Stochastic evolution systems with constant coefficients

S. V. Lototsky · J. Zhong

Abstract While solvability of a single stochastic hyperbolic or parabolic equation is well known, the problem remains mostly open for stochastic evolution systems. The paper investigates well-posedness and stability in Sobolev spaces on $\mathbb{R}^d$ of the initial value problem for systems of stochastic evolution equations with constant coefficients and multiplicative time-only Gaussian white noise. A general criterion for well-posedness is derived in terms of sums of certain Kronecker products of the system matrices, and a stochastic analogue of the Petrowski parabolicity condition is proposed.

Keywords Kronecker product · Spectral abscissa · Stability · Stochastic parabolic systems · Stochastic parabolicity condition

Mathematics Subject Classification Primary 60H15 · Secondary 35E15

1 Introduction

What is a natural extension of the Petrowski[1] parabolicity condition to stochastic systems? Can we solve a system if it is neither hyperbolic nor parabolic? The answers are non-trivial even for simple examples.

As an example of a parabolic system, consider

$$du = u_{xx}dt; \ dv = 2v_{xx}dt + \sigma u_x dw(t), \quad t > 0, \ x \in \mathbb{R}, \ \sigma > 0. \qquad (1.1)$$

---

[1] We follow the spelling from [7].

S. V. Lototsky (✉) · J. Zhong
Department of Mathematics, USC, Los Angeles, CA 90089, USA
e-mail: lototsky@usc.edu
URL: http://www-rcf.usc.edu/~lototsky

J. Zhong
e-mail: jiezhong@usc.edu

On the one hand, if $u_0, v_0 \in L_2(\mathbb{R})$, then standard results for parabolic equations imply that $u \in L_2(\Omega \times (0, T); H^1)$, so that $\sigma u_x \in L_2(\Omega \times (0, T) \times \mathbb{R})$, and therefore $v \in L_2(\Omega \times (0, T); H^1)$, that is, the system is reasonable to call parabolic for every $\sigma > 0$. On the other hand, if we ignore the lower-triangular structure of the system and try to estimate $|u_x|^2 + |v_x|^2$ using the Itô formula and integration by parts, then the usual parabolic-type estimate will be possible only when $\sigma^2 < 2$—a condition consistent with [12, Assumption 2.2].

As far as equations that are neither hyperbolic nor parabolic, let us start with a deterministic example

$$u_{tt} - a u_{txx} = c^2 u_{xx}, \quad a > 0, \ c > 0. \tag{1.2}$$

Even though it is a damped wave equation, it is not hyperbolic because its order in space-time (three) is different from its order in time (two); cf. [14, Theorem 2.7]. Similarly, even though the function $v = u_t$ satisfies what looks like an inhomogeneous heat equation, (1.2) is not parabolic: there are several equivalent conditions that can be checked and turn out not to hold; see, for example [5, Sect. 1.3.3] or [7, Sect. 9.1]. Still, direct computations, either using integration by parts or the Fourier transform, show that (1.2) is well-posed for every $a > 0$ in a way similar to the usual wave equation.

Now consider a stochastic version of (1.2),

$$u_{tt} - a u_{txx} = c^2 u_{xx} + \sigma u_{xx} \dot{w}, \quad a > 0. \tag{1.3}$$

While integration by parts no longer works, the Fourier transform approach shows that the equation is well-posed for every $\sigma \in \mathbb{R}$, and if $2ac^2 > \sigma^2$, then the zero solution is asymptotically stable; see Sect. 4 for details.

To state the main result of the paper, consider two matrix partial differential operators A and B with constant coefficients and corresponding symbols $A(y)$, $B(y)$; a symbol in this case is a square matrix and the entries of the matrix are polynomials in y with complex coefficients. Let $w = w(t)$ be a standard Brownian motion and consider the system of stochastic Itô equations for the unknown vector function $\boldsymbol{u}(t, x) \in \mathbb{R}^N$:

$$\boldsymbol{u}(t, x) = \boldsymbol{u}_0(x) + \int_0^t \big(\mathrm{A}\boldsymbol{u}(s, x) + \boldsymbol{f}(s, x) \big) ds$$

$$+ \int_0^t \big(\mathrm{B}\boldsymbol{u}(s, x) + \boldsymbol{g}(s, x) \big) dw(s), \quad 0 \le t \le T, \ x \in \mathbb{R}^{\mathrm{d}}. \tag{1.4}$$

We say that system is well-posed if, for every collection of input $\boldsymbol{u}_0, \boldsymbol{f}, \boldsymbol{g}$ in some Sobolev spaces, there exists a unique solution $\boldsymbol{u}(t, x)$ with values in some Sobolev space, and the norm of the solution continuously depends on the norm of the input.

Define the matrix $\mathcal{M}_{A,B}(y) = \overline{A(y)} \otimes I + I \otimes A + \overline{B(y)} \otimes B(y)$, where‾means complex conjugation, I is the N-by-N identity matrix, and $\otimes$ denotes the Kronecker product of two matrices.

Theorem 1.1 *System* (1.4) *is well-posed on a fixed time interval* $[0, T]$ *if and only if the real parts of the eigenvalues of* $\mathcal{M}_{A,B}(y)$ *are uniformly bounded in* $|y|$: *there exists a number* $C_0 \in \mathbb{R}$ *such that, for every eigenvalue* $\lambda_{\mathcal{M}_{A,B}}(y)$ *of the matrix* $\mathcal{M}_{A,B}(y)$ *and all* $y \in \mathbb{R}^d$,

$$\Re\lambda_{\mathcal{M}_{A,B}}(y) \le C_0. \tag{1.5}$$

In Sect. 4 we show that, according to this theorem, Eq. (1.1) is well-posed for all $\sigma \in \mathbb{R}$, and Eq. (1.2) is well-posed for $a > 0$ and all $\sigma, c \in \mathbb{R}$.

Denote by B^* the complex conjugate of the matrix B and denote by $\| \cdot \|$ any matrix norm. The stochastic parabolicity condition is as follows.

Definition 1.2 System (1.4) is called parabolic of order $2p$ if $\|A(y)\| + \|B^*(y)B(y)\|$ is a polynomial of degree $2p$, and, for some $\varepsilon > 0$ and $L \in \mathbb{R}$, all eigenvalues of the matrix $\mathcal{M}_{A,B}(y)$ satisfy

$$\Re\lambda\big(\mathcal{M}_{A,B}(y)\big) \le -\varepsilon|y|^{2p} + L.$$

The rest of the paper is organized as follows. Section 2 introduces the main notations and notions, including the Kronecker product, necessary to state and prove the main result. The main result is in Sect. 3. Section 4 presents various ramifications of the main result for special types of systems, including parabolic ones. Section 5 discusses alternative approaches and possible generalizations of the main result. The overall summary is in Sect. 6.

2 Function spaces, operators, and matrices

Our main object is a function of a d-dimensional real variable $x = (x_1, \ldots, x_d)$, taking values in the real N-dimensional space $\mathbb{R}^N$ or in the complex N-dimensional space $\mathbb{C}^N$. Such a function will usually be denoted by a lower-case bold Latin letter, e.g. $\boldsymbol{h}$, whereas upper-case regular Latin letter, e.g. A, will usually mean an N-by-N matrix. Then $|\boldsymbol{h}|$ is the Euclidean norm of $\boldsymbol{h}$ and $\|A\|$ is the induced matrix norm

$$\|A\| = \max\big\{|A\boldsymbol{h}| : |\boldsymbol{h}| = 1\big\}.$$

For a matrix $A \in \mathbb{C}^{N \times N}, \overline{A}$ is the matrix with complex conjugate entries, $A^\top$ means transposition, and A^* denotes the conjugate transpose: $A^* = \overline{A^\top} = \overline{A}^\top$. The same notations, $\bar{}$, $\top$, and *, will also be used for vectors in $\mathbb{C}^N$.

The Sobolev space $H^\gamma = H^\gamma(\mathbb{R}^d; \mathbb{R}^N)$ of $\mathbb{R}^N$-valued functions on $\mathbb{R}^d$ is the collection of generalized functions $\boldsymbol{h} \in \mathcal{S}' = \mathcal{S}'(\mathbb{R}^d; \mathbb{R}^N)$ such that

$$\|\boldsymbol{h}\|_\gamma^2 = \int_{\mathbb{R}^d} (1 + |y|^2)^\gamma |\widehat{\boldsymbol{h}}(y)|^2 dy < \infty,$$

where $\widehat{h}$ is the Fourier transform of h; for $h \in L_1(\mathbb{R}^d; \mathbb{R}^N)$,

$$\widehat{h}(y) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-ixy} h(x) dx, \quad i = \sqrt{-1}.$$

Clearly, $H^{\gamma_2} \subset H^{\gamma_1}, \gamma_2 > \gamma_1$. If $H^\infty = \bigcap_{\gamma > 0} H^\gamma$, $H_\infty = \bigcup_{\gamma < 0} H^\gamma$, and $\mathcal{S} = \mathcal{S}(\mathbb{R}^d; \mathbb{R}^N)$ is the space of rapidly decreasing test functions, then

$$\mathcal{S} \subset H^\infty \subset H_\infty \subset \mathcal{S}',$$

with all inclusions strict: for example, in the case $N = d = 1$, the function $h(x) = 1/(1 + x^2)$ is in H^∞ but not in $\mathcal{S}$. For $h \in H_\infty$ and $z \in H^\infty$, define

$$\langle h, z \rangle = \int_{\mathbb{R}^d} \widehat{h}^*(y) \widehat{z}(y) dy. \tag{2.1}$$

By the Cauchy-Schwarz inequality,

$$|\langle h, z \rangle| \le \|h\|_{-\gamma} \|z\|_\gamma$$

with sufficiently large $\gamma > 0$.

