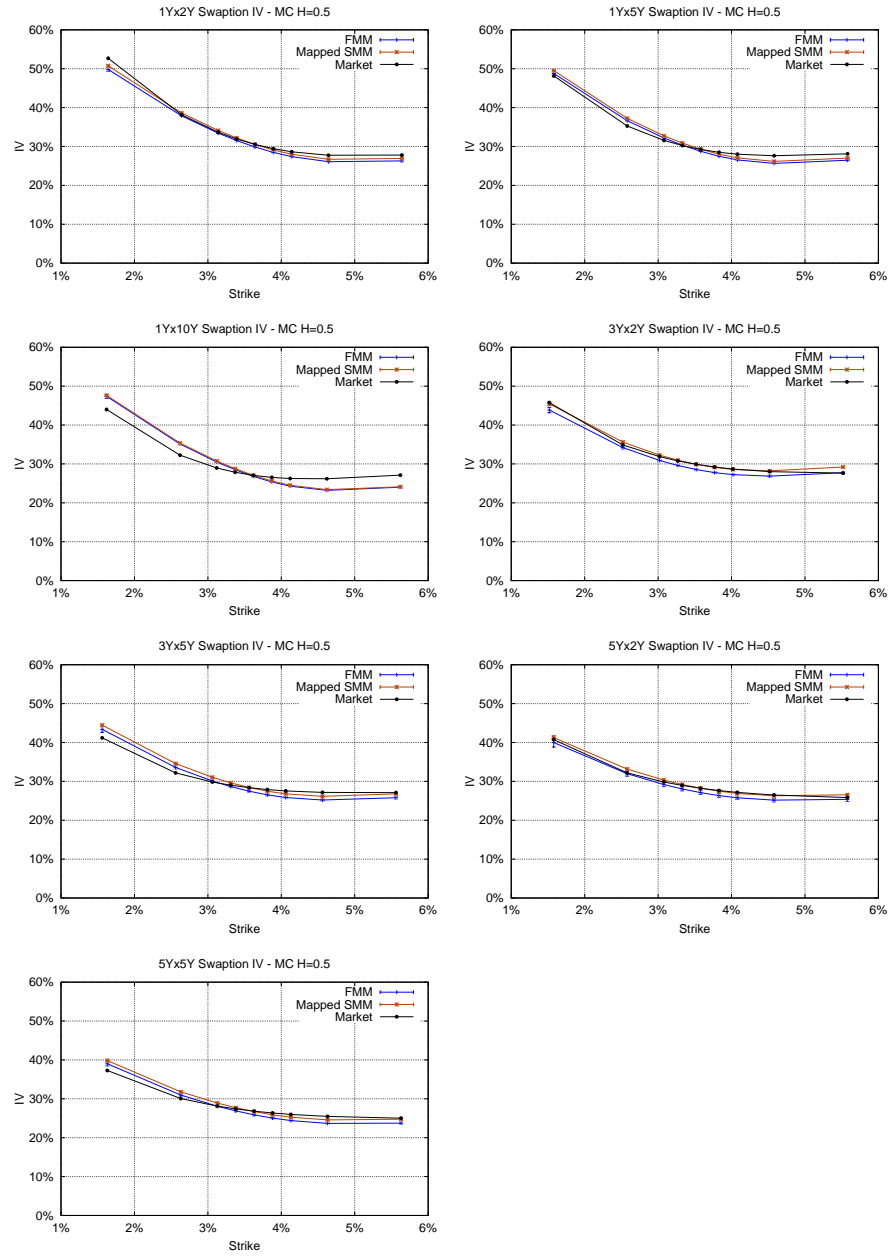
Fig. 3 Model and Market IVs of more than 2Y tenor at $H = 0.5$
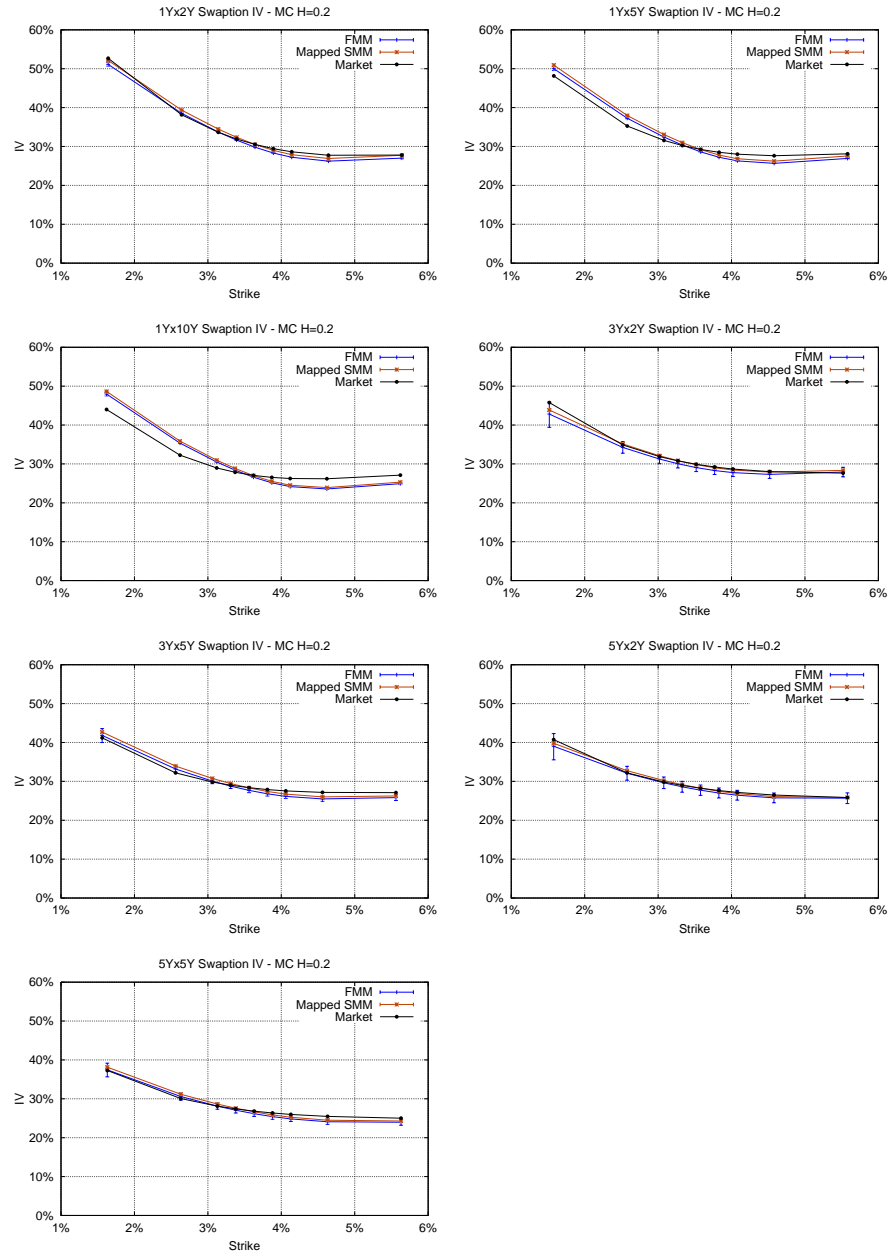
Fig. 4 Model and Market IVs of more than 2Y tenor at $H = 0.2$