Definition 2.1 An operator A on H_∞ is called pseudo-differential opera-tor with constant coefficients if there exists a matrix-valued function $A = A(y) \in \mathbb{C}^{N \times N}$, $y \in \mathbb{R}^d$, such that,

(1) the entries of A have at most polynomial growth:

$$\|A(y)\| \le C_A(1 + |y|)^a \tag{2.2}$$

for some $C_A > 0$ and $a \ge 0$;
(2) for every $h \in H_\infty$,

$$\widehat{Ah}(y) = A(y)\widehat{h}(y), \tag{2.3}$$

(3) all entries of A are continuous functions of y.

The matrix A is called the symbol of the operator A.

In the above definition, the polynomial growth condition (2.2) is standard (without it, the operator might not map H_∞ to itself), while both (2.3) and the "constant coef-ficients" part of the name come from the observation that if A is a partial differential operator

$$Ah(x) = \sum_{|\alpha| \le m} A_\alpha \frac{\partial^{|\alpha|} h(x)}{\partial x_1^{\alpha_1} \cdots x_d^{\alpha_d}}, \quad A_\alpha \in \mathbb{R}^{N \times N}, \quad |\alpha| = \alpha_1 + \cdots + \alpha_d,$$

then

$$A(y) = \sum_{|\alpha| \le m} i^{|\alpha|} y_1^{\alpha_1} \cdots y_d^{\alpha_d} A_\alpha, \quad i = \sqrt{-1}.$$

Our definition makes it possible to consider more general operators such as $(-\nabla^2)^r, r > 0$, with symbol $|y|^{2r}$. The continuity requirement is for purely technical reasons, to avoid unnecessary complications in the future; it is still much weaker than the usual C^∞ requirement if the operator is to act on $\mathcal{S}'$ instead of H_∞.

Proposition 2.2 *Under* (2.2)*, the operator A is bounded from H^γ to $H^{\gamma-a}$ for every* $\gamma \in \mathbb{R}$.

Proof This follows directly from (2.3) and the definition of the norm $\| \cdot \|_\gamma$. □

Next, we present a brief summary of the Kronecker product and related topics; possible references are [4] or [10, Chap. 4].

Kronecker product is a construction allowing a product of three matrices to be written in an equivalent matrix-vector form. Consider a matrix equation $BXA^\top = C$ with known matrices A, B, C, and the unknown matrix X; given the nature of our applications, we assume that all matrices are in $\mathbb{C}^{N \times N}$. Introduce a column vector $\text{vec}(X) = \boldsymbol{X} \in \mathbb{C}^{N^2}$ by stacking together the columns of X, left-to-right:

$$\text{vec}(X) = \boldsymbol{X} = (X_{11}, \dots, X_{n1}, X_{12}, \dots, X_{n2}, \dots, X_{1n}, \dots, X_{nn})^\top. \quad (2.4)$$

Then direct computations show that the matrix equation $AXB^\top = C$ can be written in the matrix-vector form for the unknown vector $\boldsymbol{X}$ as

$$(A \otimes B)\boldsymbol{X} = \boldsymbol{C}, \quad \boldsymbol{C} = \text{vec}(C), \quad (2.5)$$

where $A \otimes B$ is the $\texttt{Kronecker product}$ of matrices A and B, that is, an N^2-by-N^2 block matrix with blocks $A_{ij}B$. For example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \otimes \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b & 2a & 2b \\ c & d & 2c & 2d \\ 3a & 3b & 4a & 4b \\ 3c & 3d & 4c & 4d \end{pmatrix}.$$

In other words, (2.5) means

$$\text{vec}(BXA^\top) = (A \otimes B)\text{vec}(X), \quad (2.6)$$

with $\text{vec}(\cdot)$ operation defined in (2.4). The point is that, at least for theoretical purposes, the matrix-vector Eq. (2.5) is a more familiar, and hence convenient, object than the corresponding matrix-matrix equation.

3 The main result

Let A and B be pseudo-differential operators with constant coefficients; denote by $A(y)$, $B(y)$ the corresponding symbols. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)T \geq 0, \mathbb{P})$ be a stochastic basis with the usual assumptions and with a standard Brownian motion $w = w(t)$. Consider the system of stochastic Itô equations for the unknown vector function $\boldsymbol{u} \in \mathbb{R}^N$:

$$\boldsymbol{u}(t) = \boldsymbol{u}_0 + \int_0^t \big(A\boldsymbol{u}(s) + \boldsymbol{f}(s)\big)ds + \int_0^t \big(B\boldsymbol{u}(s) + \boldsymbol{g}(s)\big)dw(s), \quad 0 \leq t \leq T. \quad (3.1)$$

To simplify the presentation, we assume that there is only one Brownian motion and it is the only source of randomness in the system: the initial condition $\boldsymbol{u}_0$, the free terms $\boldsymbol{f}$, $\boldsymbol{g}$, and the operators A, B are all non-random. Many of these restrictions can be removed, and we will discuss this in Sect. 5.

Sometimes it is convenient to write (3.1) in a less formal differential way

$$\dot{\boldsymbol{u}} = A\boldsymbol{u} + \boldsymbol{f} + (B\boldsymbol{u} + \boldsymbol{g})\dot{w}. \quad (3.2)$$

Definition 3.1 Given

$$\boldsymbol{u}_0 \in H^r, \ \boldsymbol{f}, \boldsymbol{g} \in L_2((0, T); H^r) \text{ for some } r \in \mathbb{R}, \quad (3.3)$$

a solution of (3.1) is an $\mathcal{F}_t$-adapted process $\boldsymbol{u} \in L_2(\Omega \times [0, T]; H^\gamma)$ for some $\gamma \leq r$ with the following property: there is a set $\Omega' \subseteq \Omega$ with $\mathbb{P}(\Omega') = 1$ such that, for every $\omega \in \Omega'$, every $t \in [0, T]$, and every $z \in H^\infty$,

$$\langle \boldsymbol{u}(t), z \rangle = \langle \boldsymbol{u}_0, z \rangle + \int_0^t \langle A\boldsymbol{u}(s) + \boldsymbol{f}(s), z \rangle ds + \int_0^t \langle B\boldsymbol{u}(s) + \boldsymbol{g}(s), z \rangle dw(s). \quad (3.4)$$

Equation (3.1) is called `well-posed` on $[0, T]$ if

- for every $r \in \mathbb{R}$ and input satisfying (3.3), there exists a unique solution $\boldsymbol{u}$, and
- there exists a $C > 0$ such that

$$\mathbb{E}\|\boldsymbol{u}(t)\|_\gamma^2 \leq C\Big(\|\boldsymbol{u}_0\|_r^2 + \int_0^t \|\boldsymbol{f}(s)\|_r^2 ds + \int_0^t \|\boldsymbol{g}(s)\|_r^2 ds\Big), \quad 0 \leq t \leq T. \quad (3.5)$$

Remark 3.2 (1) The definition allows both γ and C in (3.5) to depend on T.
(2) While it is possible that (3.5) can hold for $t > 0$ with $\gamma > r$, we need to assume $\gamma \leq r$ if (3.5) is to hold for $t = 0$.
(3) Since (3.1) is linear with constant coefficients, it is enough to establish (3.5) for $r = 0$.

Using (2.1) and the notation

$$(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{z}}) = \int\limits_{\mathbb{R}^d} \widehat{\boldsymbol{h}}^*(y)\widehat{\boldsymbol{z}}(y)dy,$$

we can re-write (3.4) as

$$(\widehat{\boldsymbol{u}}(t), \widehat{\boldsymbol{z}}) = (\widehat{\boldsymbol{u}}_0, \widehat{\boldsymbol{z}}) + \int\limits_0^t \big(A(y)\widehat{\boldsymbol{u}}(s) + \widehat{\boldsymbol{f}}(s), \widehat{\boldsymbol{z}}\big)ds + \int\limits_0^t \big(B(y)\widehat{\boldsymbol{u}}(s) + \widehat{\boldsymbol{g}}(s), \widehat{\boldsymbol{z}}\big)dw(s).$$

(3.6)

Consider the linear system of stochastic ordinary differential equations depending on the parameter y:

$$\boldsymbol{v}(t, y) = \widehat{\boldsymbol{u}}_0(y) + \int\limits_0^t \big(A(y)\boldsymbol{v}(s, y) + \widehat{\boldsymbol{f}}(s, y)\big)ds + \int\limits_0^t \big(B(y)\boldsymbol{v}(s, y) + \widehat{\boldsymbol{g}}(s, y)\big)dw(s).$$

(3.7)

Of course, (3.7) is equivalent to (3.6) if a solution of (3.1) exists. The precise connection between the two systems is as follows.