First, we notice that FMM and Mapped SMM are close, meaning that our approximative model (23) works well, even though the case $H = 0.5$ is not included in our theoretical justification by Theorem 2.

Second, we observe that these model IVs successfully predict Market values, suggesting that the proposed model (15) that describes the joint distribution of the forward term rate processes is useful to relate the caplet skew (that is, the swaption skew with 1 year tenor) to the swaption skew with different tenor.

Table 6 presents the RMSE values in the swaption volatility fit to compare the performances under $H = 0.5$ and $H = 0.2$. We observe that the values under $H = 0.2$ are superior to those under $H = 0.5$ most of the case, suggesting rough volatility.

Table 6 RMSE Comparison between FMM and Mapped SMM under Different Hurst Exponents

	FMM		Mapped SMM	
	H = 0.5	H = 0.2	H = 0.5	H = 0.2
1Yx1Y	1.42%	1.03%	0.87%	0.61%
3Yx1Y	1.09%	0.71%	1.00%	0.52%
5Yx1Y	1.06%	0.63%	0.55%	0.37%
10Yx1Y	0.34%	0.54%	0.28%	0.49%
1Yx2Y	1.35%	1.02%	0.85%	0.67%
1Yx5Y	1.18%	1.43%	1.15%	1.56%
1Yx10Y	2.27%	2.26%	2.29%	2.40%
3Yx2Y	1.24%	1.20%	0.61%	0.70%
3Yx5Y	1.42%	1.03%	1.50%	1.01%
5Yx2Y	0.98%	0.73%	0.49%	0.42%
5Yx5Y	1.25%	0.82%	1.15%	0.71%

In summary, this calibration experiment suggests that the proposed model (15) for forward term rates and the approximation (23) to forward swap rates effectively reproduce the swaption market volatilities.

7 Conclusion

In this study, we propose the rough SABR Forward Market Model, which incorporates the rough nature of interest rate volatility, with the aim of effectively capturing the term structure of swaption skew. A key theoretical advancement is the derivation of the relationship between forward term rates and swaps over infinitesimal time intervals without relying on approximations. We further develop an approximate skew formula based on the proposed model and construct an approximate forward swap rate model that preserves the original model's asymptotic behavior. This leads to the derivation of parameter interrelationships.

Our numerical analysis and calibration to USD SOFR swaption market data demonstrate the model's practical validity and robustness. Notably, even with limited data, the model successfully reproduces implied volatility skews across a wide range of maturities and tenors. The results indicate that lower Hurst exponent yield

better alignment with market observations. Moreover, the introduced approximation techniques are shown to be accurate and simplify the parameter estimation process.

Overall, this research bridges advanced probabilistic modeling and practical derivative pricing, offering valuable tools and insights for both academics and market practitioners.

Appendix A

Here we describe the drift terms of the forward term rates. Note first that

$$\mathsf{E}_{\mathsf{Q}}\left[\left.\frac{\mathrm{d}\mathsf{Q}^{T}}{\mathrm{d}\mathsf{Q}}\right|\mathscr{F}_{t}\right] = \frac{1}{P_0(T)}\frac{P_t(T)}{M_t},$$

$$\mathsf{E}_{\mathsf{Q}}\left[\left.\frac{\mathrm{d}\mathsf{Q}^{*}}{\mathrm{d}\mathsf{Q}}\right|\mathscr{F}_{t}\right] = \frac{1}{A_0}\frac{A_t}{M_t}$$

for all $t \geq 0$. By the Girsanov-Maruyama theorem, a continuous adapted process X is a local martingale under Q^{T} if and only if

$$X + \left\langle X, \log\frac{P(T)}{M}\right\rangle$$

is a local martingale under Q. In particular,

$$R^{j} + \left\langle R^{j}, \log\frac{P(T_j)}{M}\right\rangle$$

is a continuous local martingale under Q because R^{j} is so under Q^{T_j}. Since

$$\frac{P(T_j)}{P(T_0)} = \prod_{i=1}^{j}\frac{P(T_i)}{P(T_{i-1})} = \prod_{i=1}^{j}\frac{1}{1 + \theta_i R^{i}} \tag{32}$$

and $P(T_0) = P(0) = M$, we have

$$-\mathrm{d}\left\langle X, \log\frac{P(T_j)}{M}\right\rangle = \sum_{i=1}^{j}\frac{\theta_i}{1 + \theta_i R^{i}}\mathrm{d}\langle R^{i}, X\rangle \tag{33}$$

for a continuous semimartingale X.

Similarly, by the Girsanov-Maruyama theorem, a continuous adapted process X is a local martingale under Q if and only if

$$X - \left\langle X, \log\frac{A}{M}\right\rangle$$

is a local martingale under Q^{*}. Note that

$$\mathrm{d}\langle X, \log A \rangle = \frac{1}{A}\mathrm{d}\langle X, A \rangle = \sum_{k=I+1}^{J} \frac{\theta_k P(T_k)}{A} \mathrm{d}\langle X, \log P(T_k) \rangle$$

$$= -\sum_{k=I+1}^{J} \frac{\theta_k P(T_k)}{A} \sum_{i=1}^{k} \frac{\theta_i}{1+\theta_i R^i} \mathrm{d}\langle R^i, X \rangle$$

$$= -\sum_{i=1}^{J} \left(\sum_{k=(I+1)\vee i}^{J} \frac{\theta_k P(T_k)}{A} \right) \frac{\theta_i}{1+\theta_i R^i} \mathrm{d}\langle R^i, X \rangle$$

by (32). This implies that

$$-\left\langle X, \log \frac{A}{M} \right\rangle = \sum_{i=1}^{I+1} \frac{\theta_i}{1+\theta_i R^i} \mathrm{d}\langle R^i, X \rangle + \sum_{i=I+2}^{J} \frac{\sum_{k=i}^{J} \theta_k P(T_k)}{A} \frac{\theta_i}{1+\theta_i R^i} \mathrm{d}\langle R^i, X \rangle. \tag{34}$$

Appendix B

In the numerical experiments in Section 6, option prices under the rough Bergomi model are computed using the Monte Carlo method. As an alternative to Monte Carlo pricing, the rough SABR formula proposed by Fukasawa and Gatheral (2022) provides a short-term approximation for plain vanilla option prices within the rough SABR framework. This section presents the results of using the rough SABR formula in place of the Monte Carlo method. We observe that the rough SABR formula does not provide sufficiently accurate approximations for our purpose.

Table 7 presents the results of separate tenor calibration using the rough SABR formula, corresponding to Table 2 in Section 6.1. The results obtained with the rough SABR formula at $H = 0.5$ are close to those from the Hagan formula. In contrast, the results deviate significantly from Monte Carlo results in Table 2, especially for longer maturities. Nevertheless, both approaches exhibit a similar tendency in which the range of κ narrows as the Hurst exponent decreases.