Proposition 3.3 *Assume* (3.3) *holds. Then*

(1) For every $y \in \mathbb{R}^d$, system (3.7) *has a unique $\mathcal{F}_t$-adapted solution, and the fundamental solution of the system is a continuous function of y;*

(2) If $\boldsymbol{v} = \boldsymbol{v}(t, y)$ is the solution of (3.7) *and*

$$\mathbb{E} \int\limits_{\mathbb{R}^d} (1 + |y|^2)^\gamma |\boldsymbol{v}(t, y)|^2 dy < \infty$$

(3.8)

for some $\gamma \in \mathbb{R}$ and all $t \in [0, T]$, then $\boldsymbol{u}(t) = \check{\boldsymbol{v}}(t)$ is the solution of (3.1)*, where $\check{}$ denotes the inverse Fourier transform;*

(3) If $\boldsymbol{u}$ is a solution of (3.1)*, then $\boldsymbol{v} = \widehat{\boldsymbol{u}}$ is a solution of* (3.7)*. In particular,* (3.1) *has at most one solution.*

Proof Recall that the fundamental solution of (3.7) is the matrix $\Phi(t, s, y)$, $t \geq s \geq 0$, such that

$$\Phi(t, s, y) = I + \int\limits_s^t A(y)\Phi(\tau, s, y)d\tau + \int\limits_s^t B(y)\Phi(\tau, s, y)dw(\tau);$$

$I \in \mathbb{R}^{N \times N}$ is the identity matrix. Existence and uniqueness of Φ is standard, and continuity of Φ with respect to y follows from continuity of $A(y)$ and $B(y)$; see [8, Corollary 2.7.1]. While existence and uniqueness of solution of (3.7) are also standard for every fixed y, representation

$$\boldsymbol{v}(t, y) = \Phi(t, 0, y)\widehat{\boldsymbol{u}}_0(y) + \int_0^t \Phi(t, s)\widehat{\boldsymbol{f}}(s, y)ds + \int_0^t \Phi(t, s)\widehat{\boldsymbol{g}}(s, y)dw(s) \quad (3.9)$$

and continuity of Φ with respect to y ensure that the exceptional set involved in the construction of $\boldsymbol{v}$ does not depend on y. In particular, if (3.8) holds, then $\check{\boldsymbol{v}}$ is well defined. The rest of the proposition now follows from (3.6).

This completes the proof of Proposition 3.3   □

Proposition 3.3 reduces analysis of (3.1) to analysis of (3.7), and more specifically, to verification of (3.8). Let us look at the homogeneous version of (3.7):

$$\boldsymbol{v}(t, y) = \widehat{\boldsymbol{u}}_0(y) + \int_0^t A(y)\boldsymbol{v}(s, y)ds + \int_0^t B(y)\boldsymbol{v}(s, y)dw(s). \quad (3.10)$$

By the Itô formula, the scalar quantity $\mathfrak{v}(t, y) = \mathbb{E}\boldsymbol{v}^*(t, y)\boldsymbol{v}(t, y) = \mathbb{E}|\boldsymbol{v}(t, y)|^2$ satisfies

$$\mathfrak{v}(t, y) = \widehat{\boldsymbol{u}}_0^*\widehat{\boldsymbol{u}}_0 + 2\Re \int_0^t \mathbb{E}\boldsymbol{v}^*(s, y)A(y)\boldsymbol{v}(s, y)ds$$

$$+ \int_0^t \mathbb{E}\boldsymbol{v}^*(s, y)B^*(y)B(y)\boldsymbol{v}(s, y)ds, \quad (3.11)$$

which is not an equation[2] for $\mathfrak{v}$. While one can still proceed with this approach and derive some sufficient conditions for solvability of (3.1), we temporarily abandon (3.11) in favor of a different approach. We will re-examine (3.11) in Sect. 5.

Instead of looking at the *norm* of the solution, let us look at the *covariance matrix* $U(t, y) = \mathbb{E}\boldsymbol{v}(t, y)\boldsymbol{v}^*(t, y)$. With $\boldsymbol{U} = \text{vec}(U)$, the Cauchy-Schwartz inequality implies

$$|\boldsymbol{U}|^2 = \sum_{i,j=1}^{N^2} \left|\mathbb{E}v_i\bar{v}_j\right|^2 \geq \sum_{i=1}^{N^2} \left(\mathbb{E}|v_i|^2\right)^2 \geq \frac{1}{N^2}\left(\mathbb{E}|\boldsymbol{v}|^2\right)^2,$$

that is,

$$\mathbb{E}|\boldsymbol{v}|^2 \leq N|\boldsymbol{U}|. \quad (3.12)$$

By the Itô formula,

$$U(t, y) = \widehat{\boldsymbol{u}}_0\widehat{\boldsymbol{u}}_0^* + \int_0^t \left(A(y)U(s, y) + U(s, y)A^*(y) + B(y)U(s, y)B^*(y)\right)ds, \quad (3.13)$$

---

2 $\Re$ denotes the real part of a complex number

which is a matrix ordinary differential equation for U. We use the Kronecker product and equality (2.6) to re-write (3.13) as

$$U(t, y) = U_0 + \int_0^t \mathcal{M}_{A,B}(y)U(s, y)ds, \tag{3.14}$$

where $U_0 = \text{vec}(\widehat{u}_0\widehat{u}_0^*)$,

$$\mathcal{M}_{A,B}(y) = \overline{A(y)} \otimes I + I \otimes A(y) + \overline{B(y)} \otimes B(y), \tag{3.15}$$

and $I \in \mathbb{R}^{N \times N}$ is the identity matrix. In other words, Eq. (3.14) and Proposition 3.3 imply that the properties of the solution of the stochastic system (3.1) are determined by the exponential matrix

$$\exp\left(t\mathcal{M}_{A,B}(y)\right),$$

with $\mathcal{M}_{A,B}$ defined in (3.15). More precisely, if $f = g = 0$, then, according to (3.14),

$$U(t, y) = \exp\left(t\mathcal{M}_{A,B}(y)\right)U_0 \tag{3.16}$$

and therefore (3.8) holds if and only if

$$\| \exp\left(t\mathcal{M}_{A,B}(y)\right)\| \leq R(1 + |y|^2)^\beta, \quad 0 \leq t \leq T, \tag{3.17}$$

for some $R > 0$ and $\beta \geq 0$.

Theorem 3.4 *System* (3.1) *is well-posed if and only if* (3.17) *holds.*

Proof Necessity follows from the analysis of the homogeneous case. To establish sufficiency, we need an analogue of (3.16) for the inhomogeneous Eq. (3.7).

In a more compact differential form (c.f. (3.2)), (3.7) becomes

$$\dot{v} = Av + \widehat{f} + (Bv + \widehat{g})\dot{w}, \tag{3.18}$$

so that, with $V = V(t, y) = vv^*$, $\mathbf{V} = \mathbf{V}(t, y) = \text{vec}(V)$, $U = U(\text{t, y}) = \mathbb{E}\mathbf{V}$, and $\dot{}$ denoting derivative with respect to time,

$$\dot{v}^* = v^*A^* + \widehat{f}^* + (v^*B^* + \widehat{g}^*)\dot{w},$$
$$\dot{V} = AV + VA^* + \widehat{f}v^* + v\widehat{f}^* + (Bv + \widehat{g})(v^*B^* + \widehat{g}^*)$$
$$+ (BV + VB^* + \widehat{g}v^* + \widehat{v}g^*)\dot{w},$$
$$\dot{\mathbf{V}} = \mathcal{M}_{A,B}\mathbf{V} + \text{vec}\left(\widehat{f}v^* + v\widehat{f}^* + \widehat{g}v^*\text{B}^* + Bv\widehat{g}^* + \widehat{g}\widehat{g}^*\right) \tag{3.19}$$
$$+ \text{vec}\left(BV + VB^* + \widehat{g}v^* + v\widehat{g}^*\right)\dot{w},$$
$$\dot{U} = \mathcal{M}_{A,B}U + \text{vec}\left(\mathbb{E}\left(\widehat{f}v^* + v\widehat{f}^* + \widehat{g}v^*\text{B}^* + Bv\widehat{g}^* + \widehat{g}\widehat{g}^*\right)\right). \tag{3.20}$$