Table 8 shows the results of the first-step calibration in Section 6.2, where option prices are computed using the rough SABR formula instead of the Monte Carlo method. Compared to Table 3, the results exhibit a tendency for the level of α_i to decrease as the Hurst exponent becomes smaller.

Tables 9 and 10 present the results of the second-step correlation calibration for $H = 0.5$ and $H = 0.15$, respectively. Compared to Tables 4 and 5, the estimated correlations tend to be higher when using the rough SABR formula.

Finally, we compare the results of the FMM and the mapped SMM. For parameter sets obtained using rough SABR formula above, we calculate the swaption IVs for the maturities and underlying where market smile data is available, and compare them with market prices. We plot the following IVs:

Table 7 Model Calibration Results for Different Hurst Exponents by Rough SABR Formula

	1Y	3Y	5Y	10Y
Hagan	1.743	1.095	1.001	0.842
H=0.5	1.730	1.102	0.994	0.851
H=0.45	1.605	1.081	1.001	0.887
H=0.4	1.484	1.057	1.004	0.921
H=0.35	1.365	1.029	1.002	0.953
H=0.3	1.250	0.997	0.997	0.982
H=0.25	1.139	0.962	0.986	1.006
H=0.2	1.031	0.922	0.969	1.025
H=0.15	0.927	0.878	0.947	0.945
H=0.1	0.827	0.830	0.919	1.005
H=0.05	0.731	0.779	0.883	0.740

Table 8 Comparison of Monte Carlo and Rough SABR Formula Calibration Results

Hurst Exponent	κ	α_2	α_4	α_6	α_{11}	ρ_{1Y}	ρ_{3Y}	ρ_{5Y}	ρ_{10Y}	RMSE
H = 0.5	1.18389	0.282	0.227	0.173	0.082	-72.9%	-37.1%	-36.4%	-26.7%	1.339%
H = 0.45	1.24617	0.274	0.217	0.165	0.082	-69.5%	-36.1%	-36.5%	-28.1%	1.159%
H = 0.4	1.30590	0.264	0.205	0.156	0.081	-66.4%	-35.3%	-36.9%	-29.8%	0.977%
H = 0.35	1.36412	0.251	0.191	0.146	0.079	-63.6%	-34.9%	-37.5%	-31.8%	0.795%
H = 0.3	1.42316	0.234	0.174	0.133	0.075	-61.1%	-34.7%	-38.5%	-34.2%	0.620%
H = 0.25	1.48775	0.210	0.152	0.116	0.067	-58.9%	-34.8%	-39.8%	-36.9%	0.463%
H = 0.2	1.56721	0.174	0.122	0.093	0.055	-56.8%	-35.3%	-41.5%	-40.2%	0.344%
H = 0.15	1.68225	0.120	0.080	0.060	0.035	-55.0%	-36.0%	-43.7%	-44.2%	0.305%
H = 0.1	1.88812	0.045	0.027	0.019	0.011	-53.3%	-37.1%	-46.3%	-48.9%	0.359%
H = 0.05	2.41007	3.37e-4	1.48e-4	9.25e-5	4.45e-5	-51.9%	-38.7%	-49.6%	-54.5%	0.464%

Table 9 Correlation Matrix Among Interest Rate Factors. (rough SABR formula, H=0.5)

1.000	0.9726	0.9060	0.9045	0.9030	0.8938	0.8844	0.8746	0.8645	0.8541
0.9726	1.000	0.9796	0.9789	0.9782	0.9736	0.9687	0.9633	0.9576	0.9516
0.9060	0.9796	1.000	0.9999	0.9999	0.9996	0.9988	0.9976	0.9960	0.9939
0.9045	0.9789	0.9999	1.000	0.9999	0.9997	0.9990	0.9978	0.9963	0.9943
0.9030	0.9782	0.9999	0.9999	1.000	0.9998	0.9991	0.9981	0.9966	0.9947
0.8938	0.9736	0.9996	0.9997	0.9998	1.000	0.9998	0.9991	0.9981	0.9966
0.8844	0.9687	0.9988	0.9990	0.9991	0.9998	1.000	0.9998	0.9992	0.9981
0.8746	0.9633	0.9976	0.9978	0.9981	0.9991	0.9998	1.000	0.9998	0.9992
0.8645	0.9576	0.9960	0.9963	0.9966	0.9981	0.9992	0.9998	1.000	0.9998
0.8541	0.9516	0.9939	0.9943	0.9947	0.9966	0.9981	0.9992	0.9998	1.000

Table 10 Correlation Matrix Among Interest Rate Factors. (rough SABR formula, H=0.15)