Using the notation

$$\Psi(t, y) = \exp\left(t\mathcal{M}_{A,B}(y)\right), \tag{3.21}$$

$$\boldsymbol{U}(t, y) = \Psi(t, y)\text{vec}\left(\widehat{\boldsymbol{u}}_0\widehat{\boldsymbol{u}}_0^*\right) + \int_0^t \Psi(t-s, y)\text{vec}\left(\mathbb{E}(\widehat{\boldsymbol{f}}\boldsymbol{v}^* + \boldsymbol{v}\widehat{\boldsymbol{f}}^*\right.$$

$$\left. +\widehat{\boldsymbol{g}}\boldsymbol{v}^*B^* + B\boldsymbol{v}\widehat{\boldsymbol{g}}^* + \widehat{\boldsymbol{g}}\widehat{\boldsymbol{g}}^*)\right)ds. \tag{3.22}$$

For every (column) vectors $\boldsymbol{h}, \boldsymbol{z} \in \mathbb{C}^N$,

$$|\text{vec}(\boldsymbol{h}\boldsymbol{z}^*)|^2 = \sum_{i,j=1}^N |h_i\bar{z}_j|^2 = |\boldsymbol{h}|^2 |\boldsymbol{z}|^2.$$

Then

$$|\boldsymbol{U}(t, y)| \leq \|\Psi(t, y)\| \, |\widehat{\boldsymbol{u}}_0|^2 + \int_0^t \|\Psi(t-s, y)\| \, \mathbb{E}\left(2|\widehat{\boldsymbol{f}}| \, |\boldsymbol{v}| + 2|\widehat{\boldsymbol{g}}| \, |B\boldsymbol{v}| + |\widehat{\boldsymbol{g}}|^2\right)ds. \tag{3.23}$$

To proceed, re-write (3.17) as

$$\|\Psi(t, y)\| \leq \rho_\beta(y), \quad \rho_\beta(y) = R(1 + |y|^2)^\beta, \quad R \geq 1, \tag{3.24}$$

use (3.12), and recall that $2|pq| \leq \epsilon|p|^2 + \epsilon^{-1}|q|^2, \epsilon > 0$. Then

$$2\rho_\beta \int_0^T \mathbb{E}\left(|\widehat{\boldsymbol{f}}| \, |\boldsymbol{v}|\right)ds \leq 2N^{1/2}\rho_\beta \int_0^T |\boldsymbol{U}(s, y)|^{1/2} \, |\widehat{\boldsymbol{f}}|ds$$

$$\leq 2N^{1/2}\rho_\beta \sup_{0<t<T} |\boldsymbol{U}(s, y)|^{1/2} \int_0^T |\widehat{\boldsymbol{f}}|ds$$

$$\leq N^{1/2}\rho_\beta \, \epsilon \sup_{0<t<T} |\boldsymbol{U}(s, y)| + N^{1/2}\rho_\beta \, \epsilon^{-1} \left(\int_0^T |\widehat{\boldsymbol{f}}|ds\right)^2$$

$$\leq N^{1/2}\rho_\beta \, \epsilon \sup_{0<t<T} |\boldsymbol{U}(s, y)| + N^{1/2}T \, \epsilon^{-1} \int_0^T \rho_\beta |\widehat{\boldsymbol{f}}(s, y)|^2 ds. \tag{3.25}$$

Similarly, since (2.2) implies

$$|B(y)\boldsymbol{v}| \leq C_B(1 + |y|^2)^{b/2} \, |\boldsymbol{v}|, \quad b \geq 0, \; C_B > 0, \tag{3.26}$$

we find

$$2\rho_\beta \int_0^T \mathbb{E}\big(|\widehat{\boldsymbol{g}}|\,|\boldsymbol{B}\boldsymbol{v}|\big)ds \le N^{1/2}\rho_\beta\,\epsilon \sup_{0<t<T}|\boldsymbol{U}(s,y)|$$

$$+ C_B^2 N^{1/2}\rho_\beta\,\epsilon^{-1}T\int_0^T \rho_b(y)|\widehat{\boldsymbol{g}}(s,y)|^2 ds. \qquad (3.27)$$

Now take $\epsilon = (4N^{1/2}\rho_\beta(y))^{-1}$. Inequality (3.23) becomes

$$\sup_{0<t<T}|\boldsymbol{U}(t,y)| \le 2\rho_\beta(y)|\widehat{\boldsymbol{u}}_0(y)|^2 + 8NTR\int_0^T \rho_{2\beta}(y)|\widehat{\boldsymbol{f}}(t,y)|^2 dt$$

$$+ 8(C_B^2 NTR^2 + 1)\int_0^T \rho_{2\beta+b}(y)|\widehat{\boldsymbol{g}}(t,y)|^2 dt; \qquad (3.28)$$

recall that we assume $\boldsymbol{u}_0$, $\boldsymbol{f}$, and $\boldsymbol{g}$ to be deterministic, but the above computations show that this assumption, while making the presentation easier, does not affect the final result.

The desired inequality (3.8) now follows from (3.28) and (3.12). In fact, keeping in mind that (3.8) implies $\boldsymbol{v} = \widehat{\boldsymbol{u}}$, (3.28) also establishes continuous dependence of the solution on the input:

$$\mathbb{E}\|\boldsymbol{u}\|_r^2(t) \le C\left(\|\boldsymbol{u}_0\|_{r+\beta}^2 + \int_0^t \|\boldsymbol{f}(s)\|_{r+2\beta}^2 ds + \int_0^t \|\boldsymbol{g}(s)\|_{r+2\beta+b}^2 ds\right), \quad t \in [0, T],$$
$$(3.29)$$

which is more informative than (3.5).

This completes the proof of Theorem 3.4.   □

To state condition (3.17) directly in terms of the matrix $\mathcal{M}_{A,B}(y)$, we need some constructions and facts from linear algebra.

Definition 3.5 The `spectral abscissa` $\boldsymbol{\alpha}(A)$ of a square matrix A is the largest real part of the eigenvalue of A:

$$\boldsymbol{\alpha}(A) = \max\{\Re\lambda : \lambda \text{ is an eigenvalue of } A\}$$

The following lemma connects the norm of the exponential of a matrix with the spectral abscissa of the matrix.

Lemma 3.6 *If $A \in \mathbb{C}^{n \times n}$ then*

$$e^{t\alpha(A)} \leq \|e^{tA}\| \leq \left(\sum_{k=1}^{n-1} \frac{(2\|A\|t)^k}{k!} \right) e^{t\alpha(A)}. \tag{3.30}$$

If $\alpha(A) = -\delta < 0$, then an alternative upper bound in (3.30) is

$$\|e^{tA}\| \leq C(n) \left(\frac{\|A\|}{\delta} \right)^{n-1} e^{-t\delta/2}. \tag{3.31}$$

If A has a complete system of eigenvectors, arranged in a matrix V, then there is a yet another upper bound

$$\|e^{tA}\| \leq \|V\| \cdot \|V^{-1}\| e^{t\alpha(A)}. \tag{3.32}$$

Proof For (3.31) and the upper bound in (3.30), see [9, Proposisions 1.3.2 and 1.3.3]. Analysis of the proof shows that the size n of the matrix in both (3.30) and (3.31) can be replaced with the degree of the minimal polynomial for A, which is always at most n. For the lower bound in (3.31), see [21, Theorem 15.3]. Inequality (3.32) follows from the equality

$$\exp(tA) = V \big[\text{diag} \big(\exp(t\lambda_i(A)) \big) \big] V^{-1}.$$

□

Proposition 3.7 *Condition* (3.17) *is equivalent to*

$$\alpha \big(\mathcal{M}_{A,B}(y) \big) \leq C_0 \ln(2 + |y|). \tag{3.33}$$

Proof If (3.17) holds, then the first inequality in (3.30) with $t = T$ implies (3.33) with

$$C_0 = \frac{2\beta}{T} + \frac{\ln R}{T \ln 2}.$$

If (3.33) holds, then (2.2) and the second inequality in (3.30) imply

$$\left\| \exp \left(t\mathcal{M}_{A,B}(y) \right) \right\| \leq C_1(T)(1 + |y|)^{C_2} \cdot (2 + |y|)^{C_0 T}$$

and (3.17) follows with $R = C_1(T)2^{C_0 T}$ and $\beta = C_2 + C_0 T$.   □

The most natural way to satisfy (3.33) is by having

$$\sup_y \alpha \big(\mathcal{M}_{A,B}(y) \big) \leq C_0. \tag{3.34}$$

A condition similar to (3.33) appears in the analysis of deterministic hyperbolic systems [14, Theorem 2.7]. If all elements of $A(y)$ and $B(y)$ are polynomials in y, then (3.33) and (3.34) are equivalent.

4 Regularity of solution

Theorem 3.4 provides the most basic information about system (3.1): if the input is in some Sobolev space, then the solution will be in some other Sobolev space, and condition (3.33) is necessary and sufficient for this to happen. Further information about the solution can be obtained with a more careful analysis of (3.19) and (3.20).