1.000	0.9939	0.9772	0.9851	0.9915	0.9916	0.9918	0.9919	0.9921	0.9922
0.9939	1.000	0.9946	0.9980	0.9998	0.9998	0.9998	0.9999	0.9999	0.9999
0.9772	0.9946	1.000	0.9991	0.9965	0.9964	0.9963	0.9962	0.9961	0.9960
0.9851	0.9980	0.9991	1.000	0.9991	0.9991	0.9990	0.9990	0.9989	0.9989
0.9915	0.9998	0.9965	0.9991	1.000	1.000	1.000	0.9999	0.9999	0.9999
0.9916	0.9998	0.9964	0.9991	1.000	1.000	1.000	1.000	0.9999	0.9999
0.9918	0.9998	0.9963	0.9990	1.000	1.000	1.000	1.000	1.000	0.9999
0.9919	0.9999	0.9962	0.9990	0.9999	1.000	1.000	1.000	1.000	1.000
0.9921	0.9999	0.9961	0.9989	0.9999	0.9999	1.000	1.000	1.000	1.000
0.9922	0.9999	0.9960	0.9989	0.9999	0.9999	0.9999	1.000	1.000	1.000

- Swaption IVs computed using Monte Carlo simulation under the FMM model "FMM",
- Swaption IVs under our approximative swap rate model (23) using the Monte Carlo method "Mapped SMM(MC)",
- Swaption IVs under our approximative swap rate model (23) using the rough SABR formula "Mapped SMM(formula)",
- Market IVs "Market".

Error bars corresponding to the ±2 standard deviations are shown for both "FMM" and "Mapped SMM (MC)".

Figure 5, 6 show the quality of in-sample fit, as the calibration was done using swaptions with 1Y tenor. The fit of Mapped SMM (formula) to Market is good for all 1Y tenor data, however, FMM and Mapped SMM (MC) are largely deviate for other maturities.

In Figure 7, 8, we present volatility smiles with tenor more than 2Y. These results indicate that the Mapped SMM (formula) tends to deviate from market observations as the tenor of the underlying increases. Moreover, significant discrepancies were observed when compared with both the FMM and the Mapped SMM (MC).
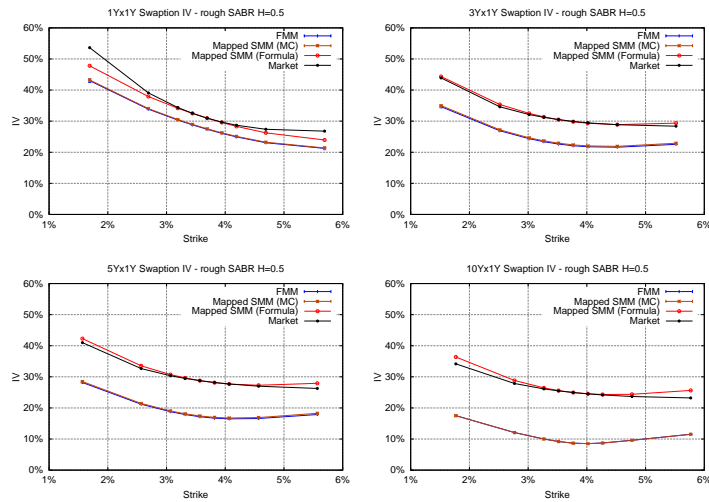

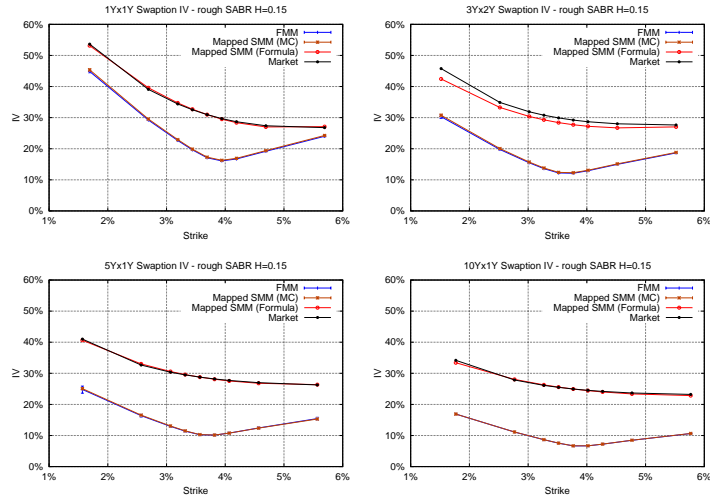**Fig. 5** Model and Market IVs of 1Y tenor calibrated with the rough SABR formula at $H = 0.5$

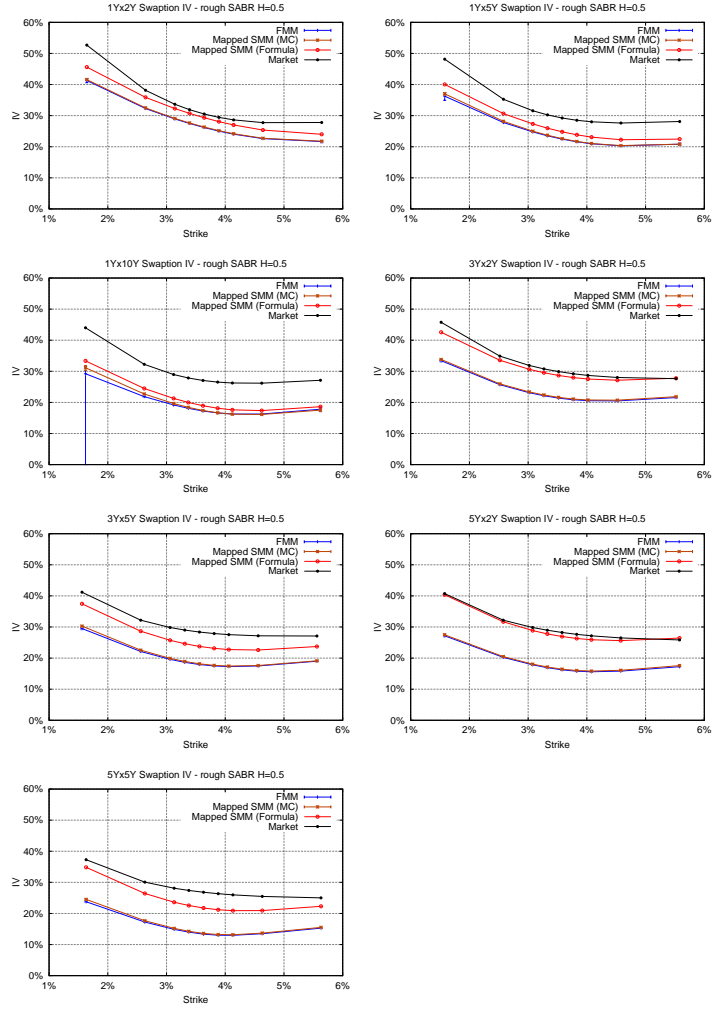**Fig. 6** Model and Market IVs of 1Y tenor calibrated with the rough SABR formula at $H = 0.15$

Fig. 7 Model and Market IVs of non 1Y tenor calibrated with the rough SABR formula at $H = 0.5$
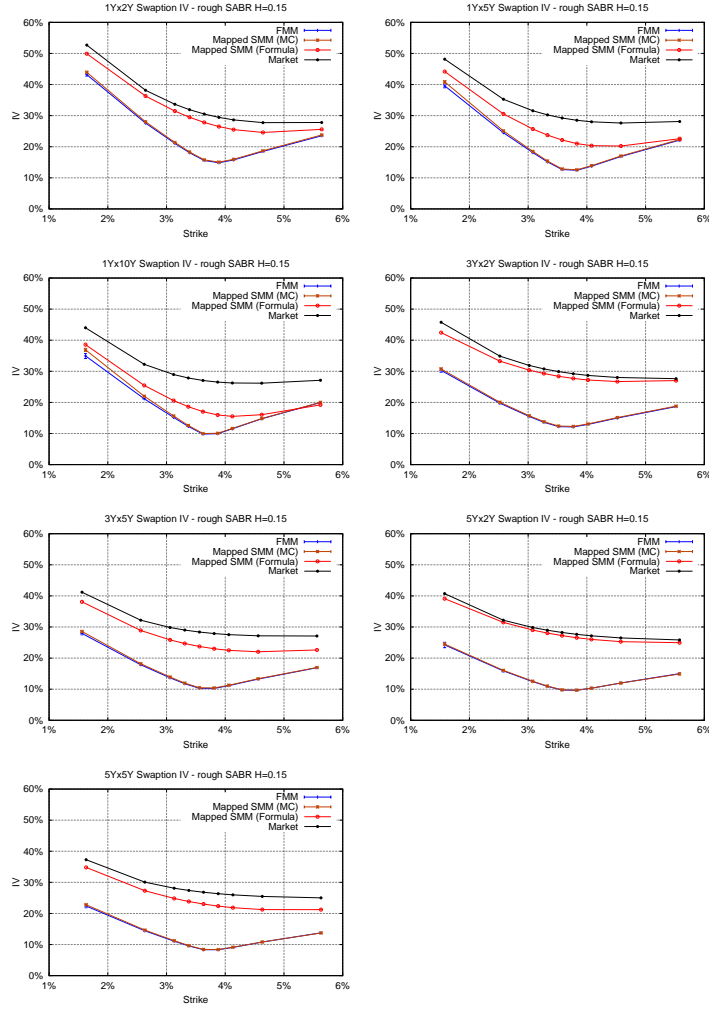
Fig. 8 Model and Market IVs of non 1Y tenor calibrated with the rough SABR formula at $H = 0.15$