Our basic analysis of solvability of (3.1) assumes a fixed time interval $[0, T]$. If we allow T to change, then (3.17) becomes

$$\left\| \exp\left(T \mathcal{M}_{A,B}(y) \right) \right\| \leq R(T)(1 + |y|^2)^{\beta(T)}. \tag{4.1}$$

In (4.1) and below, $R = R(t)$ is a positive continuous function.

To begin, we define several classes of stochastic systems, depending on the behavior of the function $\left\| \exp\left(T \mathcal{M}_{A,B}(y) \right) \right\|$ for large T and $|y|$.

Definition 4.1 System (3.1) is called

(1) `pseudo-hyperbolic` of order (β, b) if (3.26) holds and (4.1) holds with β independent of T;
(2) `neutrally stable pseudo-hyperbolic` of order (β, b) if (4.1) holds with β independent of T and with a continuous uniformly bounded function $R = R(T)$:

$$\sup_{T \geq 0} R(T) \leq R_0; \tag{4.2}$$

(3) `stable pseudo-hyperbolic` of order (β, b) if (4.1) holds with β independent of T and with a continuous exponentially decaying function $R = R(T)$:

$$R(T) \leq R_0 e^{-\phi T}, \quad \phi > 0; \tag{4.3}$$

(4) `parabolic` of order $2p$ if there exists a positive number p such that

$$0 < \lim_{|y| \to \infty} \frac{\|A(y)\| + \|B^*(y)B(y)\|}{|y|^{2p}} < \infty; \tag{4.4}$$

$$\boldsymbol{\alpha}\left(\mathcal{M}_{A,B}(y) \right) \leq -\varepsilon |y|^{2p} + L, \quad \varepsilon > 0, \quad L \in \mathbb{R}; \tag{4.5}$$

(5) (q, β, b)-`dissipative`, if (3.26) holds and

$$\left\| \exp\left(t \mathcal{M}_{A,B}(y) \right) \right\| \leq R(t)(1 + |y|^2)^{\beta} e^{-t\varepsilon|y|^{2q}}, \quad q > 0, \ \beta \geq 0. \tag{4.6}$$

A parabolic system of order $2p$ is $(p, 0, p)$-dissipative. Indeed, (4.4) implies (3.26) with $b = p$. Also, if (4.4) and (4.5) hold, then

$$\|\mathcal{M}_{A,B}(y)\| \leq C_M (1 + |y|^2)^p,$$

and (3.31) with $n = N^2$ becomes

$$\left\| \exp\left(t\mathcal{M}_{A,B}(y)\right) \right\| \le e^{(L+\varepsilon)t}(C_M/\varepsilon)^{N^2-1}C(N^2)e^{-t(\varepsilon/2)|y|^{2p}}. \tag{4.7}$$

The reason for the name "pseudo-hyperbolic" is that certain hyperbolic systems satisfy a condition similar to (3.33), even though system (3.1) is usually not hyperbolic. If the matrices $A(y)$ and $B(y)$ are real, then conditions for stability can be derived using the results from [11, Sect. 6.10]: stability of the PDE (3.1) is equivalent to the mean-square stability of the ODE (3.7) uniformly in y.

Conditions (4.4) and (4.5) can be considered a stochastic analogue of parabolicity in the sense of Petrowski, as defined, for example, in [7, Sect. 9.1]. This connection with the classical deterministic setting is the reason for using $2p$ rather than p in (4.4) and (4.5), even though p no longer has to be an integer: Eq. (3.14) describes the square of the solution rather than solution itself. For the same reason, in (4.6) we have $2q$ and β rather than q and $\beta/2$.

Consider the equation

$$u_{tt} - au_{txx} - c^2 u_{xx} = \sigma u_{xx}\dot{w}, \quad x \in \mathbb{R}, \tag{4.8}$$

first encountered in Introduction. The equivalent system formulation is

$$u_t = v, \ v_t = c^2 u_{xx} + av_{xx} + \sigma u_{xx}\dot{w}. \tag{4.9}$$

We have $d = 1$, $N = 2$, and

$$A(y) = \begin{pmatrix} 0 & 1 \\ -c^2 y^2 & -ay^2 \end{pmatrix}, \ B(y) = \begin{pmatrix} 0 & 0 \\ \sigma y^2 & 0 \end{pmatrix}, \tag{4.10}$$

$$\mathcal{M}_{A,B}(y) = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -c^2 y^2 & -ay^2 & 0 & 1 \\ -c^2 y^2 & 0 & -ay^2 & 1 \\ \sigma^2 y^4 & -c^2 y^2 & -c^2 y^2 & -2ay^2 \end{pmatrix}. \tag{4.11}$$

Analysis of the matrix $\mathcal{M}_{A,B}(y)$ can be simplified by observing that there is an eigenvector $(0\ 1\ -1\ 0)^\top$ with eigenvalue $\lambda_0 = -ay^2$. All other eigenvectors of the matrix have the form $(h_1\ h_2\ h_2\ h_3)^\top$, and the vector $(h_1\ h_2\ h_3)^\top$ is an eigenvector of the matrix

$$\widetilde{\mathcal{M}}_{A,B}(y) = \begin{pmatrix} 0 & 2 & 0 \\ -c^2 y^2 & -ay^2 & 1 \\ \sigma^2 y^4 & -2c^2 y^2 & -2ay^2 \end{pmatrix}, \tag{4.12}$$

corresponding to the same eigenvalue. We will see in Sect. 5 that transition from (4.11) to (4.12) is a particular case of the reduction possible for many second-order systems.

We conclude that the remaining three eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of $\mathcal{M}_{A,B}(y)$ are the roots of

$$\lambda^3 + 3ay^2\lambda^2 + (4c^2y^2 + 2a^2y^4)\lambda + 2(2ac^2 - \sigma^2)y^4 = 0.$$

Straightforward perturbation analysis shows that, as $|y| \to \infty$, these eigenvalues satisfy

$$\lambda_1 = -\frac{2ac^2 - \sigma^2}{a^2} + o(1), \quad \lambda_2 = -ay^2 - \frac{2\sigma^2}{a^2} + o(1),$$

$$\lambda_3 = -2ay^2 + \frac{2ac^2 + \sigma^2}{a^2} + o(1).$$

Similar computations provide asymptotic of the eigenvectors, and we conclude that, as long as $a > 0$,

$$\left\| \exp\left(t\mathcal{M}_{A,B} \right) \right\| \le C(1 + |y|^2)e^{-\phi t}, \quad \phi = \frac{2ac^2 - \sigma^2}{a^2}.$$

Then system (4.9) is

- pseudo-hyperbolic of order (1,2) if $a > 0$;
- stable pseudo-hyperbolic of order (1,2) if $2ac^2 > \sigma^2$ (cf. [11, Sect. 6.10]);
- neutrally stable pseudo-hyperbolic of order (1,2) if $2ac^2 = \sigma^2$.
  To illustrate the difference between parabolic and dissipative systems, consider

$$u_t = u_{xx}, \quad v_t = 2v_{xx} + \sigma(-\nabla^2)^{\beta/2}u\dot{w}, \quad x \in \mathbb{R}, \quad \beta \ge 0, \quad \sigma > 0. \quad (4.13)$$

We have $d = 1, \; N = 2$, and

$$A(y) = \begin{pmatrix} -y^2 & 0 \\ 0 & -2y^2 \end{pmatrix}, \quad B(y) = \begin{pmatrix} 0 & 0 \\ \sigma|y|^\beta & 0 \end{pmatrix}, \quad (4.14)$$

$$\mathcal{M}_{A,B}(y) = \begin{pmatrix} -2y^2 & 0 & 0 & 0 \\ 0 & -3y^2 & 0 & 0 \\ 0 & 0 & -3y^2 & 0 \\ \sigma^2|y|^{2\beta} & 0 & 0 & -4y^2 \end{pmatrix}. \quad (4.15)$$

For Eq. (1.1) in the introduction, the matrix $\mathcal{M}_{A,B}(y)$ corresponds to (4.15) with $\beta = 1$. The matrix $\exp\left(t\mathcal{M}_{A,B}(y) \right)$ can be easily computed, and it follows that system (4.13) is

- parabolic of order 2 if $\beta \le 1$;
- $(1, \beta - 1, \beta)$-dissipative if $\beta > 1$.

In this example, the coefficient σ has no influence on the type of the system. The possibility to have $\beta > 1$ illustrates the importance of (4.4) as part of the stochastic parabolicity condition; it also shows that, unlike a single equation, the order of the operator B in a well-posed system can be arbitrarily larger than the order of A.

The following definition refines Definition 3.1.