Appendix C

Here we discuss the well-posedness of the rough SABR forward market model (15).
We start by fixing a probability space $(\Omega, \mathscr{F}, Q^{T_N})$ and noting the well-posedness
of the corresponding stochastic differential equation under Q^{T_N}:

$$dR_t^j = \gamma_j(t)\eta_j(R_t^j)\sqrt{V_t^j}dW_t^{j\dagger},$$

$$\log V_t^j = \log\xi_j(t) - \frac{1}{2}\int_0^t \zeta_j(t-s)^2 ds + \int_0^t \zeta_j(t-s)d\bar{W}_s^0, \tag{35}$$

where

$$W^{j\dagger} = \bar{W}^j - \sum_{i=j+1}^N \int_0^{\cdot} \frac{\theta_i}{1+\theta_i R_t^i}\gamma_i(t)\eta_i(R_t^i)\sqrt{V_t^i}\rho_{ij}dt$$

and $(\bar{W}^0, \bar{W}^1, \ldots, \bar{W}^N)$ is a correlated Brownian motion under Q^{T_N} with

$$\langle \bar{W}^j \rangle_t = t, \quad \langle \bar{W}^i, \bar{W}^j \rangle_t = \rho_{ij}t.$$

Lemma 2 *If η_j is Lipschitz continuous with (17) and $R_0^j > 0$ for each $j = 1, \ldots, N$, then there exists a unique strong solution $(R^1, \ldots, R^N)$ to (35).*

Proof. Noting that V^j is explicit, the result follows from the standard argument. $\square$

Lemma 3 *Let $(R^1, \ldots, R^N)$ be a solution to (35). Then, under (17) and (18),*

$$\prod_{j=1}^N (1 + \theta_j R^j)$$

is a martingale under Q^{T_N}.

Proof. Let

$$U = \sum_{j=1}^N U^j, \quad U_t^j = \int_0^t \frac{\theta_j}{1+\theta_j R_s^j}\gamma_j(s)\eta_j(R_s^j)\sqrt{V_s^j}d\bar{W}_s^j.$$

Since

$$d\log\prod_{j=1}^N(1+\theta_j R^j) = \sum_{j=1}^N d\log(1+\theta_j R^j)$$

$$= \sum_{j=1}^N \left\{ \frac{\theta_j dR^j}{1+\theta_j R^j} - \frac{1}{2}\frac{\theta_j^2 d\langle R^j\rangle}{(1+\theta_j R^j)^2} \right\}$$

$$= \sum_{j=1}^N dU^j - \frac{1}{2}\sum_{i,j=1}^N d\langle U^i, U^j \rangle$$

$$= dU - \frac{1}{2}d\langle U\rangle,$$

the product is a nonnegative local martingale. It suffices then to show

$$\mathsf{E}_{Q^{T_N}} \left[\prod_{j=1}^{N} (1 + \theta_j R_{T_N}^j) \right] = 1.$$

Let

$$\tau_n = \inf \left\{ t \geq 0; \max_{j=1,\dots,N} V_t^j = n \right\}.$$

for integers n and define a sequence of probability measures Q^n by

$$\frac{dQ^n}{dQ^{T_N}} = \exp \left\{ U_{\tau_n} - \frac{1}{2} \langle U \rangle_{\tau_n} \right\}.$$

Here we have used (17). Then, we have

$$\begin{aligned}
1 &= \mathsf{E}_{Q^{T_N}} \left[\exp \left\{ U_{\tau_n \wedge T_N} - \frac{1}{2} \langle U \rangle_{\tau_n \wedge T_N} \right\} \right] \\
&= \mathsf{E}_{Q^{T_N}} \left[\exp \left\{ U_{T_N} - \frac{1}{2} \langle U \rangle_{T_N} \right\}; \tau_n > T_N \right] + Q^n(\tau_n \leq T_N).
\end{aligned} \tag{36}$$

By the Girsanov-Maruyama theorem,

$$\hat{W}^n = \bar{W}^0 - \langle \bar{W}^0, U_{\cdot \wedge \tau_n} \rangle = \bar{W}^0 - \sum_{j=1}^{N} \int_0^{\cdot \wedge \tau_n} \frac{\theta_j}{1 + \theta_j R_t^j} \gamma_j(t) \eta_j(R_t^j) \sqrt{V_t^j} \rho_{0j} dt$$

is a Brownian motion under Q^n. Under (18),

$$\begin{aligned}
\log V_t^j &= \log \xi_j(t) - \frac{1}{2} \int_0^t \zeta_j(t-s)^2 ds + \int_0^t \zeta_j(t-s) d\bar{W}_s^0 \\
&\leq \log \xi_j(t) - \frac{1}{2} \int_0^t \zeta_j(t-s)^2 ds + \int_0^t \zeta_j(t-s) d\hat{W}_s^n =: \log \hat{V}_t^{j,n}
\end{aligned}$$

and so,

$$Q^n(\tau_n \leq T_N) = Q^n \left(\sup_{0 \leq t \leq T_N} \max_{j=1,\dots,N} V_t^j \geq n \right) \leq Q^n \left(\sup_{0 \leq t \leq T_N} \max_{j=1,\dots,N} \hat{V}_t^{j,n} \geq n \right).$$