Definition 4.2 System (3.1) is called (ℓ, p, q)-well posed, with $\ell, p, q \in \mathbb{R}$, if, for every $r \in \mathbb{R}$, $T > 0$, and the input data satisfying

$$\boldsymbol{u}_0 \in H^{r+\ell}, \ \boldsymbol{f} \in L_2((0, T), H^{r+p}), \ g \in L_2((0, T); H^{r+q}), \qquad (4.16)$$

there exists a unique solution $\boldsymbol{u}$ such that $\boldsymbol{u}(t) \in L_2(\Omega; H^r), t \in (0, T)$ and

$$\mathbb{E}\|\boldsymbol{u}(t)\|_r^2 \le C(t)\left(\|\boldsymbol{u}_0\|_{r+\ell}^2 + \int_0^t \|\boldsymbol{f}(s)\|_{r+p}^2 ds + \int_0^t \|g(s)\|_{r+q}^2 ds\right). \quad (4.17)$$

If, in addition, $\sup_{t>0} C(t) < \infty$, then (3.1) is called (ℓ, p, q)-stable.

The two theorems below connect Definitions 4.1 and 4.2.

Theorem 4.3 *A pseudo-hyperbolic system of order (β, b) is $(\beta, 2\beta, 2\beta + b)$-well posed; a stable pseudo-hyperbolic system of order (β, b) is $(\beta, 2\beta, 2\beta + b)$-stable.*

Proof Well-posedness follows from (3.29).

To establish stability, we need a more delicate analysis of the right-hand-side of (3.23). Accordingly, we write

$$\|\Psi(t, y)\| \le R(t)\rho_\beta(y), \ \ \rho_\beta(y) = (1 + |y|^2)^\beta \qquad (4.18)$$

(not to be confused with (3.24)), and then modify (3.25) to

$$\rho_\beta \int_0^T 2R(T-s)\mathbb{E}(|\widehat{\boldsymbol{f}}|\,|\boldsymbol{v}|)ds \le 2N^{1/2}\rho_\beta \int_0^T R(T-s)|U(s, y)|^{1/2}\,|\widehat{\boldsymbol{f}}|ds$$

$$\le 2N^{1/2}\rho_\beta \sup_{0<t<T} |U(s, y)|^{1/2} \int_0^T R(T-s)|\widehat{\boldsymbol{f}}|ds$$

$$\le \rho_\beta N^{1/2}\epsilon \sup_{0<t<T} |U(s, y)| + N^{1/2}\epsilon^{-1}\left(\int_0^T R(T-s)\rho_\beta|\widehat{\boldsymbol{f}}|ds\right)^2$$

$$\le \rho_\beta N^{1/2}\epsilon \sup_{0<t<T} |U(s, y)| + N^{1/2}\epsilon^{-1}\int_0^T R(s)ds \int_0^T \rho_\beta R(T-s)|\widehat{\boldsymbol{f}}(s, y)|^2 ds.$$

$$(4.19)$$

After a similar modification of (3.27), (3.28) becomes

$$
\sup_{0<t<T} |U(t, y)| \le 2R_0 \rho_\beta(y)|\widehat{\boldsymbol{u}}_0(y)|^2
$$

$$
+2 \left(\int_0^T R(t)dt \right) \left(4N \int_0^T \rho_{2\beta}(y) R(T-s)|\widehat{\boldsymbol{f}}(t, y)|^2 dt \right. \tag{4.20}
$$

$$
\left. +4(C_B^2 N + 1) \int_0^T R(T-s)\rho_{2\beta+b}(y)|\widehat{\boldsymbol{g}}(t, y)|^2 dt \right);
$$

Condition (4.3) implies

$$
\int_0^\infty R(t)dt \le R_0/\phi,
$$

and then the number $C = C(t)$ in (3.29) satisfies $C(t) \le 8N(C_B+1) \max(R_0, R_0/\phi)$. *This completes the proof of Theorem 4.3.* □

Remark 4.4 If (3.1) is neutrally stable pseudo-hyperbolic of order (β, b) and if $\boldsymbol{f} = \boldsymbol{g} = 0$, then (3.28) implies

$$
\mathbb{E}\|\boldsymbol{u}(t)\|_r^2 \le 2R_0\|\boldsymbol{u}_0\|_{r+\beta}^2.
$$

While a (q, β, b)-dissipative system is pseudo-hyperbolic of order (β, b), the additional exponentially decaying factor in (4.6) leads to additional regularity of the solution.

Theorem 4.5 *A (q, β, b)-dissipative equation is $(0, 2\beta - q, 2\beta + b - q)$-well posed.*

Proof Combining (4.6) and (4.18) results in

$$
\|\Psi(t, y)\| \le R(t)\rho_\beta(y)e^{-t\varepsilon(1+|y|^2)^q}.
$$

Since

$$
\rho_\beta(y) \int_0^T e^{-\varepsilon t(1+|y|^2)^q} dt \le C(T, \varepsilon, \beta, q)\rho_{\beta-q}(y),
$$

the result follows from a computation similar to (4.20). □

Corollary 4.6 *A parabolic system of order $2p$ is $(0, -p, 0)$-well posed.*

Proof By (4.4) and (4.7), a parabolic system of order $2p$ is $(p, 0, p)$-dissipative. □

A single stochastic parabolic equation of order $2p$ is certainly $(0, -p, 0)$-well posed, but also has additional regularity properties. It is natural to expect this additional regularity for parabolic systems as well.

Theorem 4.7 *If system* (3.1) *is parabolic of order* $2p$, *then, for every* $r \in \mathbb{R}$, $T > 0$, *and the input data satisfying*

$$\boldsymbol{u}_0 \in H^r, \ \boldsymbol{f} \in L_2((0, T), H^{r-p}), \ \boldsymbol{g} \in L_2((0, T); H^r), \tag{4.21}$$

the solution is an element of $L_2\big(\Omega \times (0, T); H^{r+p}\big) \bigcap L_2\big(\Omega; \mathcal{C}((0, T); H^r)\big)$ *and*

$$\mathbb{E} \sup_{0<t<T} \|\boldsymbol{u}(t)\|_r^2 + \int_0^T \mathbb{E}\|\boldsymbol{u}(t)\|_{r+p}^2 dt$$

$$\leq C(T)\bigg(\|\boldsymbol{u}_0\|_r^2 + \int_0^T \|\boldsymbol{f}(s)\|_{r-p}^2 ds + \int_0^T \|\boldsymbol{g}(s)\|_r^2 ds\bigg). \tag{4.22}$$

Proof Since the operators in (3.1) have constant coefficients, it is enough to establish (4.22) for $r = 0$. Continuity follows in the usual way by showing weak continuity and continuity of the norm. It remains to show that (4.7) implies

$$\int_0^T (1 + |y|^2)^p |U(t, y)| dt$$

$$\leq C(T)\bigg(|\widehat{\boldsymbol{u}}_0(y)|^2 + \int_0^T (1 + |y|^2)^{-p} |\widehat{\boldsymbol{f}}(s, y)|^2 ds + \int_0^T |\widehat{\boldsymbol{g}}(s, y)|^2 ds\bigg) \tag{4.23}$$

and

$$\sup_{0<t<T} \int_{\mathbb{R}^d} |U(t, y)| dy$$

$$\leq C(T) \int_{\mathbb{R}^d} \bigg(|\widehat{\boldsymbol{u}}_0(y)|^2 + \int_0^T (1+|y|^2)^{-p} |\widehat{\boldsymbol{f}}(s, y)|^2 ds + \int_0^T |\widehat{\boldsymbol{g}}(s, y)|^2 ds\bigg) dy. \tag{4.24}$$

While the arguments are identical to the Fourier-analytic proof of the similar results for the one-dimensional heat equation, some explanations might be necessary because, for the single equation, the proof is usually carried out directly in the physical space using integration by parts.

With the notations from the proof of Theorem 4.3, and with a suitable function $R(t)$, (4.7) becomes

$$\|\Psi(t, y)\| \leq R(t)e^{-(t\varepsilon/2)(1+|y|^2)^p}. \tag{4.25}$$

Define

$$\psi(y) = \frac{\varepsilon(1 + |y|^2)^p}{2}.$$

To establish (4.23), we re-write (4.20) [taking $\beta = 0$, $b = p$, and replacing $R(t)$ with $Ce^{-t\psi(y)}$] as

$$|U(t, y)| \leq 2Ce^{-t\psi(y)}|\widehat{u}_0|^2 + \frac{2C}{\psi(y)}\int_0^t e^{-(t-s)\psi(y)}|\widehat{f}(s, y)|^2 dt$$

$$+ \frac{2C(1 + C_B^2)}{\psi(y)}\int_0^t \psi(y)e^{-(t-s)\psi(y)}|\widehat{g}(t, y)|^2 dt. \qquad (4.26)$$

Inequality (4.23) now follows after multiplying both sides of (4.26) by $\psi(y)$, integrating both sides in time from 0 to T and changing the order of integration on the right-hand side. Note that

$$\int_s^T e^{-(t-s)\psi(y)} dt = \frac{1 - e^{-(T-s)\psi(y)}}{\psi(y)} \leq \frac{1}{\psi(y)}.$$

To establish (4.24), we go back to (3.19) and estimate the stochastic integral using the Burkholder-Davis-Gundy inequality. The argument is identical to the derivation of (3.28). An interested reader should be able to fill in the details.

This completes the proof of Theorem 4.7.   □

Remark 4.8 Analysis of the proof shows that a version of (4.22) can be derived for a (q, β, b) dissipative equation:

$$\mathbb{E} \sup_{0 < t < T} \|u(t)\|_r^2 + \int_0^T \mathbb{E}\|u(t)\|_{r+q}^2 dt$$

$$\leq C(T)\left(\|u_0\|_{r+\beta}^2 + \int_0^T \|f(s)\|_{r+2\beta-q}^2 ds + \int_0^T \|g(s)\|_{r+2\beta+b-q}^2 ds\right).$$

Again, an interested reader can fill in the details.

5 Further developments

To begin, let us discuss connections between the results of this paper and some of the existing results.

A single equation. If $N = 1$, then (3.17) becomes

$$e^{t(2A(y)-B^2(y))} \leq R(1 + |y|^2)^\beta, \ t \in [0, T].$$

In particular, for partial differential operators A, B, the condition on the corresponding symbols is

$$2\Re(A(y)) - B^2(y) \le C_0;$$

equations satisfying this condition are called *degenerate parabolic* (although terminology can vary) and indeed form the largest class of well-posed equations [19, Sect. IV.2].

A system of deterministic equations. By taking $B(y) = 0$, we should be able to recover the results for the deterministic equations, for example, the Petrowski parabolicity. In the deterministic case, the conditions do not involve the Kronecker product and are stated in terms of $\boldsymbol{\alpha}(A(y))$. The reason is that, when $B(y) = 0$, the matrix

$$\mathcal{M}_{A,0}(y) = \overline{A(y)} \otimes I + I \otimes A(y),$$

known as the Kronecker sum of $\overline{A}$ and A, is much easier to study. In particular, it is known [10, Theorem 4.4.5] that all eigenvalues of $\mathcal{M}_{A,0}(y)$ are of the form $\overline{\lambda_i(A)} + \lambda_j(A), i, j = 1, \ldots, N$, where $\lambda_i(A)$ is an eigenvalue of $A(y)$. In particular, $\boldsymbol{\alpha}(\mathcal{M}_{A,0}(y)) = 2\boldsymbol{\alpha}(A(y))$. In the deterministic setting, the Petrowski parabolicity condition is

$$\boldsymbol{\alpha}(A(y)) \le -\varepsilon|y|^{2p} + L,$$

which is equivalent to (4.5) with $B = 0$.

A sufficient condition for parabolicity. Let us return to (3.11). Define the matrix

$$\mathcal{N}_{A,B}(y) = A^*(y) + A(y) + B^*(y)B(y).$$

Since $\mathcal{N}_{A,B}$ is Hermitian,

$$\boldsymbol{v}^*(t, y)\mathcal{N}_{A,B}(y)\boldsymbol{v}(t, y) \le \boldsymbol{\alpha}(\mathcal{N}_{A,B}(y)) \, \boldsymbol{v}^*(t, y)\boldsymbol{v}(t, y).$$

If we assume that

$$\boldsymbol{\alpha}(\mathcal{N}_{A,B}(y)) \le -\varepsilon|y|^{2p} + L, \quad \varepsilon > 0, \tag{5.1}$$

then (3.11) and the Gronwall inequality imply

$$\mathfrak{v}(t, y) \le |\widehat{\boldsymbol{u}}_0(y)|^2 \, e^{Lt} e^{-t\varepsilon|y|^{2p}}, \quad \mathfrak{v}(t, y) = \boldsymbol{v}^*(t, y)\boldsymbol{v}(t, y).$$

By Proposition 3.3, this is enough to establish well-posedness of (3.1), and in fact, implies that the system is parabolic of order $2p$. Indeed, for systems of second-order stochastic parabolic equations ($p = 1$), condition (5.1) is a slightly weaker version of the parabolicity condition from [12]. Our analysis shows that (5.1) is too restrictive and can be replaced with a weaker condition (4.5) (to have parabolicity) or even with (3.33) (to have well-posedness). In fact, it is known [18] that

$$2\boldsymbol{\alpha}(A) \le \boldsymbol{\alpha}(A + A^*)$$

for every matrix A; equality holds if $AA^* = A^*A$; for more on the subject, see [20]. In other words,

$$\boldsymbol{\alpha}(\mathcal{M}_{A,0}(y)) \leq \boldsymbol{\alpha}(\mathcal{N}_{A,0}(y)),$$

meaning that condition (5.1) is not optimal even in the deterministic case.

Next, let us take a closer look at our results.

Verification of condition (3.33). The location of eigenvalues of a matrix can be analyzed using the Routh-Hurwitz criterion, a necessary and sufficient condition, in terms of the coefficients of a real polynomial, for the roots of the polynomial to have negative real parts. Although a classical result, its application to (3.33) requires two modifications: (a) The region of interest is not exactly the left half-plane; this is resolved with a suitable shift of the variable in the polynomial; (b) The coefficients of the polynomial are complex rather than real; a suitable modification of the criterion exists for this case as well [15, Theorem 40.1].

A complete description of the procedure is rather long, technical, and mostly irrelevant to the main subject of this paper. Below are the main ideas.

Suppose that $P(\lambda)$ is the characteristic polynomial of the matrix $\mathcal{M}_{A,B}$; it is a polynomial of order N^2 with coefficients depending on the parameter y. Then (3.33) means that there exists a $C > 0$ such that, for all $y \in \mathbb{R}^d$, all roots of the equation

$$P(\lambda + C \ln(2 + |y|)) = 0$$

have negative real parts.

If the coefficients of P are real, then Routh-Hurwitz criterion requires analysis of N^2 determinants of orders $1, 2, 3, \ldots N^2$, each constructed using the coefficients of P. If the coefficients of P are complex, then the Routh-Hurwitz criterion requires analysis of N^2 determinants of orders $2, 4, 6, \ldots 2N^2$, each constructed using the real and imaginary parts of the coefficients of P. Presence of the parameter y makes the complexity of the computations formidable indeed.

On the one hand, for a general polynomial, the Routh-Hurwitz criterion is, in a certain sense, optimal [2, Theorem 1].

On the other hand, the special structure of the matrix $\mathcal{M}_{A,B}$ does allow for certain simplifications. The general result is as follows: the matrix $\mathcal{M}_{A,B}$ has $N(N+1)/2$ eigenvectors of the form vec(H) for a Hermitian matrix H, and $(N-1)N/2$ eigenvectors of the form vec(S) for a skew-Hermitian matrix S [6, Sect. 3]. Let us see how this result can indeed simplify the analysis of second-order systems with real matrices A, B. Direct [straightforward and not especially insightful] computations show that, for real 2-by-2 matrices A, B, the matrix $\mathcal{M}_{A,B}$ has an eigenvector $(0 \ 1 \ -1 \ 0)^{\top}$ corresponding to the eigenvalue trace$(A) + \det(B)$. All other eigenvectors of $\mathcal{M}_{A,B}$ are of the form $(h_1 \ h_2, \ h_2 \ h_3)^{\top}$, so that $(h_1 \ h_2 \ h_3)^{\top}$ is an eigenvector of the matrix

$$\begin{pmatrix} 2A_{11} + B_{11}^2 & 2(A_{12} + B_{11}B_{12}) & B_{12}^2 \\ A_{21} + B_{11}B_{21} & A_{11} + A_{22} + B_{11}B_{22} + B_{12}B_{21} & B_{12}B_{22} + A_{12} \\ B_{21}^2 & 2(A_{21} + B_{21}B_{22}) & 2A_{22} + B_{22}^2 \end{pmatrix}.$$

Reduction of a 4-by-4 matrix (4.11) to a 3-by-3 matrix (4.12) during the analysis of (4.8) provides an illustration of this idea.

Optimality of the results. While condition (3.17) is necessary and sufficient for solvability of the (3.1), a regularity result such as (3.29) is only as good as the estimate (3.17). Lemma 3.6 provides some information, but the resulting bounds are not always optimal. For example, for the (relatively simple) matrix (4.15) in the case $\beta > 1$, (3.30) gives

$$\left\| e^{t\mathcal{M}_{A,B}(y)} \right\| \leq C\big((1+|y|^2)^{\beta-1}\big)^3 e^{-2y^2},$$

(3.32) gives

$$\left\| e^{t\mathcal{M}_{A,B}(y)} \right\| \leq C\big((1+|y|^2)^{\beta-1}\big)^2 e^{-2y^2},$$

and a direct computation gives

$$\left\| e^{t\mathcal{M}_{A,B}(y)} \right\| \leq C\big((1+|y|^2)^{\beta-1}\big) e^{-2y^2}.$$

One reason parabolic systems are of special interest is that (4.7) all but eliminates the need to bound the exponential matrix, leading to essentially optimal regularity result (4.22).