The law of $\hat{V}^{j,n}$ under Q^n does not depend on n. Therefore, the right hand side converges to 0 as $n \to \infty$. Together with the monotone convergence theorem, this implies that the right hand side of (36) converges to

$$\mathsf{E}_{Q^{T_N}} \left[\exp \left\{ U_{T_N} - \frac{1}{2} \langle U \rangle_{T_N} \right\} \right] = \mathsf{E}_{Q^{T_N}} \left[\prod_{j=1}^{N} (1 + \theta_j R_{T_N}^j) \right],$$

which concludes the proof.   □

Lemma 4 *Let $(R^1, \ldots, R^N)$ be a solution to (35) under (17) and (18). Define a probability measure* Q *by*

$$\frac{d\mathsf{Q}}{d\mathsf{Q}^{T_N}} = \prod_{j=1}^{N} \frac{1 + \theta_j R_{T_N}^j}{1 + \theta_j R_0^j}.$$

Then,

$$W^j := \bar{W}^j - \sum_{i=1}^{N} \int_0^{\cdot} \frac{\theta_i}{1 + \theta_i R_t^i} \gamma_i(t) \eta_i(R_t^i) \sqrt{V_t^i} \rho_{ij} dt, \quad j = 0, 1, \ldots, N$$

is a Brownian motion under Q *with (16), and* $(R^1, \ldots, R^N)$ *is a solution to (15) with* $\bar{W} = \bar{W}^0$.

Proof. By the previous lemma, Q is a probability measure. The result follows from the standard application of the Girsanov-Maruyama theorem. $\square$

Theorem 3 *If η_j is Lipschitz continuous with (17) and $R_0^j > 0$ for each $j = 1, \ldots, N$, then, there exists a weak solution to (15) under (18) which is unique in law.*

Proof. The existence of a weak solution is from the preceding lemmas. For any solution,

$$\prod_{j=1}^{N} \frac{1}{1 + \theta_j R^i}$$

is a bounded local martingale and hence a martingale. We can therefore define a probability measure Q^{T_N} by

$$\frac{d\mathsf{Q}^{T_N}}{d\mathsf{Q}} = \prod_{j=1}^{N} \frac{1 + \theta_j R_0^j}{1 + \theta_j R_{T_N}^j},$$

under which $(R^1, \ldots, R^N)$ solves (35) with $\bar{W}^0 = \bar{W}$. Consequently, the law of $(R^1, \ldots, R^N)$ is determined by that under Q^{T_N} that is unique by Lemma 2. $\square$

References

Bennedsen, M., Lunde, A., & Pakkanen, M. S. (2017). Hybrid scheme for brownian semistationary processes. Finance and Stochastics, 21, 931–965.

El Euch, O., Fukasawa, M., & Rosenbaum, M. (2018). The microstructural foundations of leverage effect and rough volatility. Finance and Stochastics, 22, 241–280.

Fukasawa, M. (2021). Volatility has to be rough. Quantitative Finance, 21, 1–8.

Fukasawa, M. (2022). On asymptotically arbitrage-free approximations of the implied volatility. Frontiers of Mathematical Finance, 1, 525–537.

Fukasawa, M., & Gatheral, J. (2022). A rough SABR formula. <u>Frontiers of Mathematical Finance</u>, <u>1</u>, 81–97.

Fukasawa, M., Horvath, B., & Tankov, P. (2023). Hedging under rough volatility. In <u>Rough volatility</u> (pp. 115–125). SIAM.

Gaspar, R. M., & Pimentel, R. (2017). On swap rate dynamics: To freeze or not to freeze? <u>International Journal of Computer Mathematics</u>, <u>94</u>(11), 2208–2222.

Gassiat, P. (2019). On the martingale property in the rough bergomi model. <u>Electronic Communications in Probability</u>, <u>24</u>(33), 9.

Lyashenko, A., & Mercurio, F. (2019). <u>Looking forward to backward-looking rates: A modeling framework for term rates replacing LIBOR.</u> Available at SSRN. Retrieved from `https://ssrn.com/abstract=3330240`

Rapisarda, F., Brigo, D., & Mercurio, F. (2007). Parameterizing correlations: A geometric interpretation. <u>IMA Journal of Management Mathematics</u>, <u>18</u>(1), 55–73.

Rebonato, R., McKay, K., & White, R. (2009). <u>The SABR/LIBOR market model: Pricing, calibration and hedging for complex interest-rate derivatives.</u> Hoboken, NJ: John Wiley & Sons.