In general, getting optimal bounds on the exponential of a non-normal matrix is a difficult problem [16,17,21, etc.]. For a normal matrix A (that is, $A^*A = AA^*$), the corresponding bound is very simple:

$$\|e^{tA}\| = e^{t\boldsymbol{\alpha}(A)};$$

see [21, Theorem 15.3]. As a result, majority of the current research on the subject is about efficient computation of e^{tA} for normal (in fact, Hermitian) matrices A [3,22, etc.] In this connection, note that if matrices A and B are Hermitian, then so is $\mathcal{M}_{A,B}$.

A related question is possibility of considering truly vector Sobolev space, with an option of having different regularity for different components of the vector function. At this point, the answer is not clear, but Eq. (4.13) provides a motivation: if $\beta \neq 1$, then it is possible to take $u_0 \in H^{r+\beta-1}$ and $v_0 \in H^r$ so that

$$u \in \mathcal{C}((0,T); H^{r+\beta-1}) \bigcap L_2((0,T); H^{r+\beta}),$$

and then $Bu \in L_2((0,T); H^r)$, so that

$$v \in L_2(\Omega; \mathcal{C}((0,T); H^r) \bigcap L_2((0,T); H^{r+1}).$$

According to Remark 4.8, if $u_0, v_0 \in H^{r+1}$, then

$$u, v \in L_2(\Omega; \mathcal{C}((0,T); H^r) \bigcap L_2((0,T); H^{r+1}),$$

which is, technically, neither weaker nor stronger, but just different.

Generality of the results. System (3.1) was intentionally chosen the simplest possible, to make sure that the key ideas are not lost in the computations. Analysis of the proof of Theorem 3.4 suggests immediate generalizations of (3.29) in the following directions:

(1) NUMBER OF BROWNIAN MOTIONS. The term $\sum_{k \geq 1} \mathrm{B}_k u d w_k(t)$ can replace the term $\mathrm{B} u d w(t)$ as long as $\sum_k \|B_k(y)\|^2 \leq C_B (1 + |y|^2)^{b/2}$. The matrix $\mathcal{M}_{A,B}$ becomes

$$\mathcal{M}_{A,B}(y) = \overline{A(y)} \otimes I + A(y) \otimes I + \sum_k \overline{B_k(y)} \otimes B_k(y);$$

the infinite sum of Kronecker products is well-defined, because $\|\overline{B_k(y)} \otimes B_k(y)\| \leq C(N)\|B_k(y)\|^2$.

(2) INPUT DATA. There are no significant difficulties extending the results to

$$\boldsymbol{u}_0 \in L_2(\Omega, \mathcal{F}_0; H^{r+\beta}), \ \boldsymbol{f} \in L_2(\Omega \times (0,T), \mathcal{F}_t; H^{r+2\beta}),$$

$$\boldsymbol{g}_k \in L_2(\Omega \times (0,T), \mathcal{F}_t; H^{r+2\beta+b}), \ \sum_k \mathbb{E} \int_0^T \|\boldsymbol{g}_k(t)\|_{r+2\beta+b}^2 dt < \infty.$$

Some technical issues arise in connection with (3.9), which can now contain an anticipating stochastic integral, and a different approach (e.g using rough path theory) might be necessary to deal with exceptional sets during the construction of $\boldsymbol{v}$.

Further generalizations are more technically demanding:

(1) TIME- DEPENDENT OPERATORS A AND B_k. Now there is no matrix exponential to consider; instead, we need to study the fundamental solution of the corresponding ODE with matrix $\mathcal{M}_{A,B}(t, y)$. Similar to the deterministic setting [7, Sect. 9.2], this difficulty can be resolved by freezing the time variable in the operators.

(2) SPACE- DEPENDENT OPERATORS. The symbol of the operator becomes a function of t, x, y, and the Fourier transform method no longer works. The analysis should still be possible with the help of standard PDE techniques such as localization (freezing the coefficients and partition of unity) and the method of continuity.

(3) RANDOM PREDICTABLE OPERATORS. This generalization presents the most challenge because now there seems to be no way around anticipating stochastic integrals. While anticipating integration can be avoided for scalar parabolic equations [13] (and then, under a more restrictive condition (5.1), for systems [12]), the method does not seem to work under a more general condition (3.33). The approach based on anticipating stochastic calculus is possible for one equation [1] and could probably be extended to systems, at least for second-order in space under the parabolicity condition

$$\sup_{\omega,t,x} \boldsymbol{\alpha}\big(\mathcal{M}_{A,B}(\omega, t, x; y)\big) \leq -\varepsilon|y|^2 + L.$$

6 Conclusions

In the analysis of stochastic evolution systems, the key object is the covariance matrix of the solution rather than the norm of the solution. When written in a matrix-vector form, the matrix differential equation satisfied by the covariance leads to a linear system of ordinary differential equations; the matrix in the equation is a sum of certain Kronecker products of the matrices from the original system. In other words, instead of the "obvious" characteristic matrix

$$\mathcal{N}_{A,B} = A + A^* + B^* B,$$

the "correct" characteristic matrix of the system is

$$\mathcal{M}_{A,B} = \overline{A} \otimes I + I \otimes A + \overline{B} \otimes B.$$

While dimension of the matrix increases from N-by-N to N^2-by-N^2, and integration by parts no longer works, the pay-off is a sharp condition for well-posedness of the system. For parabolic systems, the condition is an extension of the Petrowski parabolicity.

References

1. Alòs, Elisa, León, Jorge A., Nualart, David: Stochastic heat equation with random coefficients. Probab. Theory Relat. Fields **115**(1), 41–94 (1999)
2. Anderson, B.D.O., Jury, E.I.: A "simplest possible" property of the generalized Routh-Hurwitz conditions. SIAM J. Control Optim. **15**(1), 177–184 (1977)
3. Botchev, M.A., Grimm, V., Hochbruck, M.: Residual, restarting, and Richardson iteration for the matrix exponential. SIAM J. Sci. Comput. **35**(3), A1376–A1397 (2013)
4. Brewer, J.W.: Kronecker products and matrix calculus in system theory. IEEE Trans. Circuits Syst. **25**(9), 772–781 (1978)
5. Egorov, YuV, Shubin, M.A.: Foundations of the Classical Theory of Partial Differential Equations. Springer, Berlin (1998)
6. Feng, J., Lam, J., Wei, Y.: Spectral properties of sums of certain Kronecker products. Linear Algebra Appl. **431**(9), 1691–1701 (2009)
7. Friedman, A.: Partial Differential Equations of Parabolic Type. Prentice-Hall, Englewood Cliffs (1964)
8. Gīhman, I.I., Skorohod, A.V.: Stochastic differential equations. Springer, New York (1972)
9. Godunov, S.K.: Ordinary Differential Equations with Constant Coefficient, Translations of Mathematical Monographs, vol. 169. American Mathematical Society, Providence (1997)
10. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1994)
11. Khasminskii, R.: Stochastic stability of differential equations. Stochastic Modelling and Applied Probability, vol. 66, 2nd edn. Springer, Heidelberg (2012)
12. Kim, K.-H., Lee, K.: A W_2^n-theory of stochastic parabolic partial differential systems on C^1-domains. Potential Anal. **38**(3), 951–984 (2013)
13. Krylov, N.V.: An analytic approach to SPDEs, stochastic partial differential equations: six perspectives. Math. Surv. Monogr. Am. Math. Soc. **64**, 185–242 (1999)
14. Lax, P.D.: Hyperbolic partial differential equations, Courant Lecture Notes in Mathematics, vol. 14. New York University, Courant Institute of Mathematical Sciences, New York (2006)
15. Marden, M.: Geometry of polynomials. Mathematical Surveys. American Mathematical Society, Providence (1996)
16. Nechepurenko, YuM: On a bound for the norm of a matrix exponential. Dokl. Akad. Nauk **377**(5), 597–600 (2001)

17. Nechepurenko, YuM: An estimate for the norm of a matrix exponential in terms of the solution of the Lyapunov equation and the bounds of the Hausdorff set of the matrix. Zh. Vychisl. Mat. Mat. Fiz. **42**(2), 131–141 (2002)
18. Pao, C.V.: Logarithmic derivates of a square matrix. Linear Algebra Appl. **6**, 159–164 (1973)
19. Rozovskiĭ, B.L.: Stochastic evolution systems. Mathematics and its Applications (Soviet Series), vol. 35. Kluwer Academic, Dordrecht (1990)
20. Ström, T.: On logarithmic norms. SIAM J. Numer. Anal. **12**(5), 741–753 (1975)
21. Trefethen, L.N., Embree, M.: Spectra and Pseudospectra. Princeton University Press, Princeton (2005)
22. Ye, Q.: Error bounds for the Lanczos methods for approximating matrix exponentials. SIAM J. Numer. Anal. **51**(1), 68–87 (2013)